



UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO  
PROGRAMA DE MAESTRÍA Y DOCTORADO EN FILOSOFÍA  
FACULTAD DE FILOSOFÍA Y LETRAS  
INSTITUTO DE INVESTIGACIONES FILOSÓFICAS  
Área: Epistemología

NECESIDAD, ESENCIA Y REPRESENTACIÓN:  
Metafísica y Epistemología de las Esencias y la Modalidad

T E S I S  
QUE PARA OPTAR POR EL GRADO DE  
MAESTRO EN FILOSOFÍA

PRESENTA:  
Carlos Alberto Romero Castillo

TUTORA:  
DRA. LOURDES VALDIVIA DOUNCE (FFyL, UNAM)

México, D.F., Agosto de 2014



Universidad Nacional  
Autónoma de México

Dirección General de Bibliotecas de la UNAM

**Biblioteca Central**



**UNAM – Dirección General de Bibliotecas**  
**Tesis Digitales**  
**Restricciones de uso**

**DERECHOS RESERVADOS ©**  
**PROHIBIDA SU REPRODUCCIÓN TOTAL O PARCIAL**

Todo el material contenido en esta tesis esta protegido por la Ley Federal del Derecho de Autor (LFDA) de los Estados Unidos Mexicanos (México).

El uso de imágenes, fragmentos de videos, y demás material que sea objeto de protección de los derechos de autor, será exclusivamente para fines educativos e informativos y deberá citar la fuente donde la obtuvo mencionando el autor o autores. Cualquier uso distinto como el lucro, reproducción, edición o modificación, será perseguido y sancionado por el respectivo titular de los Derechos de Autor.

## CONTENIDOS

Resumen de la tesis	4
Convenciones	6
Agradecimientos	7
0. Introducción: motivación de la investigación	9
0.1. Comentario metodológico	11
I. Metafísica: Modalidad, Esencias y Fundamentalidad	
1. La concepción estándar de la modalidad y la esencialidad	14
1.1. Aspectos básicos de la semántica estándar para la lógica modal cuantificada	14
1.2. La semántica estándar: teoremas de correspondencia, definibilidad y caracterización	16
1.3. Interpretando filosóficamente a la semántica modal: <i>mundismo</i> y <i>ersatzismo</i>	24
1.4. Esencia y modalidad	29
2. Esencialismo, modalidad y reducción	33
2.1. Reducción y fundamentación	34
2.2. Teorías “no-no” de la modalidad	38
2.3. Problemas con las teorías no-no de la modalidad	42
2.3.1 Problemas con el ersatzismo lingüístico	42
2.3.2. Generalizando, y el caso de la teoría de las potencias de Jacobs	44
2.4. Contra una teoría reductiva: el esencialismo de Fine	46
2.4.1. Presentando la propuesta	48
2.4.2. Evaluando la propuesta	50
2.4.3. Un argumento general contra toda reducción esencialista	56
2.5. Conclusión	57
3. Una teoría de la modalidad: <i>estructuralismo absolutista</i>	58
3.1. Formulando afirmaciones sobre la estructura fundamental	58
3.2. Un marco formal para el estructuralismo absolutista	61
3.3. Algunas consecuencias de la postura y una breve comparación	68
3.4. Conclusión: estructuralismo absolutista y las esencias	75
II. Epistemología: Conocimiento, Fiabilidad, y Representación	
4. Contra el escepticismo del conocimiento modal y de esencias	78
4.1. Motivaciones para el anti-escepticismo modal	80
4.2. Contra van Inwagen	86
4.3. Contra Nozick	98
4.4. Conclusión, y una nota sobre el problema de Benacerraf	105
5. Fiabilismo externista	107
5.1. Preliminares: ontología/atribución, y la función de las intuiciones	109
5.2. Motivando un externismo mínimo	112
5.3. El fiabilismo	120

5.4. El problema de la generalidad	128
5.5. Contra el problema de la generalidad	130
5.6. Conclusión: a favor del fiabilismo externista	140
6. Una epistemología de la modalidad y la esencialidad: <i>representacionalismo fiabilista</i>	141
6.1. La metáfora de las etiquetas que se pegan al mundo	142
6.2. Tipos de representación	150
6.3. Semántica general: condiciones de adecuación representacional	156
6.4. Representaciones fiables	159
6.4.1. De la rigidez...	159
6.4.2. ...Pasando por un <i>excursus</i> sobre el bidimensionalismo...	162
6.4.3. ...A la representación fiable	164
6.5. El conocimiento modal y de esencias mediante la doble fiabilidad	168
6.5.1. El fiabilismo de proceso revisitado y generalizado	168
6.5.2. Fiabilismo de representación	169
6.5.3. Conociendo esencias y necesidades: la teoría central	171
6.5.4. Creando representaciones fiables	173
6.6. Después de la destrucción del Triángulo Dorado: Quine (otra vez), y la metafísica a oscuras	176
Conclusiones	180
Referencias	182

## Resumen de la tesis

En la introducción motivo brevemente los dos problemas centrales de esta tesis: uno epistemológico y uno metafísico. El metafísico consiste en el problema de definir qué son la necesidad y la posibilidad, por un lado, y la esencialidad, por otro. El epistemológico consiste en tener una teoría acerca de cómo sabemos que algo es posible o necesario, o que tiene una determinada esencia.

El resto de la tesis se divide en dos grandes partes: en la primera (capítulos 1-3) expongo los fundamentos lógico-metafísicos de la modalidad, la esencialidad, y su interrelación; mientras que en la segunda (capítulos 4-7) expongo los fundamentos epistemológicos de nuestro conocimiento de la modalidad y la esencialidad. Hay independencia entre las tesis epistemológicas propuestas en la segunda parte y las tesis metafísicas de la primera parte.

El capítulo 1 es, sobre todo, *expositivo*: expongo los aspectos básicos de la semántica estándar (dada en términos de los *marcos de Kripke*) para la lógica modal (proposicional y de primer orden), arguyendo sobre la base de resultados en la lógica que esta semántica *sólo* describe *estructuras relacionales*, que de hecho también pueden describirse con el lenguaje de la lógica clásica. Dado este hecho, surge la pregunta de si hay algo específicamente *modal* acerca de la lógica modal. Esto me lleva a exponer una popular interpretación filosófica (o conceptual) del formalismo, que llamo "*mundismo*". De acuerdo a esta idea, los *mundos posibles* satisfacen la estructura de la semántica estándar y además son el fundamento de las propiedades modales (incluyendo a los *modos* de ser verdad, como la *verdad posible*). Noto varios tipos de mundismo, y examino una popular variante—lo que David Lewis llamó *ersatzismo*. Finalmente, expongo una concepción de la esencialidad bastante popular y que se condice con el ersatzismo, que llamo "*concepción estándar de la esencialidad*". De acuerdo a esta concepción, básicamente, la esencialidad se reduce a la posesión necesaria de propiedades.

El capítulo 2 es, sobre todo, *crítico*: critico a las teorías *no-reductivas* que *no* toman a la modalidad como *absolutamente* fundamental (a estas teorías las llamo "*teorías no-no*"). Para esto, aíslo nociones de *reducción* y de *fundamentación* desde el inicio del capítulo. Noto que el ersatzismo regularmente se reconoce como *no reductivo*, pero que algunos teóricos buscan explicar desde esa postura a la naturaleza metafísica de la modalidad. Critico esta idea. También critico la idea de que el ersatzismo puede usarse como una representación perspicua de la modalidad. Después generalizo ambas críticas. Para dar más evidencia a favor de que esta generalización es correcta para toda teoría *no-no*, aplico la estructura de la crítica a la teoría de las potencialidades de Jonathan Jacobs. También critico una teoría *reductiva*: el esencialismo de Kit Fine. Arguyo que esta teoría falla en su pretendida reducción de la modalidad a la esencialidad, pues la esencialidad debe presuponer hechos modales previamente fijados. Finalmente, doy un argumento para la idea de que es altamente implausible que *alguna* teoría esencialista logre reducir la modalidad.

El capítulo 3 es, sobre todo, *propositivo*: propongo una nueva teoría sobre la modalidad. En el capítulo pasado motivé el realismo acerca de la modalidad notando que ciencias fundamentales (como la física y la matemática) lo asumen, y rechacé a las teorías *no-no* y a una teoría reductiva. Aquí argumento que todo esto nos da una motivación para creer que la modalidad es *absolu-*

*tamente fundamental*, en el sentido en que existen hechos o características modales que no se fundamentan en *absolutamente ningún otro* aspecto de la realidad. Expongo algunas breves consideraciones sobre cómo “insertar” a la metafísica de la modalidad dentro de la metafísica fundamental, en especial la desarrollada recientemente por Theodore Sider, de tal manera que hagamos sentido de la idea de la modalidad como fundamental. Después expongo una estructura formal para cierta variante de la teoría absolutista de la modalidad (que llamo *estructuralismo absolutista*), tal estructura induce lo que la llamo *marcos jerárquicos*. Entonces muestro cómo los modelos para la lógica modal estándar, la lógica de la esencialidad de Fine, y la lógica pura de la fundamentación de Fine, así como el acercamiento neo-platonista de Øystein Linnebo a la filosofía de las matemáticas, resultan ser *aplicaciones* de los marcos jerárquicos, la estructura formal que propongo. Finalmente, expongo tres consecuencias importantes del estructuralismo absolutista para problemas tradicionales de la metafísica modal, además de comparar mi teoría con, primero, otras teorías que también toman a la modalidad como absolutamente fundamental y, segundo, teorías anti-realistas de la modalidad; finalmente muestro cómo mi teoría se condice con la teoría esencialista de Fine, si a este se le remueven sus aspiraciones reductivistas.

El capítulo 4 inicia la segunda parte—la concerniente a la epistemología—de esta tesis, y es sobre todo, una sección *crítica*. Primero, motivo—apelando a evidencia de la ciencia cognitiva, así como filosófica—la postura según la cual deberíamos aceptar que tenemos *al menos* cierto conocimiento modal, y que por ello el escepticismo acerca de *todo* nuestro conocimiento modal es altamente implausible. Después analizo críticamente dos escepticismos *locales* sobre nuestro conocimiento modal: el de Peter van Inwagen, de acuerdo al cual no conocemos hechos sobre posibilidades “remotas”, y el de Robert Nozick, de acuerdo al cual no conocemos ningún tipo de necesidades. Argumento que ambos tipos de escepticismos están insuficientemente justificados; sin embargo, extraigo de la discusión tres problemas que toda epistemología modal debería resolver. Finalmente, considero brevemente un escepticismo basado en el problema de Benacerraf en epistemología de las matemáticas, y argumento que la metafísica de la modalidad que propuse en el capítulo 3 está libre de tal problema epistemológico.

El capítulo 5 es *crítico-propositivo*: expongo una teoría sobre las propiedades epistémicas—como la justificación o el conocimiento—que es externista y fiabilista. Primero hago la importante diferencia entre teorías sobre *la naturaleza* de las propiedades epistémicas, y las teorías sobre su *atribución*. En el camino, expongo dos acercamientos opuestos a la función filosófica de las intuiciones, y favorezco uno según el cual los juicios intuitivos, aunque pueden brindar evidencia para tesis filosóficas, pueden resultar falsos tras un proceso de investigación, y por ello no hay garantía *a priori* de que la teoría correcta sea la más intuitiva. Después, expongo y argumento a favor de una teoría *externista*, de acuerdo a la cual nuestro conocimiento puede estar en muchos casos determinado por factores a los que no tenemos acceso consciente. Noto cómo este externismo es compatible con varios tipos de tesis internistas, por lo que lo llamo *mínimo*. Finalmente, expongo el fiabilismo de Alvin Goldman. Este fiabilismo se enfrenta con el *problema de la generalidad*, motivado más fuertemente por Earl Conee y Richard Feldman. Argumento que tal problema tiene presuposiciones que están en fuerte tensión—y que, suplementadas con asunciones plausibles, llevan a incoherencias. Esto me lleva a rechazar tal problema, de parte del fiabilista.

El capítulo 6 cierra la tesis, y es, sobre todo, *propositivo*. Construyo una epistemología modal

que sólo necesita de la asunción de que *podemos representarnos* y que *podemos simular* (la evidencia empírica acerca de la segunda se revisa al inicio del capítulo 4). Utilizo herramientas modelo-teóricas y probabilísticas para dar una epistemología de la modalidad y la esencialidad que sigue el espíritu fiabilista y externista que fue defendido en el capítulo anterior. Comienzo caracterizando metafóricamente a las representaciones como “etiquetas” que se “pegan” al mundo, lo cual nos dará indicios para la epistemología que sigue. Entonces caracterizo con más exactitud a las representaciones, a partir de la filosofía del lenguaje, la mente, la ciencia y el arte. Esto nos da las bases para investigar la *teoría general de la representación*. Luego paso a motivar la asunción de una *semántica general*: una teoría al estilo de la teoría de modelos, que caracterice matemáticamente la asignación de contenidos semánticos a cualquier tipo de representación—en cuanto que *es* representación. Lo siguiente es aplicar la semántica general para generalizar las ideas fiabilistas acerca de la creencia a toda representación. La sección 6.5.3 da la propuesta epistemológica como tal, de acuerdo a la cual nuestro conocimiento modal y de esencias depende de una estructura *doblemente fiabilista*: nuestro conocimiento modal y de esencias está fijado por la fiabilidad de los procesos productores de representaciones, y por la fiabilidad de tales representaciones. Muestro también cómo la propuesta puede responder a los problemas generales extraídos de la discusión con van Inwagen en el capítulo 4. Cierro el capítulo, y la tesis, dado algunos comentarios acerca de las implicaciones del externismo para la investigación metafísica, argumentando que el “Triángulo Dorado” de Chalmers (la idea de que hay una conexión esencial entre significado, modalidad y aprioridad) ha sido definitivamente roto.

## Convenciones

- (1) Cuando una cita esté originalmente en inglés y no haya una traducción al español disponible, pongo mi traducción en el cuerpo principal del texto, llamando con un asterisco (o con dos, y así recursivamente) a una nota al pie, donde pondré la cita en el inglés original.
- (2) A lo largo del texto, cuando me refiero a la tercera persona del singular, salto de manera poco uniforme entre los géneros femenino y el masculino.
- (3) Pongo en VERSALES el nombre de definiciones y tesis a las que estaré regresando en la discusión.
- (4) Para nombrar proposiciones, uso “<” y “>”, de tal manera que  $\lceil \langle p \rangle \rceil$  nombra a la proposición que *p*.
- (5) Las referencias son en formato Autor, Año de publicación; aunque cuando el autor es mencionado en el contexto, sólo Año.
- (6) Los epígrafes son los rebeldes en esta tesis, por razones estilísticas. En primera, los dejo en su idioma original, contra la convención (1). En segunda, las referencias no van con el formato de la convención (5), sino directamente al texto citado. En tercera, en ellos hago énfasis con *itálicas*, que son todos míos—los uso para anticipar un punto que defenderé en la sección—y no del autor citado.
- (7) Una convención particularmente importante para los capítulos 1, 2 y 6: el concepto al que en inglés se suele referir con el término “*intended*” (como en “*intended model*”), en español se suele referir diciendo “intencional” (como en “modelo intencional”). Es un hecho des-

afortunado, pienso; una mejor palabra sería “pretendido”. Sin embargo, sigo el uso ya corriente en español. Por ejemplo, cuando en inglés diríamos “*intended referent*”, y cuando yo preferiría usar “referente pretendido”, diré “referente intencional”.

- (8) Finalmente, otra convención terminológica. Traduzco “*actuality*” por *realidad*—a diferencia de la práctica, desafortunadamente común, de traducirla como “actualidad”. “*Actuality*” es un falso cognado, no significa lo que significa “actualidad” en español: contemporaneidad. No me importaría este infortunio si no fuera por las consecuencias filosóficas que puede tener: cuando escucha, en inglés, “*non actual* possible”, uno tiene una sensación de cuasi-incoherencia que no provoca “*posibles no actualizados*”—pero que “*posibles no reales*” sí provoca. La decisión de traducir “*actuality*” por *realidad* está, entonces, motivada por la idea de no manchar las intuiciones filosóficas. Sin embargo, en muchos casos, he decidido notar que hablo de la *actuality* cuando sea relevante.

## Agradecimientos

En primera y de manera central, fue gracias al inquebrantable apoyo de mi familia—mis papás Laura y Oscar, fieles, protectores y perseverantes; mis hermanos Oscar y Diana; la memoria de mi abuelo, en cuya casa estudié por semanas seguidas—que pude, en primer lugar, dejarme de preocupar casi totalmente por las dificultades que la vida en nuestro país presenta. Es gracias al núcleo de mi familia que las necesidades primordiales de todo ser humano—emocionales y hasta económicas—nunca peligraron, y hasta fueron—diría mi mamá—defendidas como el león que defiende a sus crías.

Es a mi familia a quien en primer lugar debo esta tesis, y a ella quien agradezco sobre todo.

Pasé horas discutiendo en el seminario de los sábados, guiado por nuestra amiga y maestra Lourdes, y los asistentes, entre los que perseveraron Afra, Anaïd, Álvaro, Chayo, Emiliano, Erick, Hugo, Melahuac, los dos Rodrigos, Víctor, Viorica... Todos ellos me comentaron, en algún momento, cada capítulo de la tesis, y gracias a sus preguntas, objeciones, dudas y preocupaciones, pude hacer más clara la tesis o corregir muchos errores. A ustedes, gracias.

En el seminario de los sábados encontré no sólo filosofía, sino amistad y apoyo constante. Lo mismo pasó en las extensas y emocionantes reuniones en casa de Moisés con Cristian, Alejandro y Víctor (y Moisés, claro está), donde discutíamos filosofía después de que ellos se entrenaban en las oscuras artes del infinito (conocidas como “teoría de conjuntos”), y donde también podíamos comer alguna garnacha o platicar de ética y política.

Fuera de la práctica filosófica analítica, pero también resultado de la carrera en filosofía, estaban mis amigos de siempre: la banda que nos juntábamos en la casa de Natalia en Yautepec y/o en el California; también con la que iba a darketear al Centro o la Roma: Rafa, Ezequiel, y Álvaro. Estuvieron también las amistades perdidas en el ágora de la Fac, en donde también filosofé durante tardes enteras (en la charla informal sobre la cultura *pop*, en los garabatos de fórmulas en mi cuaderno).

A todos ellos: gracias, y que sean más años. (Y más filosofía.)

Además de la familia y amigos, está también mi gratitud en lo profesional, que va para mucha



gente en el posgrado—a los profesores que me dieron clase, al personal administrativo del posgrado que siempre tuvo buena cara ante mi ineptitud para los trámites y, sobre todo, a mis sinodales: Alessandro, Axel, Edgar y Eduardo, quienes me dieron muchos y excelentes comentarios que me ayudaron a ver mi tesis desde otros puntos de vista, además de corregir varios de sus puntos flacos.

Gracias especiales van para Lourdes Valdivia, que además de amiga personal siempre ha sido guía en lo filosófico y lo profesional. También para el CONACyT y la UNAM, quienes me apoyaron durante los dos años que dura la maestría para que lograra sobrevivir, así como al proyecto PAPIIT IN403311 (*Racionalismo en Metafísica, Semántica y Epistemología*).

Como todo lo humano, la filosofía es una empresa que depende de condiciones sociales específicas. De ahí que quiera agradecer no sólo a mis más cercanas, sino a toda aquella persona que, incluso de manera indirecta, contribuyó para que yo escribiera esto (en la manera más obvia, pagando sus impuestos para que yo pudiera disfrutar dos años de beca). Pero no sólo, quizá.

## 0. Introducción

ESTRAGON: No use struggling.

VLADIMIR: *One is what one is.*

ESTRAGON: No use wriggling.

VLADIMIR: *The essential doesn't change.*

—Samuel Beckett, *Waiting for Godot*

**M**e interesan dos grandes cuestiones aquí: una epistemológica y una metafísica. Ambas se dividen en dos cuestiones más específicas, pero interrelacionadas. Para la cuestión epistemológica, en primera, me interesa la cuestión de si podemos conocer las esencias de los objetos, su naturaleza: la cuestión de si podemos saber lo que una cosa *es*, en un sentido distintivamente metafísico. Junto a esto, también, la cuestión de si podemos saber lo que *es necesario*: la cuestión de si podemos saber que algo *no podría* cambiar, *no podría* ser distinto. La cuestión metafísica se relaciona con las cuestiones epistemológicas. Se divide en, por un lado, la pregunta de *qué es* una naturaleza, una esencia; por el otro, en la pregunta de *qué es* que algo sea necesario.

Lo que una cosa *es* en el sentido específicamente metafísico del que he hablado no está determinado por cualquier propiedad que la cosa en cuestión tenga—por ejemplo, no porque ahora yo esté hambriento diremos que yo *soy* un ser con hambre, en este sentido metafísico. Como contraste, lo que un conjunto *es*, en este sentido metafísico, es ser una cosa que tiene tales y cuales determinados miembros, una cosa tal que hay otras que están en la relación  $\in$  con ella.<sup>1</sup> Me interesa saber cómo es que podemos—si es que podemos—saber lo que las cosas *son* en este sentido metafísico.

¿Qué es lo que una cosa *es*, en este tan mencionado sentido metafísico? Lo que estoy preguntando aquí no es, puesto en otras palabras, *cuál* es la naturaleza de una cosa dada, sino *qué es* esa naturaleza—puesto en términos filosóficos, *qué es una esencia*. Varias secciones de la primera parte de esta tesis están dedicadas a esa cuestión (a diferencia de la teoría filosófica sobre *cuáles* esencias son poseídas por cuáles objetos). Pero, por el momento, quiero dar pie a una idea más o menos intuitiva de la noción de esencia.

La esencia de un objeto es lo que el objeto *realmente es*, lo que *define* al objeto, lo que le da *identidad* al objeto, lo que el objeto *no podría dejar de ser* sin también dejar de existir; incluso, quizá, lo que *explica* lo que el objeto *es*. La esencia de una cosa  $x$ , si de hecho existe tal esencia, es lo que diferencia a  $x$  de absolutamente cualquier otra cosa, ya sea meramente posible o real.

Así que la noción de esencia sirve una función teórica doble: una *estabilizadora*, y una *identificadora*. Sirve una función estabilizadora porque marca los límites del cambio posible: un objeto puede—por todo lo que la realidad dicta—cambiar todo lo que sea, *excepto* por aquello que el

---

1 Debemos, por supuesto, incluir la sutileza acerca del conjunto vacío, que aunque es un conjunto, no  $\in$ -contiene a ninguna otra cosa. Sin embargo, creo que el ejemplo es más intuitivo cuando dejamos de lado esta idea; el conjunto vacío resulta un caso límite de la generalización natural, donde un conjunto puede estar en la relación  $\in$  con alguna o ninguna cosa.

objeto *es*, en este sentido metafísico privilegiado. Sirve una función identificadora porque permite una especificación objetiva de la cosa: nada sino  $x$  puede tener la esencia de  $x$ . Ambos papeles son complementarios: es en parte porque  $x$  no puede cambiar lo que es por naturaleza, que su naturaliza permitirá identificarlo; es en parte porque  $x$  es identificado por su naturaleza, que tal naturaleza no admite cambios que preserven la identidad. Así, la cuestión de la necesidad está cercanamente relacionada con la cuestión de la esencialidad—aunque, como veremos, terminen siendo dos cosas distintas.

Si de hecho hay cosas que tienen esencias, o hechos que suceden por necesidad, es una cuestión posterior que no trataré en esta tesis.<sup>2</sup> Pero ya antes de investigar esa cuestión, parece interesante preguntarse qué sería que algo fuera necesario—*investigar qué es la necesidad*—y qué sería para un objeto el poseer una esencia—*investigar qué es una esencia*. Una vez teniendo una idea de ello, podríamos pasarnos a preguntar *cómo podríamos saber que un objeto tiene una esencia, o que algo le es necesario*. Estas dos—las preguntas ontológicas y epistemológicas—son exactamente las cuestiones centrales a mi tesis. Ahora bien, ¿por qué deberíamos interesarnos filosóficamente en estas preguntas—en la pregunta de qué son, y cómo podríamos conocer, las esencias y la necesidad?

Una primera aproximación para justificar el interés filosófico en la definición de lo que es una esencia podría consistir en notar lo central que tal concepto es para la filosofía misma: plausiblemente, la epistemología busca la esencia del conocimiento (su *definición real*; su *lo que es ser*); la estética, la esencia de lo bello (independiente de esta o aquella obra que resulta ser bella); la filosofía de la mente busca la esencia de la mente (independientemente de esta o aquella configuración neural o, en general, material, que resulta tener mentalidad); la filosofía del arte, la esencia del arte (independientemente de este o aquel medio que resulta brindar una obra de arte). Si esto es así, la filosofía busca asegurar afirmaciones de un tipo cuyas consecuencias suelen ser necesarias.

Bueno, quizá hoy en día no toda filósofa se sentiría cómoda con caracterizar así las aspiraciones de su rama favorita. Pero ¿no es, de cualquier manera, el caso que muchas veces, aún si no todas, de hecho la filosofía *sí* tiene tales aspiraciones? En empresas metafísicas, al menos, regularmente intentamos encontrar la esencia de ciertos objetos o propiedades, o la necesidad de tal o cual hecho. Es verdad, por ejemplo, que las personas *suelen vivir en ciudades o en entornos rurales*, o también *aspirar a un buen nivel de vida*; pero ¿es eso lo que es *ser* una persona? Es cierto también, por ejemplo, que a veces *sentimos* que el tiempo va más lento que de costumbre, o más rápido que de costumbre, pero ¿caracteriza eso lo que *es* el tiempo *en sí mismo*? Pero entonces regresamos a la cuestión: en primer lugar, ¿qué es “ser algo en sí mismo”? Incluso: ¿qué es ser algo “objetivamente”? Y estas preguntas son maneras de formular a la pregunta por la definición de la noción de *esencia*.

Incluso fuera de la filosofía, creo que la noción de esencia tiene cabida en nuestra vida pre-teórica. Sentimos, por ejemplo, que hay algo en la persona amada que la define de tal manera que la hace diferente de cualquier otra; sentimos que amamos a la persona por lo que esa persona *es*. Sentimos, a veces también, que *teníamos* que conocer a esa persona—que nuestro encuentro fue

<sup>2</sup> He tratado estas cuestiones en mi 2014 y mi tesis de licenciatura (Romero 2010).

cuestión de *necesidad* marcada por el destino. Quizá y muy probablemente, claro está, resulte que la idea de que el amor conoce esencias o necesidades no tiene cabida en un marco metafísico despojado de toda sentimentalidad y armado de rigor; pero al menos muestra que las ideas mismas de algo metafísicamente inevitable, y de este sentido metafísico en que las cosas *son*, no es completamente ajena al sentido común.

Especificar qué es la esencialidad y qué es la necesidad, es la tarea de la primera parte de mi tesis.

Ahora bien, si aceptamos que las nociones de esencia y de necesidad no están sacadas de la manga de algún metafísico sin mayor ocupación que especular a ver a dónde llega, y que pueden incluso llegar a jugar un papel teórico en la filosofía, surge inmediatamente la pregunta de cómo podemos *conocer* lo necesario y lo esencial (*si es que* hay tales cosas). Porque seguramente pocas entidades, si es que algunas, llevan su esencia a flor de piel. Por ejemplo, supongamos que aceptamos que la esencia de los números naturales está dada por los axiomas de Peano, de tal manera que los números naturales *son esencialmente* ésa estructura que ejemplifican todos los modelos isomorfos al modelo intencional de esos axiomas. La cuestión aquí es que, por más que los números naturales eran conocidos desde la antigüedad (incluso sabemos que algunos animales no-humanos pueden usar un pequeño fragmento de ellos para contar), los axiomas que describen su naturaleza básica tuvieron que esperar a la época de Peano para ser conocidos. Es decir, el trato milenario con ellos no fue suficiente para que pudiésemos conocer su esencia. Lo cual nos indica que la pregunta de cómo, en general, conocemos esencias, es *altamente no-trivial*.

Esa tarea es el objetivo de la segunda parte de mi investigación.

## 0.1. comentario metodológico

Cierro la introducción con un *comentario metodológico*: En varios lugares de la tesis, he hecho uso de herramientas matemáticas. Esto no quiere decir que mi objetivo sea hacer una tesis “meramente técnica”, cuando por ello se entiende afinar los detalles de un cierto aparato matemático que ha probado su utilidad filosófica. Tampoco busco ocultar los problemas filosóficos “bajo la alfombra” de un formalismo que parezca resolver todo—pero que resuelva aquello que resuelve sólo al precio de asumir soluciones a problemas que no se solucionan por estipulación. Cuando utilizo matemáticas—que, por lo demás, están, en su enorme mayoría a un nivel elemental: la lógica que aprendemos en cursos de licenciatura, la teoría de conjuntos que apenas presupone los axiomas básicos de ZFC y que también se aprende en la licenciatura, y la probabilidad que se enseña a nivel bachillerato en casi todo el país—lo hago con ánimos de aclaración: es justamente para no confundir una posición filosófica con una metáfora con contenido tan amplio que puede interpretarse de mil maneras. Pero esos usos no confunden formalización con argumento. Debemos, por un lado, evitar resolver problemas filosóficos pensando que un aparato matemático meramente estipulado y con definiciones injustificadas nos brinda la salida; debemos, por otro, evitar pensar que el uso de las matemáticas en filosofía es necesariamente “meramente técnico” y no sirve para aclarar la estructura de lo que investigamos filosóficamente—así sea la modali-

dad o el conocimiento. Después de algunos siglos en los que no pocas vagas metáforas se han aceptado como profundos *insights*, ya tenemos alguna advertencia. (Esto es especialmente relevante para los capítulos 1, 3 y 6, donde se ofrecen modelos matemáticos—relativamente simples—de tesis filosóficas).

Lo mismo para el recurso a la investigación empírica cuando esto sea posible. No asumo que la ciencia natural nos dará una respuesta a los problemas filosóficos—sí asumo que, en cuanto sea posible, hay que buscar extraer, *filosóficamente*, propuestas filosóficas de la ciencia. La filosofía no se acaba, ni se convierte en mera recapitulación de lo que los científicos dicen, si en vez de basarnos en la pura imaginación y la creatividad, tomamos los *datos* que la ciencia nos da. Al revés—con este *naturalismo moderado*, integramos a la filosofía en la esfera de conocimiento que las ciencias, con tanto éxito, han estado forjando. (Esto es especialmente relevante en el capítulo 4).

## I. METAFÍSICA:

FUNDAMENTALIDAD, MODALIDAD Y ESENCIA

## 1. La concepción estándar de la modalidad y la esencialidad

El primer tema sobre la mesa es definir las nociones centrales para esta tesis: *esencia*, *necesidad* y *posibilidad*. Cualquier lector informado sabrá que estas han sido centrales para muchos proyectos y tradiciones filosóficas, por lo que un análisis con pretensiones de exhaustividad histórica sería una empresa monstruosa. En lugar de ello, en este capítulo me restringiré a la manera en que muchos filósofos en la tradición analítica del siglo XX las entendieron, a partir de la década de los '70. Le llamaré a esa manera de entender tales nociones “la concepción estándar”, debido a lo popular que fue en su momento (aunque hoy día no sea la única y, quizá, tampoco la más favorecida, cf. §2.4, abajo). Según esta concepción, la teoría formal de la esencialidad es un caso especial de la lógica modal alética cuantificada, interpretada filosóficamente. Esta interpretación filosófica toma, básicamente, a la semántica modal como una descripción de la estructura del *pluriverso de mundos posibles*, pluriverso que a su vez fundamenta (o, según algunas teorías, *es lo mismo que*) el que haya modalidad en el mundo. Veamos.

### 1.1. aspectos básicos de la semántica estándar para la lógica modal cuantificada

We never had any reason to expect metaphysics to be easy. We should be grateful to have the technical power of modern logic at our command, to give us some grounds for hope that we can do better than our illustrious predecessors.

—T. Williamson, *Modal Logic as Metaphysics*

Como, según la concepción que estaremos analizando en este capítulo, lo *esencial* es un caso especial de lo *necesario* y como, además, lo necesario obedece ciertos constreñimientos estructurales que son estudiados por la lógica modal, resulta natural examinar cómo esta funciona. En especial, resulta importante examinar la semántica estándar para la lógica modal, junto con la interpretación filosófica que suele conectarla con cuestiones metafísicas.

Por ello, vamos a revisar algunos aspectos básicos de la semántica estándar para la lógica modal (la semántica de *marcos*).<sup>3</sup> En §1.3 expondré la manera en que la concepción estándar de la modalidad *interpreta metafísicamente* al aparato de la lógica modal.

Vamos a definir un lenguaje completamente estándar para la lógica modal de primer orden con identidad, lenguaje al que llamaré “ $\mathcal{L}^1(=, \Box)$ ”. En particular, para futuras referencias,  $\mathcal{L}^1(=, \Box)$  incluye:

<sup>3</sup> Cualquier libro de texto sobre lógica modal servirá como referencia para esta sección; aunque para la que viene no todos sirven. Véase, por ejemplo, Blackburn & van Benthem 2007, Braüner & Ghilardi 2007, y Jansana 1990.

$\{\rightarrow, \neg, \forall, =, \square\}$ , el conjunto de las *constantes lógicas*,<sup>4</sup>  
 $\{x_1, x_2, \dots\}$ , el conjunto de las *variables de primer orden*,  
 $\{f^1, f^2, \dots\}$ , el conjunto de los *símbolos de función n-ádicas* (para cada  $n$ ),  
 el conjunto de los *símbolos de función de cualquier aridad*, (la unión de los anteriores),  
 $\{a_1, a_2, \dots\}$ , el conjunto de las *constantes de individuo*,  
 el conjunto de los *términos*: la unión del conjunto de funciones, constantes y variables,  
 $\{P^1, P^2, \dots\}$ , el conjunto de los *símbolos de predicados n-ádicos* (para cada  $n$ ),  
 el conjunto de los *símbolos de predicado de cualquier aridad*, (la unión de los anteriores).

Las reglas de formación de  $\mathcal{L}^1(=, \square)$  son completamente estándar, y las omito por brevedad.

Ahora pasemos a la semántica. La conexión entre el lenguaje modal y la semántica estándar—la basada en marcos—surge desde la aceptación de los “bicondicionales leibnizianos” (cf. Sider 2002, §3):

Posiblemente,  $p$  sii  $p$  es verdadera en al menos un mundo posible,

Necesariamente,  $p$  sii  $p$  es verdadera en todo mundo posible.

La mayoría de los metafísicos de la modalidad consideran a los bicondicionales leibnizianos como una buena guía para la estructura de la modalidad.<sup>5</sup> Es cuando hemos dado por sentado que la modalidad metafísica—el que haya algo necesario o posible—puede ser iluminada por los bicondicionales leibnizianos, que la semántica estándar para la lógica modal parece una buena aproximación a su estructura. Veamos.

Definimos un *marco* como un par  $F = \langle W, R \rangle$ ; mientras que un *marco de primer orden con dominios variantes* es una 4-tupla  $F^1 = \langle W, R, I, Q \rangle$ , donde:

$W \neq \emptyset$ ; (que será nuestro conjunto de *mundos posibles*)

$R \subseteq W^2$ ; (la *relación de accesibilidad* entre mundos posibles)

$\emptyset \neq I \subseteq W$ ; ( $I$  es el conjunto de los *individuos*)

$Q: W \rightarrow \text{Pot}(I)$ , asignando a cada mundo  $w$  el *dominio de cosas existentes en  $w$* ,

(donde  $\text{Pot}(X)$  es el conjunto potencia de  $X$ ). Entonces definimos un *modelo de dominios variantes*  $M = \langle F^1, \varphi \rangle$ , donde:

$F^1$  es un marco de primer orden de dominios variantes; y  $\varphi$  es una función diádica tal que:

para  $c$  una constante de nuestro lenguaje,  $\varphi(c, w) \in Q(w)$  tal que  $\varphi(c, w) = \varphi(c, v)$ , para cualesquiera  $w, v \in W$ ;

para  $f$  un símbolo de función  $k$ -ádica,  $\varphi(f, w)$  es una función  $k$ -ádica,  $g: Q(w)^k \rightarrow Q(w)$ ;

para  $P$  un símbolo de predicado  $k$ -ádico,  $\varphi(P, w) \subseteq Q(w)^k$ .

Decimos que  $M$  está *basado en* o está *sobre* el marco de primer orden con dominios variantes  $F^1$ . Dado un modelo de dominios variantes  $M$ , definimos un conjunto de asignaciones de variables  $a$ , que van de las variables a  $I$ , tal que:

para cualquier asignación  $a$  y variable  $v$ :  $a(v) \in I$ .

Ahora definimos *verdad en un modelo  $M$  de acuerdo a un mundo  $w$  bajo una asignación de variables  $a$*  (relación que denotamos  $w \models_M, a$ ). Vamos a dejar implícito el índice para el modelo.

4 La disyunción y conjunción, el cuantificador existencial y el diamante (el operador de posibilidad) se asumen introducidos de manera estándar, mediante la negación.

5 Aún si no necesariamente una guía para su reducción a aspectos no modales. Sobre esto, véase las secciones posteriores de este trabajo.



Así, para  $w \in W$  y  $\alpha, \beta, \gamma$  fórmulas de nuestro lenguaje modal  $\mathcal{L}^1(=, \Box)$ :

- $w \models_a P(c_1, \dots, c_n)$  sii  $\langle \varphi(c_i) \rangle_{i \leq n} \in \phi(P, w)$ ; donde  $c_1, \dots, c_n$  son o constantes o símbolos de función, y  $P$  es un predicado  $n$ -ádico;
- $w \models_a (\forall v)(\alpha)$  sii para cualquier asignación  $a^*$  que difiera de  $a$  a lo más en lo que le asigna a  $v$ ,  $w \models_{a^*} \alpha$ ;
- $w \models_a (\alpha \rightarrow \beta)$  sii: o no  $w \models_a \alpha$ , o  $w \models_a \beta$ ;
- $w \models_a \neg \alpha$  sii no  $w \models_a \alpha$ ;
- $w \models_a \Box \alpha$  sii para todo  $u \in W$  tal que  $R(w, u)$ :  $u \models_a \alpha$ .<sup>6</sup>

Como se puede ver, la definición recursiva de las condiciones de verdad para un lenguaje de primer orden con operadores modales refleja las ideas que se extraen de los condicionales leibnizianos:  $\alpha$  es necesaria cuando es verdadera en todo mundo posible (accesible desde el mundo en el que se evalúa), y es posible (dado que  $\Diamond \alpha$  se toma como abreviación de  $\neg \Box \neg \alpha$ ) si es verdadera en al menos un mundo posible (accesible desde el mundo de evaluación).

Finalmente, definimos la relación *ser verdad en un modelo*  $M$ ,  $M \models$ , como sigue:

$M \models \alpha$  sii  $w \models_a \alpha$  para toda asignación de variables  $a$  y todo mundo  $w$  del modelo.

Decimos que  $\alpha$  es una *verdad de la lógica de*  $\mathcal{L}^1(=, \Box)$  sii  $M \models \alpha$ , para todo modelo  $M$  de dominios variantes.

En la siguiente sección voy a exponer (junto con algunas breves pruebas) algunos teoremas que caracterizan la relación entre los marcos, los modelos, y nuestro lenguaje modal. Eso nos llevará a preguntarnos qué exactamente es lo que le da una *importancia metafísica* a la lógica modal, junto con su semántica estándar, como para llegar a ser una parte esencial de lo que después describiré como *la concepción estándar de la modalidad*.

## 1.2. la semántica estándar: teoremas de correspondencia, de definibilidad y de caracterización

[...] a frame  $F$  is a couple  $\langle W, R \rangle$  [...]. There is nothing intrinsically 'modal' about all this, of course. Frames are just the 'directed graphs' of Graph Theory.

—Johan van Benthem, "Correspondence Theory"

Comencemos notando una importante característica de las estructuras lógico-matemáticas que hemos caracterizado en sus aspectos más básicos, los marcos y modelos de Kripke: ellos *no* son objetos *intrínsecamente modales*: son *estructuras relacionales*, objetos conjunto-teóricos. Pero es en términos de ellos que la lógica modal obtiene sus condiciones de verdad.<sup>7</sup> Si esto es así,

<sup>6</sup> Dadas las definiciones de las demás conectivas que asumimos arriba, las cláusulas de verdad para ellas son las obvias.

<sup>7</sup> Por supuesto, la semántica en términos de marcos kripkeanos no es la única semántica para el lenguaje modal—están al menos las semánticas de entornos [*neighborhoods*], que es una generalización de la semántica de marcos; además de la topológica y la algebraica. Aquí me enfoco en la semántica de marcos porque, para bien o para mal, ha sido la que ha ocupado el papel protagónico en las aplicaciones de la lógica modal a la metafísica de la modalidad. Para algunos casos, esta elección puede justificarse de manera muy obvia—por ejemplo, la semántica topológica requiere inmediatamente la validez del axioma **4** y por ello es inadecuada para analizar las características de lógicas más débiles—pero en otros casos la motivación es mucho menos clara: por ejemplo, hay una vas-

entonces cabe la posibilidad de que haya *más* en la modalidad que la estructura descrita por la lógica modal—aquello que trasciende el ser una estructura relacional. En el resto de esta sección exploraré esa posibilidad; la siguiente sección busca explicar qué *más* se ha asumido que hay en la modalidad, tradicionalmente.

(En la discusión siguiente, me enfocaré sobre todo en los lenguajes modales proposicionales; debido a que su semántica es mucho más fácil de tratar. Aunque la conclusión también puede argumentarse para los lenguajes modales cuantificacionales, es más difícil, dado que para ellos existen menos resultados análogos a los expuestos abajo.)<sup>8</sup>

Notemos primero que los modelos modales, *qua* objetos matemáticos, basados en los marcos, *no* necesitan del lenguaje modal para ser descritos—de hecho, un fragmento de la lógica clásica de primer orden es suficiente: ese fragmento es conocido por los lógicos modales como el *lenguaje de correspondencia* (con el lenguaje modal proposicional) *de primer orden*. Ese lenguaje está definido por la *traducción estándar*, que básicamente nos da una manera de pasar del lenguaje modal proposicional al lenguaje de correspondencia. Veamos.

Sea  $L^0(\Box, P)$  el lenguaje modal proposicional clásico (con un operador “caja” y uno “diamante”), donde  $P$  es cualquier conjunto denumerable de letras proposicionales. Sea  $L^1(\mathbb{R}^\Box, P^1)$  el fragmento del lenguaje de la lógica clásica cuantificacional de primer orden determinado por un conjunto denumerable de variables  $\{x, y, \dots\}$ , una letra binaria que llamaremos ‘ $\mathbb{R}^\Box$ ’, y un conjunto  $P^1$  de letras monádicas de al menos la cardinalidad de  $P$ . Definimos entonces una familia de *funciones de traducción estándar*  $\{TE_x\}_{x \in L^1(\mathbb{R}^\Box, P^1)}$  de  $\mathcal{L}^0(\Box, P)$  a  $\mathcal{L}^1(\mathbb{R}^\Box, P^1)$  como sigue (para  $p \in P$ , y  $\alpha, \beta$  fórmulas de  $\mathcal{L}^1(\mathbb{R}^\Box, P^1)$ ):

$$TE_x(p) = Px,$$

$$TE_x(\neg\alpha) = \neg TE_x(\alpha),$$

$$TE_x(\alpha \ \& \ \beta) = TE_x(\alpha) \ \& \ TE_x(\beta),$$

$$TE_x(\Box\alpha) = (\forall y)(\mathbb{R}^\Box xy \rightarrow TE_y(\alpha))$$

(donde  $y$  es la primera variable después de  $x$  en el orden estándar)

La idea básica es hacer corresponder a las variables cuantificacionales con los “mundos” (los elementos de  $W$ ) de los modelos modales. Así, el conjunto de “mundos” donde una letra proposicional  $p$  del lenguaje modal es verdadera (o expresa una verdad), corresponderá al conjunto de “mundos” donde la correspondiente fórmula cuantificacional  $Px$  es verdadera. Es por eso que la traducción estándar para una fórmula modalizada  $\Box\alpha$  emula estructuralmente el comportamiento semántico de  $\Box$ .<sup>9</sup> Con esto, para cada modelo de la lógica modal proposicional  $M$ , obtenemos un modelo para la lógica de primer orden  $M^*$ , donde  $M^* = \langle W, R, \langle \varphi_{p1} \rangle$ : para cada  $p \in \mathcal{L}^0(\Box, P)$  hay una  $P^1 \in \mathcal{L}^1(\mathbb{R}^\Box, P^1) \rangle \rangle$ . Ahora es fácil probar la siguiente proposición:

TEOREMA DE LA TRADUCCIÓN ESTÁNDAR. *Sea  $\alpha$  una fórmula del lenguaje modal proposicio-*

---

ta clase de lógicas que son incompletas respecto a la semántica kripkeana, pero que son completas respecto a  $M$ -otras semánticas (Braüner & Ghilardi 2007). La cuestión de cómo balancear entre generalidad y manejabilidad en la elección de una semántica *metafísicamente privilegiada* para la lógica modal es una cuestión que pocos metafísicos modales, si alguno, se han tomado en serio; pero que desafortunadamente yo tampoco puedo tocar en esta tesis.

8 Me baso fundamentalmente en Blackburn & van Benthem 2007, Blackburn & de Rijke & Venema 2001, caps. 2 y 3, Jansana 1990, y van Benthem 1984.

9 Y, por supuesto, del operador de posibilidad,  $\Diamond$ : asumiendo la definición estándar, de la condición para  $\Box$  obtenemos la traducción para fórmulas conteniendo al diamante:  $TE_x(\Diamond\alpha) = (\exists y)(\mathbb{R}^\Diamond xy \ \& \ TE_y(\alpha))$ .

nal  $\mathcal{L}^0(\Box, P)$ ,  $M$  un modelo para este lenguaje, y  $w \in W \in M$ .<sup>10</sup> Sea  $M^*$  el modelo de primer orden obtenido a partir de  $M$ . Entonces  $w \models_M \alpha$  sii  $M^* \models TE_x(\alpha)$ , cuando a  $w$  le asignamos la variable  $x$ .

*Prueba:* Dado como definimos nuestra traducción estándar,<sup>11</sup> lo único que tenemos que hacer es una inducción sobre la complejidad de las fórmulas. Por ejemplo, para cada  $p$  tenemos una  $P^1$  tal que:  $w \models_M p$  sii  $w \in \phi_{P^1}$  (es decir, cuando la traducción le asigna  $P^1$  a  $p$ ,  $p$  es verdadera en  $w$  sii  $w$  está en la extensión de  $P^1$  bajo  $\phi_{P^1}$ ). Los casos de la negación y la conjunción son obvios (para la negación, tomamos complementos (relativos a  $W$  y bajo la  $\phi$  relevante), para la conjunción, tomamos intersecciones). Suponiendo que hemos mostrado el caso de  $\alpha$ , para el operador de caja tenemos que: primero, sabemos que  $w \models_M \Box\alpha$  sii para todo  $v$  tal que  $wRv$ :  $v \models_M \alpha$ . La fórmula asignada a  $\Box\alpha$  por la traducción es  $(\forall y)(R^\Box xy \rightarrow TE_y(\alpha))$ . A su vez,  $M^* \models (\forall y)(R^\Box xy \rightarrow TE_y(\alpha))$  sii para cualquier objeto  $v'$  en  $W$  que sustituyamos por la variable  $y$ , si  $R^\Box xv'$ , entonces  $v' \in$  la extensión de  $TE_y(\alpha)$  bajo  $\phi$ . Dado que interpretamos a la letra  $R^\Box$  como la relación  $R$  sobre el modelo original  $M$ ,  $R^\Box uu'$  va a ser verdad sólo si  $\langle u, u' \rangle$  está en  $R$ . Por lo tanto si en  $M^*$  es verdad que  $R^\Box xv'$ , entonces en  $M$  es verdad que  $\langle x, v' \rangle$  está en  $R$ . Y dado eso, dadas las cláusulas de verdad del lenguaje modal proposicional,  $x \models_M \Box\alpha$  sii  $v' \models_M \alpha$ . Pero entonces  $v' \in$  la extensión de  $TE_y(\alpha)$  bajo  $\phi$ .  $\dashv$

Este teorema nos dice que la lógica modal proposicional corresponde a *cierto fragmento* de la lógica cuantificacional clásica—el fragmento al cual hacemos la traducción desde el lenguaje modal proposicional. Sabemos, de entrada, que ese fragmento sólo necesita de una variable libre—la variable sobre la cual se hace la traducción, pues todas las demás variables son ligadas por un cuantificador existencial (para el caso del operador  $\Diamond$ ), o por uno universal (para  $\Box$ ). ¿Podemos individuar este fragmento de una manera más informativa? El famoso *teorema de correspondencia* de van Benthem responde a esta pregunta.

TEOREMA DE CORRESPONDENCIA EN LA TRADUCCIÓN ESTÁNDAR DE VAN BENTHEM. *Sea  $\alpha(x)$  una fórmula del lenguaje de correspondencia de primer orden. Entonces  $\alpha(x)$  es invariante bajo bisimulaciones sii es equivalente a la traducción estándar de una fórmula modal.*<sup>12</sup>

El elegante teorema de van Benthem nos dice qué parte del lenguaje de primer orden con una variable libre es aquel que corresponde al lenguaje modal proposicional: aquel lenguaje que es invariante bajo bisimulaciones. La idea intuitiva es que dos modelos modales son bisimilares cuando sus dominios (*i.e.*, sus conjuntos de mundos) están relacionados de tal manera que cada mundo en un dominio tiene un “espejo” en el otro dominio: si un mundo hace ciertas proposiciones verdaderas, también su espejo, y si un mundo  $R$ -accesa a otro  $v$ , entonces su espejo  $R$ -accesa a otro  $v^*$ , que a su vez es espejo de  $v$ . La definición es:

BISIMULACIÓN: Dos modelos  $M = \langle W, R, \phi \rangle$  y  $M' = \langle W', R', \phi' \rangle$  para la lógica modal

10 Un modelo proposicional es un modelo de dominios variantes (definido arriba) sin  $I$  ni  $Q$ . Es decir:  $\langle W, R, \phi \rangle$ , donde  $\phi$  le asigna a cada letra proposicional  $p$  un subconjunto  $\phi(p)$  de  $W$ .

11 En realidad, definimos una *familia* de traducciones—una para cada variable de nuestro fragmento de lenguaje de primer orden. Aquí voy a fijar una traducción arbitraria, para una variable  $x$  fija.

12 La prueba de este teorema es algo larga y técnica, por lo que excede el alcance de este trabajo. Para ello, véase Blackburn & de Rijke & Venema 2001: capítulo 2, teorema 2.68, p. 103.

proposicional son *bisimilares* sii existe una relación binaria no-vacía,  $B$ , entre sus dominios ( $B \subseteq W \times W'$ ) tal que:

$B(w, w')$  sii:

- (1)  $w$  y  $w'$  satisfacen las mismas letras proposicionales (son “espejos”),
- (2) si  $R(w, v)$ , entonces hay un mundo  $v'$  en  $W'$  tal que  $B(v, v')$  y  $R'(w', v')$
- (3) si  $R'(w', v')$ , entonces hay un mundo  $v$  en  $W$  tal que  $B(v, v')$  y  $R(w, v)$

Ahora tomemos nuestro lenguaje de correspondencia  $\mathcal{L}^1(\mathbf{R}^\square, P^1)$ . Lo que el teorema de van Benthem nos dice es que una fórmula de este lenguaje corresponderá a un fórmula modal proposicional sii su valor de verdad se preserva a través de modelos bisimilares: si  $M$  y  $M'$  son dos modelos bisimilares (no importa bajo cuál bisimulación en específico), una fórmula  $\alpha$  del lenguaje de correspondencia es verdadera en  $M$  sii es verdadera en  $M'$ .<sup>13</sup>

Es decir, el lenguaje modal proposicional, visto como descripción de los modelos modales, corresponde al lenguaje de primer orden que sólo da para describir estructuras relacionales de primer orden que sean bisimilares: aquellas que sean *bisimilares* no podrán ser distinguidas por la lógica modal (*i.e.*, las estructuras bisimilares son equivalentes desde el punto de vista de la lógica modal proposicional).<sup>14</sup> Si es así, no es claro que el lenguaje modal describa, en y por sí mismo, fenómenos *modales*; no al menos, más de lo que el lenguaje cuantificacional clásico lo hace, al describir *estructuras bisimilares*. Entonces, si lo único que hubiera en la modalidad fuera lo que es descrito por la lógica modal, la modalidad se agotaría en ser un fenómeno que se agota en la bisimilaridad.

Voy a hacer otro punto relacionado con el anterior. Para verlo, notemos que, así como podemos ver al lenguaje modal proposicional como aquel fragmento del lenguaje cuantificacional que describe *modelos*, individuados (desde el “punto de vista” del lenguaje) por bisimulaciones, podemos verlo también como un lenguaje que describe directamente los *marcos* Kripkeanos. La idea sería que diferentes *fórmulas* en el lenguaje modal *definen* diferentes estructuras en el nivel de los marcos. Es decir: una fórmula de la lógica modal proposicional típicamente también puede verse como una fórmula que, agregada al sistema  $\mathbf{K}$ ,<sup>15</sup> *define cierto tipo de estructura relacional*. Para esto, dos definiciones:

VALIDACIÓN EN MARCOS. Un marco  $\mathbb{F}$  *valida a una fórmula*  $\alpha$  ( $\mathbb{F} \models \alpha$ ) sii para todo modelo  $M$  tal que  $M = \langle \mathbb{F}, \varphi \rangle$  (para alguna  $\varphi$ )  $M \models \alpha$ . Una clase  $\mathcal{C}$  de marcos *valida a*

13 Cuando a las variables de  $\alpha$  se le asignan los mundos relevantes. La definición exacta es esta: una fórmula  $\alpha(x)$  del lenguaje de correspondencia cuya única variable libre es  $x$ , es *invariante bajo bisimulaciones* sii para cualesquiera modelos  $M$  y  $M'$ , y cualesquiera “mundos”  $w$  en  $M$  y  $v$  en  $M'$ , y para toda bisimulación  $B$  tal que  $B(w, v)$ :  $w \models_M \alpha(x/w)$  sii  $M' \models \alpha(x/v)$ .

14 Aquí hay que ser cuidadosos. Que dos modelos sean bisimilares implica que harán exactamente las mismas fórmulas modales verdaderas (que son *modalmente equivalentes*), pero la converso no es verdad en lo general. Así, la equivalencia modal no implica bisimilaridad. Sólo hay condiciones especiales en las que la converso es verdad, véase Blackburn & de Rijke & Venema 2001, seccs. 2.3-2.5. (Una caracterización más exacta de la equivalencia modal es esta: *dos modelos son modalmente equivalentes sii sus extensiones por ultrafiltros son bisimilares* (cf. *op. cit.*, teorema 2.62, p. 97). Otra condición equivalente a la equivalencia modal es la bisimilaridad de ciertas ultrapotencias de los modelos equivalentes, véase el teorema 2.74, p. 106.)

15 Como es bien sabido, el sistema proposicional  $\mathbf{K}$  resulta de agregarle, a una axiomatización de la lógica proposicional clásica (con regla de sustitución), el axioma:  $\Box(p \rightarrow q) \rightarrow (\Box p \rightarrow \Box q)$  y la regla de necesidad (que nos dice que todo teorema es necesario). A su vez, tal axioma se conoce como el axioma  $\mathbf{K}$ .

una fórmula  $\alpha$  ( $\mathfrak{C} \models \alpha$ ) sii cada marco en  $\mathfrak{C}$  valida a  $\alpha$ . Un marco  $\mathbb{F}$  (clase de marcos  $\mathfrak{C}$ ) valida a un conjunto de fórmulas  $\Sigma$  sii  $\mathbb{F} \models \alpha$  ( $\mathfrak{C} \models \alpha$ ), para cada  $\alpha$  en  $\Sigma$ .

DEFINIBILIDAD DE MARCOS. Una fórmula modal proposicional  $\alpha$  define una clase  $\mathfrak{C}$  de marcos sii la clase de marcos que validan a  $\alpha$  es *exactamente*  $\mathfrak{C}$  (i.e., todos y sólo los marcos en  $\mathfrak{C}$  validan a  $\alpha$ ).

Algunos ejemplos clásicos de teoremas de definibilidad son los siguientes.

TEOREMAS DE DEFINIBILIDAD. El axioma **K**, es decir, la fórmula  $\Box(p \rightarrow q) \rightarrow (\Box p \rightarrow \Box q)$  define a la clase de todos los marcos. El axioma **T**:  $\Box p \rightarrow p$ , define a la clase de los marcos reflexivos (i.e., tales que su relación de accesibilidad es reflexiva). El axioma **4**:  $\Box p \rightarrow \Box \Box p$  define a la clase de los marcos transitivos.

Para tener una idea de cómo se prueban este tipo de hechos, probemos rápidamente el teorema para el axioma **4** y la transitividad (más pruebas pueden encontrarse en Jansana 1990: cap. 4).

*Prueba del teorema de definibilidad para el caso del axioma 4:* Sea  $\mathbb{F}$  un marco transitivo que no valida **4**, derivamos una contradicción ahora. Por asunción,  $\mathbb{F} \models \Box p$  pero no  $\mathbb{F} \models \Box \Box p$ . Esto quiere decir que hay alguna valuación  $\varphi$  sobre  $\mathbb{F}$  y mundo  $w$  en  $\mathbb{F}$  tales que  $w \in \varphi(\Box p)$ , pero  $w \notin \varphi(\Box \Box p)$ . Esto quiere decir que, para todo  $v$  tal que  $Rwv$ ,  $v \in \varphi(p)$ , pero que hay algún  $v'$  con  $Rwv'$  tal que  $v' \notin \varphi(p)$ . Este  $v'$  será entonces tal que para algún  $v''$  con  $Rv'v''$ , no  $v'' \in \varphi(p)$ . Pero como  $\mathbb{F}$  es transitivo, y como  $Rwv'$  y  $Rv'v''$ , entonces  $Rwv''$ , y el que  $v'' \notin \varphi(p)$  contradice el hecho de que todo  $v$  tal que  $Rwv$ ,  $v \in \varphi(p)$ . Por lo tanto, si  $\mathbb{F}$  es transitivo, valida **4**. Para la converso, supongamos que  $\mathbb{F}$  no es transitivo. Mostramos ahora que hay mundos donde **4** es falso: mundos  $w$  tal que  $w \in \varphi(\Box p)$  pero  $w \notin \varphi(\Box \Box p)$ . Que  $\mathbb{F}$  no sea transitivo es que hay mundos  $w$  en  $\mathbb{F}$  tal que  $Rwv$  y  $Rvv'$ , pero no  $Rwv'$  (para algunos  $v, v'$  en  $\mathbb{F}$ ). Fijemos unos  $w, v, v'$  así, y supongamos que  $w \in \varphi(\Box p)$ . Ahora bien, todavía puede pasar que  $w \notin \varphi(\Box \Box p)$ , puesto que nada impide que  $v \notin \varphi(\Box p)$  al pasar que  $v' \notin \varphi(p)$ , con  $Rvv'$ . Pero como  $R$  no es transitiva ( $Rwv$  y  $Rvv'$ , pero no  $Rwv'$ ),  $v' \notin \varphi(p)$  es compatible con que  $w \in \varphi(\Box p)$ . Así,  $\mathbb{F}$  valida a **4** sólo si es transitivo. Juntando los dos resultados: **4** define a los marcos transitivos.  $\dashv$

Ahora podemos notar otra característica importante del lenguaje modal: si puede definir *propiedades de las relaciones* en los marcos, entonces tiene el mismo poder expresivo que, al menos, un fragmento del lenguaje de la lógica clásica *de segundo orden*. Otra manera de verlo es notando que “[...] si pensamos en lo que significa que una fórmula modal sea válida en un marco, veremos que este concepto es esencialmente de segundo orden: cuantificamos sobre todas las valuaciones posibles, y las valuaciones asignan subconjuntos de marcos a los símbolos de proposición”\* (Blackburn & van Benthem 2007: p. 38). Esto, por supuesto, sucede porque al pasar del nivel de los modelos al nivel de los marcos, abstraemos la información que nos provee la función de valuación que constituye a los modelos—abstraemos cuantificando sobre *todas las valuaciones posibles definidas sobre un tipo de marco* (e.g., transitivo, simétrico, etcétera) *dado*.

Esto nos lleva a un segundo lenguaje de correspondencia: un *lenguaje de correspondencia* (con el lenguaje modal proposicional) *de segundo orden*. Ése lenguaje es el lenguaje de corresponden-

\* “[...] if you think about what it means for a basic modal formula  $\phi(p_1, \dots, p_n)$  to be valid on a frame, we see that this concept is essentially second-order: we quantify across all possible valuations, and valuations assign subsets of frames to proposition symbols.”

cia de primer orden  $\mathcal{L}^1(\mathbb{R}^n, P^1)$  que revisamos arriba, suplementado con variables de predicado para cada  $P^1$ , la constante  $=$  para la identidad, y los correspondientes cuantificadores de segundo orden. Entonces:

TEOREMA DE CORRESPONDENCIA DE SEGUNDO ORDEN. *Sea  $\alpha(p_1 \dots p_n)$  cualquier fórmula modal proposicional donde aparecen las letras proposicionales  $p_1 \dots p_n$  en ese orden. Entonces:  $\mathbb{F} \models \alpha(p_1 \dots p_n)$  sii  $\mathbb{F} \models (\forall P_1 \dots P_n)(\forall x)(TE_x(\alpha))$ .*<sup>16</sup>

*Prueba:* Sabemos por el TEOREMA DE LA TRADUCCIÓN ESTÁNDAR (de primer orden) que  $w \models_M \alpha$  sii  $M \models TE_x(\alpha)$ , cuando a  $w$  le asignamos la variable  $x$ . De esto y por las cláusulas de verdad modal,  $M \models \alpha$  sii  $M \models (\forall x)(TE_x(\alpha))$ . Ahora bien,  $\mathbb{F} \models \alpha(p_1 \dots p_n)$  sii para toda  $\varphi$  sobre  $\mathbb{F}$ ,  $\langle \mathbb{F}, \varphi \rangle \models \alpha(p_1 \dots p_n)$ . De lo que  $\mathbb{F} \models \alpha(p_1 \dots p_n)$  sii para toda  $\varphi$  en  $\mathbb{F}$ ,  $\langle \mathbb{F}, \varphi \rangle \models (\forall x)(TE_x(\alpha(p_1 \dots p_n)))$ . Pero cada  $\varphi$  simplemente asocia un conjunto de puntos tomados de  $W$  a cada  $p_i$ , que por el TEOREMA DE LA TRADUCCIÓN ESTÁNDAR es lo mismo que asociarle tal tipo de conjuntos a cada  $P^1_i$ . Así, que para toda  $\varphi$  en  $\mathbb{F}$ ,  $\langle \mathbb{F}, \varphi \rangle \models (\forall x)(TE_x(\alpha))$  es lo mismo que para cualquier asignación de extensiones a las  $P^1_i$ ,  $\mathbb{F}$  valide  $(\forall x)(TE_x(\alpha(p_1 \dots p_n)))$ . Pero entonces, por razonamiento de segundo orden, esto último es lo mismo que  $\mathbb{F} \models (\forall P_1 \dots P_n)(\forall x)(TE_x(\alpha))$ .

†

Ya podemos ver cómo las fórmulas modales proposicionales pueden definir la estructura de una relación en un marco: lo hacen al tener el poder expresivo de un fragmento del lenguaje de la lógica clásica *de segundo orden*.<sup>17</sup>

Así, vistas como describiendo modelos, las fórmulas de la lógica modal proposicional describen estructuras bisimilares de primer orden; vistas como describiendo marcos, las fórmulas describen estructuras de segundo orden.<sup>18</sup> Parafraseando el epígrafe de van Benthem que abre esta

16 Como en el caso de primer orden, en estricto sentido tenemos que definir una estructura  $F^{**}$  de segundo orden, que es la que valida las fórmulas de segundo orden; pero como todo lo necesario ya se encuentra en nuestro marco  $F$  original, seguiré diciendo que es este el que hace verdaderas a las fórmulas de segundo orden.

17 Es interesante notar que todo el poder expresivo de la lógica de segundo orden *no siempre* es necesario: por poner un ejemplo obvio, podemos definir implícitamente a los marcos transitivos diciendo que son aquellos cuyas relaciones  $R$  satisfacen:  $(\forall x)(\forall y)(\forall z)(Rxy \ \& \ Ryz \rightarrow Rxz)$ , y lo hicimos sin cuantificar sobre predicados.

Hay ciertos métodos para determinar cuándo una fórmula modal, al nivel de los marcos, es esencialmente de primer orden (*i.e.*, determina relaciones que también pueden ser determinadas por fórmulas del lenguaje clásico de primer orden), y también tenemos un importante teorema para saber cuándo la estructura de un marco (su ser transitivo, o reflexivo, etcétera) que puede ser definida por el lenguaje de primer orden, puede también ser definida por el lenguaje modal proposicional. Para lo primero, tenemos el teorema de van Benthem de correspondencia en la definibilidad (según el cual fórmula modal define una propiedad de marcos que es definible en el lenguaje de primer orden si y sólo si la fórmula modal es preservada bajo cierta operación—la construcción de ultrapotencias); para lo segundo, está el famoso teorema de caracterización de Goldblatt y Thomasson, según el cual una clase de marcos que es definible con el lenguaje de primer orden es definible con fórmulas modales si y sólo si esa clase está cerrada bajo ciertas operaciones (uniones disjuntas, submarcos generados, e y imágenes mórficas acotadas). Véase Blackburn & de Rijke & Venema 2001: secc. 3.5. para las definiciones y las pruebas.

18 De nuevo, es claro que dependiendo de qué semántica le imponamos a nuestro lenguaje, diferentes estructuras serán las descritas por él. Por ejemplo, una semántica algebraica asume que el lenguaje describe cierto tipo de *álgebras booleanas*, y una semántica topológica asume que el lenguaje describe cierto tipo de *espacios topológicos*. La elección de la semántica de marcos como la “metafísicamente privilegiada” es una asunción que ha pasado a formar parte de la metafísica modal desde el trabajo de Kripke, pero que no parece haber sido completamente justificada. Me parece que aquí hay bastante que explorar, pero desafortunadamente no puedo tratarlo

sección, *no hay nada intrínsecamente modal en todo esto*. El lenguaje modal proposicional describe estructuras relacionales—ya sea con el poder expresivo de un fragmento del lenguaje de primer orden, o de uno del lenguaje de segundo orden.

Todavía no hemos revisado con el mismo cuidado al lenguaje modal de primer orden, pero revisamos brevemente su semántica en la sección anterior: el único cambio en los marcos es que a cada punto  $w$  de  $W$  le asignamos un conjunto de puntos (disjunto de  $W$ )  $Q(w)$ , tomado de  $I$ , que llamamos “el dominio de  $w$ ”. Los cuantificadores entonces corren sobre estos dominios, y las evaluaciones de las fórmulas predicativas se hacen relativas a los mundos. Ahora podremos ver que la interacción de los operadores modales con los cuantificadores de primer orden describe la estructura de  $R$  junto con la de  $I$  (y por ello, la de cada  $Q(w)$ ), así como en el caso proposicional la interacción de los operadores modales junto con las conectivas lógicas describía la estructura de  $R$ .

Como en el caso proposicional, existe una traducción estándar de la lógica modal de primer orden, pero ahora a la lógica clásica de primer orden *con dos tipos* (“two-sorted”): un tipo lo reservamos para mundos, otro para individuos dentro de esos mundos; reservamos también dos símbolos relacionales binarios:  $R^\square$  (que, por supuesto, se interpretará como la relación de accesibilidad) y  $E$  (la relación de *existir en un mundo*, que tiene como extensión exactamente el dominio del mundo donde se evalúa) y como antes, definimos una familia de traducciones (las meta-variables  $a, b, c, \dots$ , correrán sobre las variables para mundos; mientras que  $\chi, \psi$  son metavariables para variables de individuo):

$$TE_a(P(\chi_1, \dots, \chi_n)) = P^*(a, \chi_1, \dots, \chi_n); P \text{ un predicado } n\text{-ádico y } P^* \text{ uno } n+1\text{-ádico};$$

$$TE_a(\chi = \psi) = (\chi = \psi)$$

$$TE_a(\neg\alpha) = \neg TE_a(\alpha),$$

$$TE_a(\alpha \ \& \ \beta) = TE_a(\alpha) \ \& \ TE_a(\beta),$$

$$TE_x(\Box\alpha) = (\forall b)(R^\square ab \rightarrow TE_b(\alpha))$$

(donde  $b$  es la primera variable después de  $x$  en el orden estándar)

$$TE_a((\forall\chi)(\alpha)) = (\forall\chi)(Ea\chi \rightarrow TE_a(\alpha))$$

Con esto es fácil probar, como lo hicimos anteriormente para el caso proposicional, que un modelo para la lógica modal de primer orden hace verdadera a una fórmula  $\alpha$  de tal lógica sii hace verdadera a su traducción estándar  $TE_w(\alpha)$  en la lógica de primer orden con dos tipos (básicamente, un modelo para esta lógica consiste de dos dominios: uno para cada tipo de término, y la función de interpretación está constreñida de la manera correspondiente).

Como ejemplo, se sabe desde el trabajo inicial de Kripke que la fórmula Barcan:

$$\forall x(\Box Ax) \rightarrow \Box \forall x(Ax)$$

define la siguiente relación (entre  $E$  y  $R^\square$ ):

$$\forall w \forall v (R^\square wv \rightarrow \forall x (Ewx \rightarrow Ewx)),$$

es decir, la fórmula Barcan es válida en una clase de modelos sii para cualesquiera mundos  $w$  y  $v$  en cualquiera de esos modelos, si  $w$  accesa a  $v$ , entonces  $Q(v)$  es un subconjunto de  $Q(w)$ —y, por ello, es válida en exactamente todos los modelos basados en marcos de primer orden con dominios *decrecientes*.

Así como arriba vimos (con el TEOREMA DE CORRESPONDENCIA EN LA TRADUCCIÓN ESTÁNDAR DE VAN

BENTHEM) que el lenguaje modal proposicional corresponde al sublenguaje de la lógica de primer orden que es invariante bajo bisimulaciones, ahora podemos ver que el lenguaje modal de primer orden corresponde al fragmento de la lógica cuantificacional con dos tipos que es invariante tanto bajo bisimulaciones (donde la noción se refina adecuadamente para el caso cuantificacional, cf. van Benthem 1984, p. 370) y para *submodelos generados*:

SUBMODELOS GENERADOS. Un modelo modal de primer orden  $M' = \langle W', R', I', Q', \varphi' \rangle$  es un *submodelo generado de otro modelo*  $M = \langle W, R, I, Q, \varphi \rangle$  sii:

- (a)  $W' \subseteq W$ ,
- (b)  $R' = R$  restringida a  $W'$ ,
- (c)  $I' \subseteq I$ ,
- (d)  $Q' = Q$  restringido a  $W'$ ,
- (e)  $\varphi' = \varphi$  restringido a  $W'$ ,
- (f) si  $w$  está en  $W'$  y  $Rwv$ , entonces  $v$  también está en  $W'$ .

Las condiciones (a)-(e) definen *submodelos*, la condición (f) nos individua a los *submodelos generados*.

Con estas nociones en mano, van Benthem probó el siguiente teorema:

TEOREMA DE CORRESPONDENCIA EN LA TRADUCCIÓN ESTÁNDAR PARA EL LENGUAJE DE PRIMER ORDEN DE VAN BENTHEM. *Una fórmula del lenguaje cuantificacional con dos tipos es equivalente a una fórmula modal cuantificacional de primer orden sii la primera es invariante para submodelos generados y bisimulaciones.*

*Prueba:* ver van Benthem 1984, §2.5.

Este tipo de resultados, así como en el caso proposicional, nos muestran, de una manera particularmente clara, que los lenguajes modales<sup>19</sup> no están esencialmente, o intrínsecamente, conectados con la *modalidad metafísica*: al menos asumiendo la semántica estándar que se da en términos de marcos kripkeanos, los lenguajes modales describen estructuras relacionales, y son “ciegos” a la cuestión de si tales estructuras informan fenómenos metafísicamente modales o de otro tipo completamente diferente. Por poner un ejemplo, podemos definir marcos cuyo dominio  $W$  consista de subconjuntos de la clase de todos los particulares (físicos) existentes en nuestro universo, y cuya relación  $R$  se interprete como la relación *estar a la izquierda de* (relativo a un marco de referencia) sobre ese dominio. Nuestro lenguaje lógico-modal (definido por las conectivas booleanas, cuantificadores e identidad, y el rombo y la caja) seguirá sirviendo igual de bien para describir tales marcos. Pero claramente no hay nada específicamente modal acerca de ellos.

Veremos en la próxima sección que la idea de que los lenguajes modales no describen, *en y por sí mismos*, fenómenos modales, y que las estructuras de los marcos kripkeanos no son, *en y por sí mismas*, estructuras modales, ha sido reconocida por algunos de los autores pioneros en la metafísica modal. Sin embargo, también argumentaré que *no es claro que los metafísicos moda-*

<sup>19</sup> Por supuesto, aquí sólo hemos revisado los lenguajes modales *del tipo básico*: aquellos que sólo contienen el operador unario de caja (y el diamante se introduce por definición). Hay muchos otros tipos de lenguaje modal: donde los operadores modales son poliádicos, o donde hay más de un tipo de operador modal (en las lógicas modales multi-dimensionales). Pero, en la medida en que para esos lenguajes haya teoremas de correspondencia, de definibilidad y de caracterización análogos a los revisados en esta sección, puntos análogos a los argumentados aquí se siguen para esos casos.



*les sigan estrictamente las consecuencias de este punto.* Como veremos en el capítulo 2, importantes teorías de la modalidad se siguen tomando demasiado a pies juntillas la idea de que la modalidad es una cuestión puramente de mundos posibles, idea fuertemente sugerida por una lectura literal de la semántica estándar para los lenguajes básicos. Esto les traerá varios problemas que, una vez notados en el capítulo 2, intentaré superar en el capítulo 3.<sup>20</sup>

Por el momento basta con quedarnos con la idea de que la lógica modal describe estructuras que no son intrínsecamente modales y que, por ello, necesitamos de *más teoría* para conectar a esa lógica con la naturaleza de la modalidad.

## 1.2 interpretando filosóficamente a la semántica modal estándar: *mundismo y ersatzismo*

¿Por qué deberíamos suponer que el aparato formal necesita de alguna interpretación? La razón es que, como vimos en la última sección, el aparato formal de la teoría de modelos modales sólo especifica *estructuras* que pueden ser ejemplificadas en una vasta variedad de dominios.

Esto nos da razones para reconocer que la semántica formal de la lógica modal alética no nos brinda una teoría sobre lo que *es* la modalidad alética. Lo más que nos podría brindar es la *estructura lógica* de la modalidad, y todavía habría que responder si la modalidad *es sólo* esa estructura (de hecho, dada la existencia de semánticas alternativas, falta ser argumentado que la semántica kripkeana debe ser la privilegiada). Consideremos algunas ideas por dos filósofos pioneros en la metafísica de la modalidad.

Lewis escribe:

Cuando digo que los mundos posibles ayudan con el análisis de la modalidad, no quiero decir que ayuden con el “análisis semántico de la lógica modal” metalógico. El interés reciente con los mundos posibles empezó ahí, por supuesto. Pero incorrectamente. Para ese trabajo, no necesitamos mundos posibles. Necesitamos conjuntos de entidades las cuales, como guía heurística, “pueden ser tomados como” mundos posibles, pero que en realidad pueden ser cualquier cosa que quieras. Estamos haciendo matemáticas, no metafísica. Donde necesitamos mundos posibles, más bien, es al aplicar los resultados de estas investigaciones metalógicas. Los resultados metalógicos, por sí mismos, no responden ninguna pregunta acerca de la lógica de la modalidad. Sólo nos dan respuestas condicionales: si los operadores modales pueden ser correctamente analizados de-tal-y-cual manera, entonces ellos obedecen tal-y-cual sistema de lógica modal. Debemos considerar si ellos de hecho podrían ser analizados así; y entonces estamos haciendo metafísica, no matemáticas. [...] ([Lewis 1986]: p. 17)\*

<sup>20</sup> Agradezco a Lourdes Valdivia el haberme presionado para poner en contexto filosófico estos resultados formales.

\* “When I say that possible worlds help with the analysis of modality, I do not mean that they help with the metalogical ‘semantical analysis of modal logic’. Recent interest in possible worlds began there, to be sure. But wrongly. For that job, we need no possible worlds. We need sets of entities which, for heuristic guidance, ‘may be regarded as’ possible worlds, but which in truth may be anything you please. We are doing mathematics, not metaphysics. Where we need possible worlds, rather, is in applying the results of these metalogical investiga-

Mientras que Kripke nos dice (2005, p. 24, nota 18):

No pienso en los “mundos posibles” como si suministraran un análisis *reductivo* en ningún sentido filosóficamente importante, esto es, como si descubrieran la naturaleza última de los operadores modales, las proposiciones, etc., desde un punto de vista epistemológico o metafísico, o como si los “elucidaran”. En el desarrollo real de nuestro pensamiento, los juicios que contienen locuciones modales directamente expresadas (“podría haber sido el caso que...”) ciertamente son anteriores. [...] La motivación principal para el “análisis de mundos posibles” —y la manera como esto aclaró la lógica modal— fue que permitió tratar la lógica modal mediante las mismas técnicas conjuntistas de la teoría de modelos que se había[n] mostrado tan exitosa[s] al aplicarse a la lógica extensional. También es útil para aclarar ciertos conceptos.

Aunque, en una página anterior (p. 21), nos recordaba:

Pero no quisiera dejar ninguna impresión exagerada de que repudio completamente los mundos posibles, ni siquiera de que los considero como un mero artificio formal. El uso que yo mismo hice de ellos debió de haber sido lo suficientemente amplio como para evitar cualquier malentendido semejante.

Tenemos, entonces, que los resultados de la lógica modal apenas nos hablan de la metalógica de cierta clase de lenguajes, y de las estructuras que se pueden postular para dar un análisis de las condiciones de verdad de tales lenguajes—pero un análisis abstracto, modelo-teórico, todavía alejado de la cuestión, menos abstracta, de cuál es la naturaleza de la modalidad metafísica.

Pero como también dice Kripke, esto no es decir que los mundos posibles sólo hayan sido postulados como mero artificio formal; el problema es que todavía no es claro qué otra función más sustantiva Kripke sí asume que cumplen.

Aquí Lewis nos da una idea (en la cita de arriba): “si los operadores modales pueden ser correctamente analizados de-tal-y-cual manera, entonces ellos obedecen tal-y-cual sistema de lógica modal”. Es decir: *si es que* los operadores pueden ser analizados como cuantificadores sobre ciertas estructuras relacionales, entonces los operadores modales del lenguaje natural tienen cierta estructura lógica. Pero esto sólo enuncia una condición suficiente, y no se ha argumentado que sea necesaria. De hecho, no puede serlo. Hay lógicas modales (proposicionales y cuantificacionales) que *no* son completas para la semántica de marcos—lo cual es decir que aunque los operadores sean regimentados mediante cierto sistema axiomático, habrá fórmulas en su lenguaje que, aunque verdaderas en todos los modelos de los axiomas y las reglas, no son derivables en el sistema (*cf.* Fine 1974, Thomasson 1974). Por lo que hay operadores que, cuando se toman como cuantificadores sobre marcos kripkeanos, su estructura lógica no es completamente especificada.

Aquí voy a seguir la suposición usual, que se toma muy en serio el poder analítico que las estructuras de Kripke tienen en el área de la lógica modal, y busca una teoría acerca de los *fundamentos metafísicos* de la modalidad que pueda *embeberse* en las estructuras kripkeanas; aunque, como he notado arriba, no me parece que tal supuesto se haya establecido concluyentemen-

---

tions. Metalogical results, by themselves, answer no questions about the logic of modality. They give us conditional answers only: if modal operators can be correctly analysed in so-and-so way, then they obey so-and-so system of modal logic. We must consider whether they may indeed be so analysed; and then we are doing metaphysics, not mathematics.”

te.<sup>21</sup>

Claramente, la interpretación filosófica más popular es la que podemos llamar *mundismo*, y consiste en la afirmación de que los puntos en las estructuras kripkeanas que nos interesan son *mundos posibles*. Estas entidades, a su vez, se conciben como maneras en las que el universo podría ser, que, si ocurrieran, implicarían una distribución completa y consistente de valores de verdad para las proposiciones. Fuera de esto, la naturaleza de estas entidades es una fuente de controversias interminables, y aquí no consideraré todas.

La primera teoría acerca de los mundos, que no consideraré aquí, es la de Lewis (1986), de acuerdo a la cual los mundos son sumas mereológicas de objetos en un continuo espaciotemporal que es topológicamente conectado y maximal. A esta teoría sobre la naturaleza de los mundos Lewis adiciona una teoría sobre la naturaleza de la realidad (*actuality*) según la cual la *actuality* no es una propiedad sustantiva, sino indexical (tan indexical como la propiedad que un día tiene al ser *hoy*); además de una teoría de las contrapartes.

Pero la interpretación filosófica más popular de la semántica formal de marcos para la lógica modal alética se conoce como *ersatzismo* (el nombre viene de Lewis mismo, 1986), una teoría cuya idea básica consiste en postular *mundos posibles*, que a su vez se identifican con ciertas entidades abstractas. Aunque esta interpretación no es la única disponible, es la que examinaré abajo debido a que parece ser la más popular entre los metafísicos de la modalidad.

Quizá la variante más importante del ersatzismo es el que Lewis llamó “ersatzismo lingüístico”. “Ersatzismo lingüístico” es en realidad una etiqueta para una familia de teorías, unidas por la siguiente idea general: los mundos posibles son “construcciones” “lingüísticas” hechas de cosas “previamente” existentes. Uso las comillas porque no se asume que los mundos posibles sean, literalmente, entidades *lingüísticas*, o *construidos*, o *construidos* a partir de cosas que existían previamente *en un sentido temporal*. Más bien, las “construcciones” son “lingüísticas” porque se llevan acabo en una muy cercana analogía con la “construcción” de un lenguaje formal: uno primero especifica las entidades primitivas y entonces especifica, de manera recursiva, reglas para formar entidades complejas. Y esto último indica que las entidades primitivas tienen que tomarse como anteriores, *en el orden del ser*, a las construidas.

En las formas más claramente lingüísticas del ersatzismo, por ejemplo, tenemos el acercamiento de Fine (1977), inspirado en Prior, quienes identificaron a los mundos posibles con ciertas proposiciones complejas; o el de quienes toman a los mundos posibles como conjuntos máximos

21 Parece que el argumento más popular para la aceptación de una metafísica modal basada en los mundos posibles es la utilidad de estos en la lógica modal; pero esto ya debería ponerse en duda debido a lo que he venido diciendo en esta y la sección pasada.

Otra idea es dar argumentos “*face-value*” desde el discurso modal (se argumenta que la manera en que hablamos nos compromete con mundos posibles), como el muy famoso argumento de Lewis en su *Counterfactuals* (1973), p. 84. Estos tiene al menos tres desventajas, que sólo señalaré aquí: lo único que parece más o menos claro es que el compromiso es con *posibilidades*, pero es una asunción extra el que estas sean lo mismo que los *mundos posibles*. Además, un proyecto nominalista podría tratar de mostrar que las condiciones de verdad de nuestro lenguaje de hecho no necesitan de los mundos posibles; o un teórico del error argumentaría que aunque tenemos tal compromiso, nuestras afirmaciones sobre posibilidades son siempre falsas porque no existen tales cosas (no me comprometo con estos proyectos, sólo digo que el argumento los ignora). Finalmente, incluso suponiendo que tuviésemos tal compromiso, me es completamente oscuro el por qué se asume que las entidades con las que nos comprometemos con nuestro lenguaje natural deben usarse para explicar la *naturaleza metafísica* de cualquier cosa fuera de tal lenguaje.

y consistentes de proposiciones (Adams 1974; Lycan & Shapiro 1986). Otros los toman como modelos tarskianos, o una variante de ellos (deRosset 2012; McGee 2006; Rayo 2012). Incluso otros deciden construir no cada mundo posible de manera recursiva, sino que postulan al pluriverso de una vez por todas, mediante una “oración del pluriverso” que los especifica de manera interdependiente (Sider 2002).

Otra importante variante del ersatzismo, de raigambre quineana (Quine 1968), consiste en postular mundos posibles que no son construidos como entidades lingüísticas, sino en conjuntos de tuplas de números, la primera de ellas representando regiones de una variedad que se identifica con un espaciotiempo, además de series de números que representan ya sea que la región o el punto está ocupada de materia, o que ciertas propiedades se ejemplifican en ella (Heller 1998a, b); sin embargo, Lewis argumenta que este tipo de ersatzismo también puede verse como un tipo de ersatzismo lingüístico (1986: §3.2).

De acuerdo a otras teorías ersatzistas, cercanas pero no idénticas a esta, los mundos posibles pueden analizarse de otra manera. Por ejemplo, la teoría de Plantinga (1979, §IV.1) dice que un mundo posible es un estado de cosas máximo y posible, *i.e.*, un estado de cosas  $w$  tal que si  $w$  hubiera ocurrido, entonces para cualquier estado de cosas  $w^*$ , ya sea este o su complemento hubiera ocurrido. Una teoría parecida es el combinatorialismo de Armstrong, de acuerdo al cual los mundos posibles son estados de cosas que resultan de combinar particulares y universales reales, aunque la teoría resultante termina siendo ficcionalista acerca de tales estados de cosas meramente posibles (Armstrong 1989). Dando algunos pasos más lejos del ersatzismo lingüístico pero cercanos todavía a la idea intuitiva de los mundos posibles como *maneras totales en las que el mundo podría ser*, llegamos a la teoría de Forrest (1986), quien identifica a los mundos posibles con propiedades complejas que no están ejemplificadas. Finalmente, quizá el caso extremo es una teoría que toma a los mundos posibles como entidades abstractas primitivas, como la de Stalnaker (1976).

Como podemos ver, todas estas teorías parecen recuperar una imagen intuitiva de los mundos posibles como *grandes hechos o maneras en que el mundo podría ser*, determinados hasta el último detalle: para cualquier  $\langle p \rangle$ , debido a la maximalidad de cada mundo  $w$ , debe haber un hecho acerca de si  $\langle p \rangle$  es o no verdadera de acuerdo a  $w$ .<sup>22</sup> Pero no puede ser que  $\langle p \rangle$  sea tanto verdadera como falsa en  $w$ , pues se asume que cada mundo posible es *consistente*.

Un primer problema con el acercamiento “mundista” es cuál se asume que es el papel de los mundos posibles. En un primer acercamiento, su papel es *semántico*, o, en general, *heurístico-representacional*; mientras que para otros teóricos, su papel es *ontológico*. A veces los dos papeles se confunden, accidental o intencionalmente.

Que a los mundos se les dé un papel *semántico* significa que se los usa simplemente en una teoría acerca de las condiciones de verdad de los fragmentos modales (subjuntivos) del lenguaje natural (o de *un* lenguaje natural). La teoría de las condiciones de verdad de un (fragmento de) lenguaje  $L$  dicen qué se necesita que exista, y cómo esto ha de ser, para que las oraciones de  $L$  sean verdaderas. En general, a los mundos se les puede dar un papel heurístico: además de usarlos para la semántica de lenguajes modales, se los postula para usarlos en teorías acerca del con-

22 Hay una controversia acerca de si los mundos posibles podrían de hecho ser máximos; pues Bringsjord (1985) argumentó que esto, más asunciones conjunto-teóricas usuales, llevaban a inconsistencia. Véase también Menzel 2012.

tenido semántico de estados mentales, o en otras teorías filosóficas.

El punto es que si hemos de usar a los mundos sin aceptar algún ficcionalismo o instrumentalismo,<sup>23</sup> es difícil entender cómo podemos dar condiciones de verdad para oraciones que en muchos casos son verdaderas, o utilizarlos para formular teorías que también buscan ser verdaderas, sin estar comprometidos a que esos mundos existieran. Una vez que el teórico mundista se ha tomado en serio la idea de dar una semántica de mundos para el lenguaje modal, parece verse comprometido con que tales mundos existen y tienen una naturaleza particular.

Pero ¿cómo todo esto se relaciona con la pregunta sobre cuál es la *naturaleza* de la modalidad misma? Es entonces que pasamos al uso *ontológico* de la modalidad. Una vez que la filósofa mundista postula sus mundos posibles para la teoría semántica, una idea es pasar a utilizar a los mundos posibles en una *explicación metafísica* de la naturaleza de la modalidad—incluyendo aquí a los proyectos reductivistas. Otra es simplemente afirmar, como Kripke parece hacerlo en la cita al inicio de esta sección, que los mundos posibles no explican totalmente la naturaleza de la modalidad, sino que ayudan a “iluminar” o elucidar tal naturaleza. Estos acercamientos, muy probablemente el más popular entre los ersatzistas, dice que los mundos posibles ayudan a explicar *parte* de la naturaleza de la modalidad—aunque no *todo* lo que hay en ella, pues regularmente estos acercamientos reconocen que utilizan ideología modal primitiva. (Esto quedará más claro en el capítulo siguiente.)

Muchas veces no es claro qué papel se asume que juegan los mundos posibles. Para algunos, como Lewis, ellos juegan un papel tanto ontológico como semántico-representacional; pero otros filósofos son mucho menos explícitos al respecto.

Voy a individuar varias tesis; de su combinación podemos obtener diferentes tipos de acercamientos:

MUNDISMO. Los mundos posibles son la herramienta fundamental (ya sea en un uso ontológico, uno representacional, o ambos) para la investigación de la modalidad.

ABSTRACCIONISMO. Los mundos posibles son entidades *abstractas*.

ACTUALISMO. Los mundos posibles son entidades reales, que existen *de hecho* (*actually*); no hay nada meramente posible: todo lo que existe, existe de hecho: todo lo que existe, tiene *actuality*.

REPRESENTACIONALISMO. Los mundos posibles *representan* meras posibilidades.

IDENTIFICACIONISMO. Los mundos posibles *son* las posibilidades.

Todas las filosofías consideradas hasta ahora, de Lewis a Plantinga, aceptan MUNDISMO: los mundos posibles existen y son aquellos objetos que dan una interpretación metafísico-modal a la semántica kripkeana de la lógica modal alética. Aunque la considera una tesis demasiado oscura como para definir su postura con su negación, Lewis parece rechazar ABSTRACCIONISMO (1986, secc. 1.7). Todos los demás mundistas la aceptan. Todos los demás mundistas también aceptan ACTUALISMO; de hecho, es difícil decir si Lewis la acepta o no, dado que su teoría ontológica de la *actuality* la entiende como una propiedad no absoluta, sino indexical.

Finalmente, Lewis también acepta IDENTIFICACIONISMO, afirmando que su postura logra reducir a la modalidad—es decir, que él puede definir exactamente todas las posibilidades usando ideología que no es modal (véase el siguiente capítulo), por lo que sus mundos posibles resultan ser

---

23 A la Rosen (1990); pero fuera de Armstrong, ninguno de los mundistas citados se reconoce como ficcionalista.

exactamente las posibilidades que hay.<sup>24</sup>

Otros filósofos, pero de vena ersatzista, parecen también aceptar IDENTIFICACIONISMO. Esta aceptación parecería problemática, dada la aceptación de ACTUALISMO. Pues si los mundos posibles son posibilidades, entonces, debido a que excepto por el mundo *actual* hay una infinidad de posibilidades no realizadas, existirían un infinito de posibilidades no realizadas. Pero eso parece ir directamente en contra de ACTUALISMO.

Parece que una mejor postura, si aceptamos ACTUALISMO, es negar IDENTIFICACIONISMO y afirmar REPRESENTACIONALISMO. Muchos ersatzistas afirman esta conjunción, aceptando con ello la introducción de ideología modal primitiva, pues los mundos posibles se toman como representaciones de posibilidades, que sólo pueden ser caracterizadas usando ideología modal.

En secciones y capítulos posteriores tendré varias cosas que decir acerca de estas tesis y sus conjunciones. Por el momento, ahora que hemos logrado un entendimiento estable de las características del ersatzismo, me interesa ver cómo va la explicación que esta postura hace de la esencia.

### 1.3 esencia y modalidad

The earlier literature on modality, arising from the work of Quine, was characterized by an unwarranted contempt for modal notions. The subsequent literature, arising from the work of Kripke, has been characterized by an unwarranted enthusiasm. [...] ‘modal mania’ is a matter of seeing everything as modal; every notion which is somehow associated with modal features is itself taken to be modal. [...] *Philosophers with a new concept are like children with a new toy; their world shrinks to one in which it takes centre stage.*

—Kit Fine, *Modality and Tense*

Parece claro que la noción más intuitiva que podemos encontrar de esencia, según la cual la esencia de una cosa es lo que la cosa *es*, lo que la *define*, la manera en que una cosa es *dejando de lado sus accidentes*, tiene un elemento modal. Pues lo que una cosa *es*, en este sentido relativamente intuitivo, no podría dejar de serlo, a riesgo de dejar de existir. Así, una vez que tenemos un entendimiento filosófico suficientemente firme de la modalidad y la lógica modal, como se alega que lo tenemos al aceptar el ersatzismo, podemos brindar una definición teórica de *esencia* usando a la lógica modal, iluminando así a la noción intuitiva. Por simplicidad voy a agrupar todas las siguientes definiciones bajo el *umbrella term* “definición modal de esencia”.

De acuerdo a Plantinga (1974: p. 72):

(ESENCIA-PLANTINGA)

Una propiedad *E* es una *esencia* sii hay un mundo *w* en donde existe un objeto *x* que (1)

<sup>24</sup> Aunque al final de su 1986, Lewis modifica esta idea y afirma que de hecho cada individuo es una posibilidad. Véase 1986, §4.4.

tiene  $E$  esencialmente, y (2) es tal que no hay otro mundo  $w^*$  en donde exista un objeto distinto de  $x$  que tenga  $E$ .

Falta saber qué significa que un objeto tenga una propiedad esencialmente. De acuerdo al mismo Plantinga, (p. 56):

(POSESIÓN ESENCIAL-PLANTINGA)

$x$  tiene  $P$  esencialmente sii  $x$  tiene  $P$  en todo mundo donde  $x$  exista.<sup>25</sup>

Plantinga no es el único en caracterizar a la noción de esencia, junto con la noción de posesión esencial de una propiedad (equivalentemente, de *propiedad esencial*) mediante el lenguaje de la lógica modal cuantificada. Salmon, por ejemplo (2005: p. 216), la asume, mientras que en 1979: n. 2, define justo así a las propiedades esenciales; y Kripke (2005: p. 45) parece presuponerla. (De hecho, parece que la definición modal de esencia ha sido la más popular desde el resurgimiento de la metafísica en la tradición analítica). Ahora, como nota Fine (1994: pp. 3-4), se han propuesto al menos tres maneras de caracterizar la posesión esencial de una propiedad en el lenguaje de la lógica modal cuantificada:

(PROPIEDAD ESENCIAL INCONDICIONAL)

$P$  es esencial a  $x =_{df} \Box(P(x))$

(PROPIEDAD ESENCIAL CONDICIONAL EN EXISTENCIA)

$P$  es esencial a  $x =_{df} \Box(E(x) \rightarrow P(x))$  (donde  $E$  es el predicado que denota a la *existencia*)

(PROPIEDAD ESENCIAL CONDICIONAL EN IDENTIDAD)

$P$  es esencial a  $x =_{df} \Box(x = x \rightarrow P(x))$ .

Si además, siguiendo a Plantinga, insistimos en caracterizar una esencia como una propiedad (quizá compleja) necesaria y suficiente para la existencia en cualquier mundo, tendríamos tres caracterizaciones análogas:

(ESENCIA INCONDICIONAL)

$S$  es la esencia de  $x =_{df} \Box(S(x)) \ \& \ \Box(\forall y)(S(y) \rightarrow y=x)$

(ESENCIA CONDICIONAL EN EXISTENCIA)

$S$  es la esencia de  $x =_{df} \Box(\forall y)(S(y) \leftrightarrow (E(y) \ \& \ y=x))$

(ESENCIA CONDICIONAL EN IDENTIDAD)

$S$  es la esencia de  $x =_{df} \Box(\forall y)(S(y) \leftrightarrow (y = y \ \& \ y=x))$

Bajo asunciones plausibles, es inmediato mostrar que algunos pares de las anteriores definiciones son necesariamente extensionalmente equivalentes. Por ejemplo, se suele identificar a la propiedad  $E$  de existencia con la propiedad de la autoidentidad, *i.e.*,  $\lambda x[x = x]$ .<sup>26</sup> Si aceptamos esa

25 De manera equivalente pero algo más compleja (y quizá por ello algo más informativa), Plantinga define (p. 32) el tener una propiedad esencial así:  $x$  tiene  $P$  esencialmente sii  $x$  tiene  $P$  y la proposición  $K(x, P)$  es necesariamente falsa, donde  $K$  es la *kernel function* definida así: donde  $x$  y  $P$  son respectivamente un objeto y una propiedad,  $K(x, P)$  es una proposición (la *proposición kernel para  $x$  y  $P$* ) que expresa el resultado de reemplazar “ $x$ ” y “ $P$ ” en “ $x$  tiene el complemento de  $P$ ” por nombres propios para  $x$  y  $P$  (cf. p. 30). Así, la proposición  $K(\text{Sócrates}, \text{Chato})$  es la proposición  $\langle \text{NEG} \langle \text{Sócrates}, \text{Chato} \rangle \rangle$ , es decir, la proposición expresada por la oración en español “Sócrates no es chato”.

26 Véase Salmon 1987, y los tratamientos estándares de las lógicas libres.

definición, entonces las definiciones (ya sean de esencia o de propiedad esencial) condicionales en la existencia y en la identidad, colapsan en una sola. Por otro lado, si asumimos que ningún objeto puede tener propiedades en los mundos donde no exista (una posición que Plantinga defiende en 1974], secc. IV.3, y en Plantinga 1976, donde la bautiza como “*Actualism*”), entonces las definiciones incondicionales colapsan en una sola con las definiciones condicionales en la existencia. Además, si asumimos que un objeto puede ser idéntico consigo incluso en mundos donde no exista, entonces las definiciones incondicionales y condicionales en la identidad colapsan en una sola. Finalmente, las definiciones incondicionales y las condicionales (ya sean de esencia, o de propiedad esencial) colapsan en una sola si el operador de necesidad en cuestión se interpreta, en la semántica formal, como un cuantificador restringido a los mundos donde  $x$  exista (lo que Kripke (1971) llama “interpretar a la necesidad de manera débil”).

No voy a argüir por ninguna de las asunciones arriba mencionadas.

En este momento tampoco me interesa preguntarme qué pares de propiedades y objetos podrían satisfacer ya sea las definiciones de *propiedad esencial* o las de *esencia*.<sup>27</sup> Para los objetivos de mi tesis, prefiero esbozar cómo va la explicación de la esencialidad en términos de la lógica modal, a su vez interpretada a la manera ersatzista.

Supongamos que aceptamos que una esencia es *sólo aquello* que tal definición nos dice, y nada más. Entonces, de acuerdo a lo visto en la sección anterior y a la definición modal de esencia, tendríamos:

(EXPLICACIÓN ESTÁNDAR DE LA ESENCIA)

La esencia de un objeto  $x$  es aquella propiedad  $E$  (quizá compleja) tal que: todo mundo posible que fundamenta al menos algún hecho posible acerca de  $x$  (*i.e.*, todo mundo posible *de acuerdo al cual*  $x$  existe), representa a  $x$  como siendo  $E$ ; además, no hay ningún mundo posible de acuerdo al cual un objeto  $y$ , diferente de  $x$ , sea representado como siendo  $E$ .<sup>28</sup> Esto es, la esencia de  $x$  es aquella propiedad que es necesariamente extensionalmente equivalente con la identidad de  $x$ .

Si nos hemos de tomar al pie de la letra la explicación modal de la esencia, la noción intuitiva que esboqué en la introducción a esta tesis debe tener elementos que, o están equivocados, o agregan características extra e innecesarias a lo esencial. Pues EXPLICACIÓN ESTÁNDAR DE LA ESENCIA sólo recupera el aspecto de “invarianza modal” de la esencia: sólo nos dice que la esencia de una cosa es aquello que la cosa no podría cambiar y que ninguna otra cosa además de la primera podría poseer. Pero recordemos cómo esboqué lo que parece ser la noción más intuitiva de *esencia*, en el sentido distintivamente metafísico que hemos estado tratando:

La esencia de un objeto es lo que el objeto *realmente es*, lo que *define* al objeto, lo que le da *identidad* al objeto, lo que el objeto *no podría dejar de ser sin* también dejar de existir;

27 Traté esta cuestión en mi tesis de licenciatura (Romero 2010), revisando el caso del esencialismo acerca de la propiedad *ser ficticio* y el esencialismo de origen, ambos motivados por Kripke; mientras que en Romero 2014 reviso un terreno algo más amplio. La literatura sobre esta cuestión de hecho es abundante. Por ejemplo: Kripke (2005: 3a. Conferencia, n. 52 y *addenda*) defiende que el origen de los objetos materiales les es esencial a ellos, y también que el ser ficticio le es esencial a todo objeto de ficción (como Sherlock Holmes o Pedro Páramo). En Mackie 2006 se revisa con más amplitud de miras varias propuestas acerca de qué propiedades son esenciales para qué tipo de individuos, y se dan argumentos en contra del esencialismo del origen y del sortal.

28 Esta última afirmación presupone, claro está, el principio conocido como NECESIDAD DE LA IDENTIDAD (*i.e.*,  $(\forall x)(\forall y)[(x = y) \rightarrow \Box(x = y)]$ ).



incluso, quizá, lo que *explica* lo que el objeto *es*. La esencia de una cosa, si de hecho existe tal cosa, es lo que diferencia a tal cosa de absolutamente cualquier otra, ya sea meramente posible o real.

Alguien que acepte EXPLICACIÓN ESTÁNDAR DE LA ESENCIA podría sostener que su explicación, debido a que nos brinda una explicación extensional que resulta ser una mera aplicación de la lógica modal cuantificada y su aparato filosófico, y recupera lo que puede ser recuperado de la noción intuitiva. Los demás elementos de la noción “intuitiva” (como la noción de *definición* o la de *explicación*) son remanentes que deben o ser eliminados o deben ser entendidos sólo sobre la base de la explicación extensional que ofrece la EXPLICACIÓN ESTÁNDAR DE LA ESENCIA.

En el siguiente capítulo revisaremos críticamente estos resultados. Argüiré que el ersatzismo y, de hecho más generalmente, toda teoría de la modalidad que comparte con el ersatzismo cierta estructura, tiene serios problemas; y recordaré las objeciones de Fine contra la definición modal de esencia. Revisaremos también la teoría de Fine, que reduce la modalidad a la esencialidad. Terminaré criticando esta reducción, y proponiendo una teoría de la modalidad de acuerdo a la cual ni la esencialidad ni la modalidad se reducen una a la otra (ambas son tomadas como *absolutamente fundamentales*), pero donde ambas nociones pueden convivir y, de hecho, librarnos de problemas filosóficos que surgen con su tratamiento estándar.

## 2. Esencialismo, modalidad y fundamentalidad

**A**quí me embarcaré en una revisión crítica de teorías que buscan fundamentar, o reducir, a la modalidad. Examinaré a la concepción ersatzista de la modalidad y su aliado, la definición modal de la esencialidad. También examinaré el reductivismo esencialista que surge del trabajo de Kit Fine. Argumentaré que ninguna opción fundamentadora o reductivista de las examinadas funciona. Si estoy en lo correcto, en el campo del realismo acerca de la modalidad pocas opciones teóricas propuestas en la literatura quedan fueran de este análisis.<sup>29</sup>

Quiero dejar claro que estaré asumiendo el *realismo* acerca de la modalidad—la idea de que los hechos acerca de la necesidad y la posibilidad, la contingencia o los subjuntivos, son hechos *independientes de nuestras mentes, lenguajes y decisiones*. Aunque no daré un argumento detallado a favor de esta asunción, me gustaría notar que tampoco es gratuita. Parece que nuestras teorías fundamentales postulan *espacios de posibilidad*: espacios en donde los puntos representan *estados o configuraciones posibles*, o donde las trayectorias representan *historias posibles*, de los sistemas. Estos espacios, a su vez, *están constreñidos por leyes*.

Por ejemplo, en física tanto clásica como cuántica, se utilizan espacios de estado, de configuración, o de fase para modelar la mecánica de sistemas físicos—tales espacios pueden ser variedades simplécticas en el caso de la mecánica clásica, espacios de Hilbert para la cuántica o espacios de Fock para la teoría cuántica de campos (Maudlin 2007, cap. 3; North 2012; Rickles 2008: §1.1; Sebens 2013; Smith 1998: cap. 1). Todos estos son *espacios de posibilidad*. Además, tales teorías suelen tener mayor poder explicativo que versiones “nominalizadas” que no apelan a espacios de posibilidad (Lyon & Colyvan 2008; North 2009). En general, cualquier teoría que postule probabilidad objetiva está comprometida con el realismo acerca de la modalidad.

Por otro lado, también los matemáticos están explorando una visión modal de los fundamentos, de acuerdo a la cual existe un *pluriverso* de universos conjuntistas, cada uno definido por las fórmulas indecidibles en ZFC que validan (Hamkins 2010; Hamkins & Löwe 2008; Mochizuki 2012, §3).<sup>30</sup>

Como vemos, diferentes teorías científicas postulan diferentes estructuras para sus espacios de posibilidad. Lo importante para mi proyecto es que, si uno acepta el realismo científico, uno está racionalmente compelido a creer que los espacios de posibilidad existen—que el mundo tiene, como repetidamente afirman Ladyman y Ross, una *estructura modal* que busca ser descubierta por la ciencia—y el éxito de esta muestra que tal estructura, a pesar del escepticismo filosófico, es descriptible y cognoscible (Ladyman & Ross 2007; DeLanda 2002, cap. 1; Giere 1985; cf. §4.1, abajo).<sup>31</sup>

29 Quizá las más notorias serían el reductivismo del realismo extremo de Lewis, o la idea (no reductivista) de parrear la necesidad con los contrafácticos (cf. Williamson 2007, cap. 5).

30 En realidad, el artículo de Mochizuki forma parte de una serie de artículos altamente especializados en teoría de números, cuyas pruebas todavía están siendo revisadas (<<http://www.nature.com/news/proof-claimed-for-deep-connection-between-primes-1.11378>>). De manera sorprendente, de todos modos, la mayoría de §3 sólo requiere un dominio no demasiado especializado de la teoría de conjuntos y lógica matemática.

31 Estrictamente hablando, falta también la premisa a favor de una *metodología naturalista*—de acuerdo a la cual nuestra mejor *guía* para la ontología es la ciencia: si una teoría exitosa postula una clase de entidades *E*, eso

La pregunta de si la ciencia postula esencias parece responderse de manera más rápida en lo negativo; aunque probablemente el debate pueda llevarse en términos de si deberíamos postular esencias en cualquier otra área teórica, incluyendo a la filosofía.<sup>32</sup> A final de cuentas, permaneceré neutral sobre la cuestión de si los objetos tienen o no esencias, y cuáles sean estas; pero mucho de lo que hago aquí resultará más interesante si algún tipo de esencialismo resulta verdadero.

Aquí voy a explorar la cuestión de si la modalidad es fundamental. Esta idea se puede entender de dos maneras. Si existen hechos acerca de si tal o cual cosa es posible o necesaria que no puedan explicarse apelando a otros hechos, parece que éso es buena razón para aceptar que la modalidad es fundamental. También, si la distinción entre lo que es posible y lo que no, o lo que es necesario y lo que no, no puede explicarse o definirse en términos que no sean modales, parece que la modalidad también resultará fundamental. Veamos.

## 2.1. reducción y fundamentación

Distingamos antes que todo a la reducción de la eliminación y la fundamentación. Una teoría *reductiva* acerca de los *Ks* encuentra un lugar en el mundo para los *Ks*, pero implica que se reducen a otras entidades (supuestamente más básicas); mientras que una teoría *eliminativa* acerca de los *Ks* implica que no hay *Ks* en absoluto (aunque tal teoría puede implicar que algo del papel teórico que iba a ser cumplido por los *Ks* puede ser llevado a cabo por otras cosas; tal cumplimiento debería ser suficientemente escaso, si es que la supuesta eliminación no ha de colapsar en una reducción.) Finalmente, una teoría acerca de los *Ks* también acepta la existencia de *Ks*, pero los toma como irreducibles—en lugar de ello, su existencia ha de ser *explicada*, en un sentido distintivamente metafísico, al apelar a otras entidades (supuestamente más básicas). Aquí dejaré de lado teorías eliminativistas de la modalidad. Me enfocaré en teorías reductivistas y de fundamentación.

Una teoría reductiva acepta la existencia de lo que busca reducir, pero nos dice que tal existencia es *nada más que*, o *simplemente es*, alguna cosa que se asume más básica. Que afirmaciones con este tipo de terminología sean implicados por las teorías en cuestión, muestra que cualquier relación a ser usada en una reducción debería al menos ser asimétrica—lo que se reduce no puede también formar una base reductiva para su propia base reductiva (asumiendo la univocidad de “reducción”).

---

constituye un argumento a favor de la necesidad racional de aceptar la existencia de *E*.

Si he de confesarlo, prefiero una metodología naturalista a la que parece ser su rival más importante en la filosofía analítica—la idea de diseñar teorías metafísicas constreñidas por las intuiciones lingüísticas que son traídas a la superficie al imaginar casos (que *se asumen* ser) posibles. Simplemente no veo por qué nuestras intuiciones lingüísticas son mejores guías para la ontología que la ciencia (véase §5.1, abajo, sobre esto). Tampoco puedo ver por qué consideraciones sobre economía, cuando están motivadas por el ideal de una ontología o ideología desérticas sin mayor fundamento que la intuición, deban pesar sobre consideraciones de utilidad teórica y poder predictivo.

32 Véase Romero 2014, donde reviso la literatura relevante, que parece mostrar que las esencias son necesarias para bloquear la paradoja de Chisholm en la metafísica modal. Véase también Ladyman & Ross 2007: 133, nota 5, y §5.6.

Ahora bien, si es que los hay, los casos no controvertidos de reducción son extremadamente escasos en la filosofía, por lo que una dificultad inicial es encontrar un caso paradigmático, respecto al cual los parecidos podrían ser evidencia de una reducción exitosa.

Consideremos a las mesas. La mayoría de los metafísicos se sienten incómodos con la idea de que las mesas son fundamentales a la realidad, y por ello son llevados a pensar que las mesas de alguna manera deben reducirse a otras entidades más fundamentales. Una idea elemental, que servirá para aclarar, es que, cualesquiera que sean las unidades físicas más pequeñas que haya (las llamaremos “partículas”), ellas están acomodadas de tal manera que *constituyen* mesas—están acomodadas “meseísticamente”.<sup>33</sup> Así que las mesas se reducen a partículas acomodadas meseísticamente (y se asume que, en la teoría final, “meseísticamente” y términos parecidos serán definidos en un vocabulario que no contenga términos como “mesa”).

En general, entonces, y en un primer acercamiento, una teoría reductiva acerca de entidades individuales implica que estas están constituidas por, compuestas de, o en general “construidas de” entidades individuales más básicas—las mesas de partículas, los números naturales de conjuntos. Pero regularmente se busca también la reducción de cosas no individuales; a veces se busca reducir relaciones o propiedades (y sin necesidad de asumir una teoría de tropos). En este caso, la reformulación necesaria es simple: las propiedades y relaciones se pueden reducir a otras propiedades y relaciones—por ejemplo, la propiedad *ser una muestra de agua* se reduce (o eso se asume) a la propiedad *ser una muestra de H<sub>2</sub>O*. En general, si  $x$  es una cosa que puede ser individuada (ya sea abstracta o concreta; general o individual), se puede decir que  $x$  se reduce a otras cosas  $y_1, \dots, y_n$ , cuando hay una relación reductiva entre  $x$  y  $y_1, \dots, y_n$ , como constitución, o identificación teórica.

Aquí uno podría tener dudas generales acerca de la relación de reducción con la identidad. Quizá la constitución es identidad; quizá, cual sea la relación reductiva entre los dos, todo número natural es un ordinal de von Neumann, y seguramente la identificación teórica es identidad. ¿Dónde, entonces, está la reducción?

Estas preguntas son mejores para un estudio mucho más detallado del que se puede hacer aquí. Pero una idea que estos casos sugieren es que, al menos en un importante rango de casos—aquellos en los que lo reducido se identifica con lo que se reduce—las teorías reductivas, cuando son exitosas, son mejor concebidas como teniendo una función primariamente *semántica*: tenemos un fragmento de discurso,  $D$ , que menciona mesas o números, y reducir mesas o números a partículas u ordinales de von Neumann es simplemente encontrar un conjunto de hacedores de verdad apropiados para  $D$  que sólo involucre a partículas u ordinales de von Neumann (cf. Dowell, 2008: §2; Maguire 2013: §1; cf. Jackson 1998: cap. 1 para una tesis relacionada). Si es exitoso, el reductivista afirmará que el mundo sólo necesita contener a la base reductiva para contener lo que se reduce: una vez que la existencia y organización meseística de las partículas está en su lugar, las verdades acerca de las mesas están en su lugar; una vez que la existencia de ordina-

---

33 Quizá no haya tales unidades físicas últimas. Quizá la cadena de fundamentalidad se acabe en el todo del espaciotiempo y no en ninguna de sus partes propias. Quizá la afirmación de que la cadena de fundamentalidad “se acaba” en algún tipo de *particular* está mal. Quizá incluso la idea misma de que el nivel más fundamental contienen *cosas* (propiedades o haces de estas, hechos, o particulares) es incorrecta, y la fundamentalidad es una cuestión de ideología (en el sentido quineano) o de representación. Pongo tales dudas entre paréntesis por propósitos expositivos.

les de von Neumann está en su lugar, las verdades acerca de los números naturales también.

Por supuesto, el compromiso con los hacedores de verdad no es universal; pero aquí podemos encontrar una fructífera aplicación de ellos que podría contar como razón para aceptarlos. De hecho, la teoría de los hacedores de verdad nos permite generalizar todavía más estas ideas acerca de la reducción. Pues los filósofos han buscado reducir cosas que no se pueden individuar *como una cosa*. Por ejemplo, algunas creen que la moralidad se reduce ultimadamente a la Naturaleza, o la modalidad a los mundos posibles, o el agua al H<sub>2</sub>O. Pero la moralidad, la modalidad y el agua no son *una cosa*, no son algo que puedas contar (hay *tipos* de modalidad, claro: nómica, o metafísica; pero no, hablando estrictamente, “modalidades”; hay *sistemas* morales, pero no, “moralidades”; hay *muestras* de agua, o ríos o mares u océanos, pero no, hablando estrictamente, “aguas”). En casos como estos, de nuevo es natural apelar a hacedores de verdad: el hacedor de verdad para cada *verdad* moral es un *hecho* puramente natural; el hacedor para cada *verdad* modal es un hecho acerca de mundos; el hacedor de verdad para cada verdad acerca del agua es un hecho acerca de las moléculas de H<sub>2</sub>O. Así que aceptar una ontología con lo que sea que se necesita para una teoría de los hacedores de verdad nos da un atractivo recurso para describir la estructura general de las teorías reductivas: una teoría que busca reducir *R* a *B* implica que cada proposición acerca de *R* se hace verdad por algún hecho acerca de *B*.<sup>34</sup>

Este acercamiento a la reducción especifica la noción de *reducción* que será asumida aquí. Con lo que una teoría reductiva de la modalidad, que busca reducir la modalidad a otro tipo de cosas *R*, tendrá la siguiente forma:

(TEORÍA REDUCTIVA)

Para toda *oración* verdadera de la forma  $\Box p$  o de la forma  $\Diamond p$ , hay un hecho *h* tal que:

(R1: *No ideología modal*) *h* no necesita de ideología modal para ser completa y perspicuamente caracterizado.<sup>35</sup>

(R2: *Hecho R*) *h* es un hecho acerca de cosas *R*.

(R3: *Hacedores de verdad*)  $\Box p$  ( $\Diamond p$ ) es hecha verdad por *h*.

Veamos una justificación para cada uno de los tres requisitos.

34 Se podrían considerar otros acercamientos a la naturaleza de la reducción. En particular, las nociones de reducción inter-teórica de Nagel (en términos de leyes puente y el modelo nomológico-deductivo de la explicación) o el de Fodor (en términos de leyes puente y propiedades físicas disyuntivas y básicas). Tales acercamientos tienen problemas bien conocidos en la literatura (cf. Bueno 2009). Otro acercamiento posible a la reducción inter-teórica sería el micro-reduccionismo de Oppenheim y Putnam (1958). De cualquier manera, mi apuesta es que la propuesta reductivista examinada aquí falla incluso si tales ideas sobre la reducción pueden aplicarse a la relación entre necesidad y esencialidad. No tengo una *prueba* de ello; pero creo que, dadas las dificultades presentes, la carga de la prueba estará sobre el reductivista de mostrar que alguna otra *reducción* bloquea los argumentos que daré abajo.

35 Por “ideología metafísico-modal” (cuando no haya peligro de confusión, simplemente hablaré de “ideología modal”), quiero decir cualquier ítem sub-oracional tal que: (a) es el operador “caja” o “diamante” de la lógica modal, interpretado como la posibilidad o necesidad metafísica, o puede ser definido en términos de ellos (y otra ideología lógica); o (b) su interpretación intencional se da en términos de la posibilidad o necesidad metafísica. Así, por ejemplo, el operador de compatibilidad, *C*, definido como:  $C(p, q)$  sii  $\Diamond(p \ \& \ q)$ , es una pieza de ideología modal, como lo son los operadores  $\Diamond_n$  (para *n* natural), interpretados como:  $\Diamond_n(p)$  es verdad sii hay al menos *n* mundos metafísicamente posibles donde *p* es verdad; mientras que los operadores de la lógica temporal de tiempo lineal, por ejemplo, no son ideología metafísico-modal en este sentido (aunque los operadores de la lógica temporal de tiempo ramificado podrían ser).

Un reductivista no *debe* necesitar apelar a ideología modal para caracterizar sus hechos básicos (R1), pues eso nos daría razones para pensar que tales hechos son modales por sí mismos. A lo más, tales hechos *podrían* ser caracterizados con ideología modal, sólo si el reductivista nos da razones para pensar que tal caracterización es, a final de cuentas, dispensable.

También podemos argumentar por analogía. Ninguna teoría reductiva acerca de las mesas podría *necesitar* caracterizar a las partículas y sus relaciones, a las que se reducen las mesas, en un lenguaje que menciona a las mesas de una manera u otra. Una teoría que no pudiera definir lo que es un ordinal de von Neumann sin apelar a los números naturales no podría afirmar reducir los números naturales a los ordinales de von Neumann. Para una analogía más cercana a nuestro tema, recordemos a Lewis—quien afirmó (1986) que cada verdad de la forma  $\Box p$  es hecha verdadera por sus mundos:  $\Box p$  ( $\Diamond p$ ) es verdad sii para todo (algún) mundo  $w$ , restringir los cuantificadores de  $p$  a  $w$  resultará en una proposición verdadera. Así que si para definir qué es un mundo, y qué mundos hay, uno no necesita ninguna ideología modal, y si la coextensividad arriba mencionada se da, la lewisiana ya tiene una teoría reductiva de la modalidad. Si, por el contrario, al postular tal pluriverso se necesita una noción modal sin reducir, la anti-reductivista le gana al lewisiano (Lycan 1979; Shalkowski 1994). De manera similar, ninguna teoría reductiva de la modalidad puede especificar los hechos acerca de las esencias que forman la base reductiva pretendida con ideología modal ineliminable.

El caso del realismo lewisiano también puede usarse para justificar, por analogía, la necesidad de cuantificación al inicio de TEORÍA REDUCTIVA. El lewisiano tiene que asegurarse que para cada verdad modal haya un hecho correspondiente acerca de mundos lewisianos. Pues si alguna verdad modal no tuviera un hecho correspondiente, esto mostraría que las verdades modales no necesitan ser hechas verdaderas por los mundos lewisianos (Divers & Melia 2002), y por ello la idea de que los mundos lewisianos de hecho satisfacen ese papel en algunos casos se pondría en duda: si al menos una verdad modal no es hecha verdad por ningún hecho acerca de los mundos, ¿qué motivación hay para decir que la verdad modal se *reduce* a los hechos acerca de los mundos? Así que el lewisiano—y, por analogía, cualquier reductivista—debe afirmar que *toda* verdad modal es hecha verdad por la base pretendida.

Consideremos ahora a R3. Tenemos razones para tal requisito, primero, si hemos aceptado las consideraciones de arriba que vinculan a la reducción con el hacer verdad (dejare de lado los detalles de la teoría relevante de hacedores de verdad). Ahora supongamos que  $h$  se da pero que  $\Box p$  (o  $\Diamond p$ ) no es verdad. Entonces  $h$ —y, en general, los hechos acerca de  $R$ —no bastan para la verdad modal. Pero entonces no se logra ninguna reducción: si las partículas acomodadas meseísticamente no bastan para fijar las verdades acerca de las mesas, entonces las mesas son algo más allá de tales partículas, y las mesas no se reducen a tales partículas—se necesita algo más que las mesas para que haya mesas.

Queda R2. Está claramente bien: una teoría reductiva toma a las cosas  $R$  como base sólo si la base reductiva tiene algo que ver con tales cosas.

En breve, hay buenas motivaciones para tomar a TEORÍA REDUCTIVA como la estructura que toda teoría reductiva de la modalidad debe tomar. Ahora pasemos al tema de la fundamentación.

Notemos primero que los teóricos de la fundamentación no intentan encontrar lo fundamentado entre lo que lo fundamenta. Lo fundamentado es distinto de sus fundamentos—pero la existencia de lo fundamentado es *fijada* por la existencia de sus fundamentos; no en el sentido de

una mera implicación, sino además en el sentido de una *explicación* de un tipo distintivamente metafísico. Así que la fundamentación también es una relación asimétrica (y tiene una clausura transitiva natural), pero es diferente de la reducción en que lo fundamentado tiene una existencia distinta—fijada y explicada por los fundamentos. Así que tenemos una noción bastante mínima de fundamentación:

(TEORÍA FUNDAMENTADORA)

Para todo hecho especificable con la forma  $\Box p$  o la forma  $\Diamond p$ , hay un hecho  $h$  tal que:

(G1: *No ideología modal*)  $h$  no necesita de ideología modal para ser completa y perspicuamente caracterizado.

(G2: *Hecho R*)  $h$  es un hecho acerca de cosas  $R$ .

(G3: *Vínculo explicativo*) Que  $\Box p$  ( $\Diamond p$ ) se dé es *explicado* por que  $h$  se dé, y de tal manera en que el que  $h$  implica que  $\Box p$  ( $\Diamond p$ ).

La justificación para G1 es análoga a la de R1. Si, para especificar  $h$ , necesitamos apelar a la ideología modal, entonces eso seguramente significa que, por todo lo que sabemos,  $h$  mismo es un hecho modal. La justificación para G2 es más clara; paso a G3. Que los fundamentos explican a lo fundamentado es universalmente compartido entre los teóricos de la fundamentación. De hecho, la noción de fundamentación se introdujo precisamente para hacer el trabajo explicativo que la mera superveniencia no pudo hacer (Clark & Liggins 2012; Schaffer 2009). También es ampliamente aceptado que los fundamentos implican (la existencia de) lo que está fundamentado (Fine 2012; Rosen 2010). Así que G3 incorpora asunciones populares sobre la fundamentación en general; y simplemente no parece haber razón alguna de por qué tales asunciones no deberían mantenerse en el caso de la modalidad.

Finalmente, notemos la diferencia entre TEORÍA REDUCTIVA y TEORÍA FUNDAMENTADORA. De acuerdo a la primera, la reducción se da entre fragmentos de discurso—conjuntos de oraciones—y sus hacedores de verdad: que las mesas se reduzcan a partículas es simplemente que todo el discurso sobre las mesas sea hecho verdad por hechos sobre partículas. Mientras tanto, de acuerdo a TEORÍA FUNDAMENTADORA, la fundamentación se da entre hechos: un hecho está fundamentado en otro cuando este explica e implica a aquel. Diferentes teorías sobre las oraciones, el hacer verdad, y los hechos, podrían resultar en que fundamentar y reducir resultan en lo mismo. Sin embargo, esto sería una consecuencia no particular para mi acercamiento, así que los argumentos que brindo ya sea contra la reducción o la fundamentación seguirían siendo aplicables.<sup>36</sup>

Habiendo entendido las nociones clave de reducción y fundamentación, podemos evaluar a las teorías sobre la modalidad que la buscan reducir o fundamentar.

## 2.2. teorías ‘no-no’ de la modalidad

En §1.2 individué varias tesis que parecen darnos una visión suficientemente general de las teorías acerca de la modalidad; ahora las repito:

<sup>36</sup> Por ejemplo, reducción y fundamentación podrían ser lo mismo si resulta que: (i) el contenido de una oración es una proposición; (ii) una proposición verdadera es lo mismo que un hecho; (iii) la relación de *hacer verdad* es la misma que la *fundamentación*, y esta se da entre hechos. Agradezco a Alessandro Torza y a Jennifer Wang el haberme hecho notar la necesidad de estas aclaraciones.

MUNDISMO. Los mundos posibles son la herramienta fundamental (ya sea en un uso ontológico, uno representacional, o ambos) para la investigación de la modalidad.

ABSTRACCIONISMO. Los mundos posibles son entidades *abstractas*.

ACTUALISMO. Los mundos posibles son entidades reales, que existen *de hecho* (*actually*); no hay nada meramente posible: todo lo que existe, existe de hecho: todo lo que existe, tiene *actuality*.

REPRESENTACIONALISMO. Los mundos posibles *representan* meras posibilidades.

IDENTIFICACIONISMO. Los mundos posibles *son* las posibilidades.

Encontramos que el ersatzismo cumple con MUNDISMO, ABSTRACCIONISMO, y ACTUALISMO, mientras que muchas veces no era claro si los filósofos ersatzistas deseaban cumplir con REPRESENTACIONALISMO o con IDENTIFICACIONISMO—pues no siempre es claro si buscan darle un papel ontológico o meramente representacional a sus mundos. (Algunas veces sí lo es, por supuesto: ver Melia, 2001).

Al leer la literatura “clásica” (de los 70’s para acá), uno tiene la impresión de que los mundos posibles de hecho se invocan con un fin ontológico: para explicar o “iluminar” la naturaleza de la modalidad. Encontramos, por ejemplo, las siguientes afirmaciones (los énfasis son míos):

Al explorar y *explicar* la naturaleza de la necesidad, Leibniz se vuelca a la idea de los mundos posibles; nosotros no podemos superarlo. (Plantinga 1974: 44)\*

Suplementé mi argumento contra *S4* con una concepción particular de qué son los mundos posibles—en conjunción con la *identificación estándar* de la necesidad con la verdad en todo mundo posible y la posibilidad con la verdad en al menos un mundo posible. (Salmon 1989: 5)\*\*

*¿En virtud de qué es que los enunciados metafísicamente modales son verdaderos o falsos? [...] Prominente entre las varias respuestas a estas preguntas son los análisis de la modalidad en términos de mundos posibles. En este acercamiento, hay algunas entidades, los “mundos”, tales que las oraciones acerca de los objetos ordinarios son verdaderas o falsas “en” [in] o “en” [at] ellos. (Roy 1993: 335)\*\*\**

Los filósofos utilizan a los mundos posibles para llevar a cabo diferentes tipos de trabajo. Uno de estos trabajos es el de explicar *en qué consiste* la verdad de las verdades modales y la falsedad de las falsedades modales. “¿En qué consiste la posibilidad de que *p*?”, preguntas. “Consiste en la existencia de un mundo en el que *p*”, alguien podría responderte.

---

\* “In exploring and explaining the nature of necessity, Leibniz turns to the idea of possible worlds; we can do no better.”

\*\* “I supplemented my argument against *S4* with a particular conception of what possible worlds are—in conjunction with the standard identification of necessity with truth in every possible world and possibility with truth in at least one possible world.”

\*\*\* “What is it in virtue of which metaphysically modal statements are true or false? [...] Prominent among the various responses to these questions are the possible worlds analyses of modality. On this approach, there are some entities, the “worlds,” such that nonmodal sentences about ordinary objects are true or false “in” or “at” them.”



(Rayo 2012)\*\*\*\*

La idea de que la verdad relativa a mundos posibles sea lo mismo que verdad necesaria, parece apuntar a que se hace cierta asunción: que los mundos posibles también han de usarse con un fin ontológico, es decir, para dar una teoría metafísica de la modalidad (a diferencia de utilizarse como recurso heurístico en cualquier otro tipo de teoría). Esto hace que sea más difícil saber si los ersatzistas están asumiendo IDENTIFICACIONISMO O REPRESENTACIONALISMO.

Pero incluso aunque se suponga que los mundos posibles han de jugar un papel ontológico en la metafísica de la modalidad, *este no puede ser completo*. Pues, como es ampliamente reconocido, la teoría ersatzista de los mundos posibles *no* logra un análisis de la ideología modal en términos no modales. El argumento es bien conocido: no toda construcción ersatzista (conjunto, modelo, etc.) corresponderá a una posibilidad. Hay dos razones para esto: que haya construcciones que correspondan a *imposibilidades*, o que haya posibilidades que no se correspondan con ninguna construcción. Bajo cualquiera de las dos opciones, el ersatzista necesitará apelar a criterios *modales*, ya sea para evitar que sus mundos resulten ser imposibles, o para lograr que a toda posibilidad le corresponda uno de sus mundos. Pero si el ersatzista apela a tales criterios modales, entonces su teoría no es reductiva, pues no ha logrado eliminar toda ideología modal a favor de sus mundos posibles suplementados con ideología no modal (Lewis 1986: §3.2; Sider 2002: §2, 2003: §3.2).

A la segunda objeción se le puede encontrar respuesta, pues de hecho ya se han ofrecido teorías ersatzistas que tienen el suficiente poder expresivo como para corresponder a posibilidades *alien* (Rayo 2012, Melia 2001 o Sider 2002, (aunque sobre esta, véase Leuenberger 2006)). Sin embargo, todas estas involucran aceptar que se necesitan criterios modales, como dice la primera objeción.

Conozco una manera de intentar responder al primer argumento, sin aceptar criterios modales primitivos, que es la de Heller (1998a, b). Para Heller, los mundos posibles son simplemente conjuntos de tuplas de números reales, que *representan* una distribución de propiedades sobre una variedad que se toma como el espaciotiempo. Tales propiedades sólo existen en un mundo, y para evitar el esencialismo extremo, Heller adopta una teoría de contrapartes para propiedades.

Ahora bien, es claro que la representación en cuestión es *impuesta*: los conjuntos relevantes no representan *por sí mismos*, sino que necesitan que *alguien* les de esa función. Como dice el mismo Heller (1998a: 296):

Por supuesto, estos conjuntos no representan intrínsecamente, sino sólo relativo a una interpretación. Esto no significa una dificultad mientras las reglas de interpretación pueda ser especificadas en un vocabulario finito—y sí pueden.\*

Sin embargo, al asumir que la existencia de los mundos posibles es *fijada* por *nuestros* actos de interpretación, Heller parece comprometerse con un fuerte tipo de anti-realismo acerca de la necesidad. Por supuesto, él considera esta objeción (p. 314). Pero, según él,

El error en este reto es asumir que cuáles conjuntos cuenten como mundos en un mundo

---

\*\*\*\* "Philosophers call on possible worlds to perform different kinds of jobs. One of these jobs is that of explaining what the truth of modal truths and the falsity of modal falsities *consists in*. "What does the possibility that *p* consist in?", you ask. "It consists in the existence of a possible world at which *p*", someone might reply."

\* "Of course, these sets do not represent intrinsically, but only relative to an interpretation. This poses no difficulty as long as the rules of interpretation can be specified in a finite vocabulary, which they can be."

$w$  depende de los interpretadores en  $w$ . [...] Lo que *nosotros* queremos es algunos objetos que estén bien equipados para el papel de los mundos posibles. Esos objetos existen incluso en los mundos sin intérpretes; y son mundos posibles porque somos *nosotros* quienes los usamos como tales. [...] Queremos ser capaces de decir con verdad que, si nadie hubiera interpretado alguna vez a los conjuntos como mundos posibles, entonces los conjuntos no hubieran representado posibilidades alternativas. Esto es verdad porque los conjuntos en un mundo sin intérpretes no habría, *en ese mundo*, representado. Pero esto es consistente con decir que en ese mundo es verdad que diferentes cosas podrían haber pasado. Esto último es verdadero porque incluso en ese mundo hay mundos—es decir, conjuntos bajo *nuestra* interpretación—en los que pasan otras cosas. (315, énfasis originales)\*\*

Confieso que simplemente no veo cómo se supone que esto resuelve el problema. La clave está en las frases “y son mundos posibles porque somos *nosotros* quienes los usamos como tales” y “[...] hay mundos—es decir, conjuntos bajo *nuestra* interpretación[...]”. Si el que algo—un conjunto, lo que sea—sea un mundo posible está fijado por cómo *nosotros lo usamos*, y si se asume que la modalidad está fijada por los mundos, bueno, no sé cómo eso *no* implica que la modalidad no está fijada *por lo que nosotros hacemos*. Si esto no es un tipo fuerte de anti-realismo, realmente no entiendo qué es el anti-realismo.

Ahora bien, al inicio del capítulo anuncié que me quedaría dentro del campo realista. Así que, por más interesante que pueda ser la teoría de Heller para un anti-realista, aquí no la voy a considerar más.

Regresemos al ersatzismo no-reductivo, y consideremos a aquellos ersatzistas que todavía buscan dar una *explicación* de la modalidad en términos de sus mundos. Diré de ellos que su teoría es *no absolutista*, en el sentido en que no asume que la modalidad es *absolutamente* fundamental—pues asume que la explicación metafísica de la modalidad se da, parcialmente, gracias a la existencia de los mundos posibles. Pero la teoría de tal ersatzista tampoco es una teoría *reductivista*, pues *no* asume que el aspecto modal del mundo se reduzca a su aspecto no modal—pues ha aceptado que para especificar sus mundos tiene que apelar a criterios que esencialmente contienen ideología modal.

Ahora consideraré más de cerca estas ideas.

---

\*\* “The mistake in this challenge is to assume that which sets count as worlds in a world  $w$  depends on the interpreters in  $w$ . [...] What *we* want is some objects that are well suited to the role of possible worlds. Those objects exist even in the worlds with no interpreters; and they are possible worlds because it is *we* who are using them as such. [...] We do want to be able to say truly that, if no one had ever interpreted sets as possible worlds, then sets would not have represented alternative possibilities. This is true because the sets in a world devoid of interpreters would not, *in that world*, have represented. But this is consistent with saying that in that world it is true that different things could have happened. This latter is a truth because even in that world there are worlds—that is, sets under *our* interpretation—in which other things happen.”

## 2.3. problemas con las teorías *no-no* de la modalidad

### 2.3.1 problemas con el ersatzismo lingüístico

El ersatzismo no es la única teoría no absolutista. El objetivo principal de lo que sigue en este capítulo es argüir que, de hecho, todas las teorías no absolutistas que además—como el ersatzismo—no sean reductivistas, enfrentan serios problemas. Llamaré a tales teorías, *teorías no-no* de la modalidad—pues *no* son absolutistas, pero *tampoco* son reductivistas.

Una vez que el ersatzista ha aceptado que su teoría no reduce a la modalidad, le queda la opción de todavía buscar *fundamentar* a la modalidad en los mundos posibles. Ahora argumento en contra de este movimiento.

El ersatzista no-reductivista que toma a las posibilidades como estando *fundamentadas* en los mundos posibles, dirá que los mundos posibles explican, de una manera distintivamente metafísica, la existencia y naturaleza de las características modales. Tal explicación sólo puede ser parcial, pues una explicación completa de la modalidad en términos de mundos posibles requeriría que estos no necesitaran ser definidos usando ideología modal. Pero entonces tal fundamentación sólo será *parcial*: no todo acerca de la modalidad podrá ser explicado con base en mundos posibles. Esto suena razonable: una vez que se acepta que los mundos posibles necesitan ser definidos modalmente, parece que se acepta también que el contenido de esa ideología modal con la que se los define no es explicada *por esos mismos mundos posibles*. Así, hay criterios modales que seleccionan algunas, pero no todas, de las construcciones ersatzistas como *mundos posibles*.

El problema entonces es que el ersatzista no puede dar cuenta de la existencia y el funcionamiento de los criterios modales en cuestión. Supongamos que un criterio modal fuera: *nada es un mundo posible a menos que represente a los axiomas de la teoría de conjuntos ZFC como siendo verdaderos*. Sea tal criterio, “Zee”. El ersatzista no puede, a menos que contradiga su afirmación de que el ersatzismo no es reductivo, buscar reducir la verdad *necesaria* de Zee a hechos acerca de mundos, pues se asumía que había criterios modales, Zee entre ellos, que eran tanto básicos como determinantes del rango de lo posible. Pero entonces tenemos todo el derecho de preguntar qué es lo que hace que Zee sea *necesariamente* verdad, si no los mundos posibles. El ersatzista podría responder al apuntar a lo primitivo de la necesidad de Zee; pero entonces uno se pregunta por qué, entonces, cualquier otra verdad modalizada debería estar fundamentada en los mundos posibles, si Zee ya por sí misma no está fundamentada así y de todos modos es necesaria. Además, Zee debe tomarse como bien definida y bien entendida—si no, el ersatzista tendrá problemas determinando el rango de lo que es posible desde el mero principio.

Afirmo que, si el ersatzista no puede dar cuenta de la existencia y funcionamiento de los criterios modales relevantes, esto nos da una motivación para rechazar el proyecto de fundamentar a la modalidad en los mundos. Pues, si una verdad o característica modal puede estar bien definida y ser completamente inteligible sin referencia a ningún mundo posible en absoluto, ¿por qué no podría *toda* característica modal tener ese estatus? Una vez que hemos visto cómo este tema se generaliza a las teorías no-no, razones análogas nos darán motivación para rechazarlas.

Una segunda manera de debilitar al ersatzismo (después de aceptar que no es reductivo, pero

que fundamenta parcialmente a la modalidad), es aceptar la primitividad absoluta de la modalidad, pero todavía tomar a las construcciones ersatzistas como una *representación fiel* de la estructura profunda de la modalidad (este camino, como apunté en secciones anteriores, parece ser el preferido de varios ersatzistas). La motivación para esta postura es la idea de que los mundos ersatz sí arrojan alguna luz sobre la estructura de la modalidad: como se suele subrayar, es mucho más fácil argumentar a favor de la transitividad de la relación *ser posible relativo a*, que se da entre mundos posibles, que argumentar por la validez de “si algo es posiblemente posible, entonces es posible *simpliciter*”. La representación extensional que nos da el aparato mundista y la semántica de marcos kripkeanos es simplemente más manejable (quizá porque nos ayuda a poner intuiciones espaciales en juego). Bajo esta variedad, el ersatzismo nos *invita* a pensar a las posibilidades como hechos completos y no-contradictorios, y a dibujar a lo que sea posible como siendo, en algún sentido, “parte” de uno de estos *grandes* hechos. Pero esta invitación no es una invitación a reducir, o fundamentar, posibilidades en mundos posibles.

Ahora, la ersatzista que no busque eliminar, o hacer dependiente de la mente, a la modalidad, aceptará que el aspecto modal es o básico o reducible o fundamentado. Si es reducible, la ersatzista nos debe una garantía de que su aparato mundista es fiel a cómo se comporta la base reductiva: la base podría inducir un sistema representacional por sí sola, incluso distinto a un sistema ersatzista de mundos—los hechos “grandotes” podrían no ser todo lo que hay en la estructura de la modalidad. Algo análogo sucede si la modalidad se fundamenta en otra cosa. Y, si la modalidad es básica, entonces la ersatzista nos debe un argumento para pensar que su representación del aspecto modal del mundo mediante su sistema de mundos es fiable. ¿Por qué?

Bajo la imagen ersatzista, la representación del aspecto modal del mundo se logra “dentro” de cada mundo por referencia a otros mundos, la relación de accesibilidad y la estructura que induce—en el sentido en que lo que un mundo posible “toma” como siendo posible está determinado por qué otros mundos posibles existen, cómo están relacionados a nuestro primer mundo, y qué portadores de verdad son hechos verdaderos por esos otros mundos. Esta imagen no puede ser *completamente* fiel, pues *se basa en una reducción que nunca se logró*. Si hay más en la modalidad que los mundos posibles y sus relaciones, entonces lo que sea el extra, no está siendo completamente representado en un mundo simplemente al apuntar a su relación con otros mundos. Así, los mundos no pueden representar fielmente los constreñimientos modales por todos y sólo aquellos portadores de verdad que sean verdaderos en cada mundo. Esto es porque ya se había concedido que la modalidad va más allá de los mundos: la verdad necesaria *no* sólo verdad en cada mundo; la existencia posible *no* es solamente la existencia en algún mundo, etcétera. El punto acerca de la representación de criterios modales también se generaliza a otras teorías no-no, como veremos.

Ahora bien, la teoría representacional que yo propondré en §3.2, abajo, *sí* que hace uso de mundos posibles como *dispositivos representacionales*. La diferencia con la imagen ersatzista que acabo de atacar es que los mundos posibles, en mi teoría, no se asumen como representando de manera fiel, *completa y perspicua*, a la modalidad—más bien, estaré usando una “función modal” que está diseñada para codificar, explícitamente, cualesquiera constreñimientos modales que estén en juego.

Pasemos ahora a ver cómo se generalizan los puntos de arriba.

## 2.3.2 generalizando, y el caso de la teoría de Jacobs

Como anticipé, hay puntos importantes, argumentados con respecto al ersatzismo, en su variedad no-no, que se generalizan para otras teorías de la modalidad.

PRIMERA. Una teoría no-no de la modalidad  $T$  tomará a la modalidad como *parcialmente fundamentada* en un tipo de entidades supuestamente más básicas  $E$  y por ello, por definición, tomará a la modalidad como siendo parcial pero no totalmente fijada por las cosas  $E$ . Para fundar aquello de la modalidad que no es fundado por las cosas  $E$ , invocará un criterio modal  $C$  para seleccionar, entre las cosas  $E$ , aquellas que fundamentan parcialmente la posibilidad de cualquier cosa posible, como opuesto a aquellas que no tienen tal función fundadora.

SEGUNDA.  $T$  tendrá que explicar cómo funciona  $C$ : ¿por qué es que sólo ciertas  $E$ -entidades son seleccionadas por  $C$  como fundando posibilidades, y no otras?

TERCERA. No puede afirmarse que la corrección de la selección hecha por  $C$  se da en virtud de las  $E$ -cosas, pues aquellas que fundamentan a la modalidad tienen que ya haber sido seleccionadas por  $C$ ; y entonces, al afirmar que las cosas  $E$  fundamentan a  $C$ ,  $T$  sería circular.

CUARTA. Si  $T$  evita la circularidad al postular que  $C$  es primitivo, esto resulta en postular modalidad absolutamente fundamental—contrario al espíritu de  $T$ , que buscaba *fundar parcialmente* a la modalidad en las cosas  $E$  (y con ello rechazar a la modalidad *absolutamente* fundamental).

Las tesis PRIMERA a TERCERA conjuntamente sugieren que o hay circularidad involucrada en la estructura de las teorías en cuestión, o que injustificadamente dejan algunos hechos sin explicar. CUARTO entonces nos dice que la única ruta de escape disponible a quien favorezca  $T$  es rechazar a la modalidad parcialmente fundada a favor de la modalidad absolutamente fundamental.

Con respecto a rechazar el alcance metafísico de una teoría no-no en favor de utilizar como una teoría representacional para la modalidad fundamental, arriba argumenté instancias de los siguientes puntos:

QUINTA. Rechazar el alcance metafísico de una teoría no-no  $T$  y en lugar de ello tomar a  $T$  como una teoría representacional para la modalidad fundamental, implica aceptar que hay más acerca de la modalidad que la existencia y características de las cosas  $E$ .

SEXTA. El defensor de  $T$  entonces tiene la carga de la prueba de mostrarnos que las entidades  $E$  son suficientes para dar una representación fiel de la naturaleza de la modalidad, pues las distinciones modales básicas no están fundamentadas ni se reducen a las cosas  $E$ , y por ello la estructura de estas podría no corresponder a la estructura de la modalidad.

Esbozemos una manera en que estos puntos se aplican a otras posturas. Para fijar ideas, tomaré el caso de la teoría de las “potencias” de Jacobs. De acuerdo a él (2010: 233-4):

Las propiedades y las relaciones entre ellas introducen conexiones modales en el mundo real [...] La postura es no-reductiva, pues las propiedades (o las relaciones entre ellas) son intrínsecamente modales. La postura es una forma fuerte de *actualismo*, pues toda la modalidad es parte del tejido fundamental del mundo real. [...]

Hablemos, de una manera general y muy abstracta [...] de las propiedades P y Q y la relación R entre ellas. P podría ser *es un caballo*, Q, *es un animal* y R, *implica* [...]. Consideremos todas las afirmaciones de la forma R(P, Q) [...] Tomadas en conjunto, todas estas afirmaciones representan la estructura de propiedades en el mundo real [*actual*], una red de propiedades interrelacionadas, por así decir. La teoría de la modalidad basada en propiedades, y todas sus versiones específicas, cuelga toda la modalidad en esta estructura de las propiedades. Alguna proposición T es posible sólo en caso de que haya alguna propiedad real P (o alguna colección de propiedades) conectadas por R a alguna otra propiedad Q (o colección de propiedades) tal que Q es un hacedor de verdad para T, o hay una cadena de tales conexiones.\*

Aunque aquí se habla del mundo como siendo *intrínsecamente modal*, o de la modalidad *siendo parte del tejido fundamental del mundo*, argumento que esto no debería confundirnos. Esta postura no implica, estrictamente hablando, que el mundo es modal *por sí mismo*. Más bien, es porque *hay conexiones modales entre propiedades*, de acuerdo a estas ideas, que hay modalidad en el mundo. Aquí, *la modalidad necesita conexiones entre propiedades*; justo como, en el uso ontológico del ersatzismo, la modalidad necesita mundos posibles.<sup>37</sup>

Ahora bien, no toda conexión modal fundamentará hechos necesarios (como diferentes de, digamos, meras posibilidades). Consideremos este ejemplo. Si *x* es humano, entonces presumiblemente es necesario que *x* sea un mamífero. Así que debe haber, de acuerdo a la teoría de Jacobs, una conexión modal R entre *ser humano* y *ser mamífero*, tal que cualquier cosa teniendo la primera propiedad tendrá, *en virtud de que* tal propiedad está R-conectada con *ser un mamífero*, esta segunda propiedad con necesidad. Pero entonces se asume que R debe ser un “hacedor de necesidad”—y no un “hacedor de contingencia”.

¿Por qué es R un hacedor de necesidad, y no un hacedor de contingencia? Los hacedores de *necesidad* claramente difieren sustancialmente de los de *contingencia*; parece que debería haber razón para esto.<sup>38</sup> Esto es lo que predice PRIMERA: una postura como la de Jacobs involucrará un criterio modal primitivo para seleccionar aquellas entidades E (que aquí son las conexiones modales entre propiedades) que fundamentan a las posibilidades de entre aquellas que no lo hacen—aquí, las conexiones no modales entre propiedades, o las conexiones modales que fundamentan imposibilidades. Ahora bien, SEGUNDA predice que la teoría de potencialidades tendrá que explicar por qué R es una conexión modal de un tipo determinado, es decir, por qué es que R es seleccionada por C como un hacedor de posibilidades (o de contingencias o necesidades o imposibilidades, como sea en cada caso). Y TERCERA dice que la teoría de potencialidades no puede afirmar que la selección hecha por C es correcta en virtud de las conexiones modales—porque las conexiones que fundan a los hechos modales tienen que haber sido seleccionadas por C. Así, el teórico de potencialidades no puede afirmar que la corrección de la selección de C está fundada en una co-

\* “Taken together, all such true claims represent the property structure of the actual world, a web of inter-related properties as it were. The properties-based theory of modality, and all of its specific versions, hangs all of modality on this structure of properties. Some proposition or truth-bearer, T, is possibly true just in case there is some actual property P (or some collection of properties) connected by R to some other property Q (or collection of properties), such that Q is a truthmaker for T, or there is a chain of such connections.”

37 Agradezco a Lourdes Valdivia y a Erick Llamas haberme hecho notar la necesidad de aclarar este punto.

38 Y claramente no puede afirmarse que la diferencia se da en virtud de la obvia diferencia entre necesidad y contingencia—pues se suponía que la teoría iba a *explicarnos* tal diferencia.

nexión modal, so pena de circularidad. Así que (como predice CUARTO) el teórico de potencialidades tiene que aceptar la primitividad de  $C$ ; pero esto resulta en aceptar modalidad absolutamente fundamental, en tensión con la naturaleza parcialmente fundada que la teoría de poderes le atribuye. Este argumento muestra que la teoría de potencialidades: o es presa de circularidad, o deja algunos hechos sin explicar—hechos que claramente necesitan ser explicados—o se reduce al absolutismo.

Aquí, como pasaba con el ersatzismo, no me estoy quejando porque la teoría de potencialidades sea no-reductiva, sino porque no sea *absolutamente no reductiva*—pues toma a la modalidad como *parcialmente dependiente*. Tanto el ersatzista que le da un uso ontológico a sus mundos como el teórico de las potencialidades dicen que la modalidad surge una vez que se fijan ciertos hechos: la existencia de mundos posibles, o la existencia de conexiones entre propiedades. Peor estos hechos resultan tener un importante aspecto modal por naturaleza. Entonces tenemos derecho a preguntar cómo es que *esos* hechos se fijaron en primer lugar, si no por un hecho modal distinto (y más fundamental).

Creo que las consideraciones argumentadas para las dos familias de teorías nos dan razones adecuadas para aceptar a las tesis de PRIMERA a CUARTA. Hay que notar que el argumento dado *no* es inductivo: *no* estoy argumentando mediante una generalización inductiva, desde la teoría de potencialidades y la ersatzista, que aceptemos PRIMERA a CUARTA. Más bien, creo que estas dos teorías son *ejemplos* de cómo, *por su misma estructura como teoría no-no*, una teoría se verá sujeta a los problemas puestos por PRIMERA a CUARTA.

Ahora, bien, aplicando las tesis QUINTA a SEXTA, vemos que el teórico de potencialidades que quiera rechazar la carga metafísica de su teoría nos debe una garantía de que la estructura de las conexiones modales se “engarza”, por así decir, de manera fiel a la estructura de la modalidad.<sup>39</sup>

Si la generalización corre como lo pretendo, los problemas puestos para las teorías no-no reducen el espacio dialéctico a ya sea el absolutismo o el reductivismo (asumiendo, claro, que queremos seguir siendo realistas acerca de la modalidad). Para achicar todavía más tal espacio, ahora voy a considerar una teoría reductiva—el esencialismo de Fine—y argumentaré que no logra la reducción. Ya para el siguiente capítulo estaré proponiendo una teoría representacional que, pretendo, capture mejor a la estructura de la modalidad—una vez que hemos aceptado su fundamentalidad.

## 24. una teoría reductiva: el esencialismo de Fine

De acuerdo con Fine (1994), lo que es esencial a  $x$  no es reducible a lo que le es necesario *de re* a  $x$ . Sus contraejemplos son bien conocidos, así que los recordaré sólo de manera rápida (*cf. ibid.*, pp. 4-6).

(SÓCRATES – UNITARIO DE SÓCRATES) Es necesario, de acuerdo a las posturas estándar en la teoría modal de conjuntos, que Sócrates pertenezca a su conjunto unitario si él existe; pues, necesariamente, el unitario existe si Sócrates existe y, necesariamente, Sócrates pertenece a su unita-

<sup>39</sup> Al menos toda característica y distinción físico-modal, dado que algunos teóricos de las potencias sólo buscan explicar la modalidad física.

rio si ambos Sócrates y su unitario existen. Se sigue (de acuerdo a la concepción modal de la esencia, cf. §1.4, arriba) que Sócrates pertenece *esencialmente* a su unitario.

(SÓCRATES – EIFFEL) No es esencial que el ejemplo apele a una entidad abstracta. Consideremos dos objetos cuyas naturalezas están desconectadas, digamos Sócrates y la Torre Eiffel. Entonces es necesario que Sócrates y la Torre sean distintos. Pero no es esencial a Sócrates que él sea distinto de la Torre; pues no hay nada en su naturaleza que lo conecte en alguna manera especial con ella.

(VERDADES NECESARIAS Y ESENCIALES, I) Consideremos cualquier verdad necesaria; podría ser una verdad matemática particular, por ejemplo, o la conjunción de todas las verdades necesarias. Entonces es necesariamente el caso que esta verdad debería darse si Sócrates existe. Pero no es parte de la esencia de Sócrates que haya infinitos números primos o que el mundo abstracto de los números sea como es.

(VERDADES NECESARIAS Y ESENCIALES, II) De acuerdo a la concepción modal de la esencia, entre las verdades necesarias están las oraciones de esencias. Pues una oración de esencia es una oración de necesidad y entonces será, como cualquier oración de necesidad, necesariamente verdadera si es verdadera del todo. Se sigue que será parte de la esencia de cualquier objeto que todo otro objeto tenga las propiedades esenciales que él tiene: “¡Oh, feliz metafísico! Pues al descubrir la naturaleza de una cosa, con ello descubre la naturaleza de todas las cosas”.\*

Con estos contraejemplos, Fine refuta a la concepción estándar de la esencialidad, pues una de las condiciones que tal concepción afirma como suficientes para la esencialidad (que la posesión de la propiedad en cuestión sea necesaria), falla.<sup>40</sup>

Aquí no voy a cuestionar los contraejemplos de Fine, pues mi intuiciones acerca de los casos que presenta son como las suyas.<sup>41</sup>

Lo que quiero cuestionar es su afirmación de que la modalidad es reducible a la esencialidad. Habiendo aislado una noción de reducción en §2.1, en lo que sigue presentaré la teoría de Fine (§2.4.1) y argumentaré que no logra reducir a la modalidad (§2.4.2). De hecho, también argumentaré que es muy poco probable que alguna teoría esencialista lo logre (§2.4.3).

Estaré tomando en cuenta casi solamente a la teoría esencialista de Fine, junto a su lógica y teoría de modelos para esta. Lo hago para tener un contrincante definido, dado que la postura esencialista de Fine es la más desarrollada hasta ahora; pero eso no significa que mis argumentos estén dirigidos a Fine mismo—quien después expresó dudas (independientes de las que presentaré) acerca de la supuesta reducción (Fine 2002).<sup>42</sup> Más bien, mis argumentos se dirigen a otros filósofos que favorecen a las esencias sobre la necesidad en el orden de la fundamentalidad (véa-

---

\* “O happy metaphysician! For in discovering the nature of one thing, he thereby discovers the nature of all things.”

40 Por supuesto, he sido demasiado rápido en la aceptación de los contraejemplos de Fine. Hay varias sutilezas que podrían considerarse, pero es probable que las teorías resultantes o compartirían el espíritu de Fine o, si no, no es claro que tengan el mismo poder explicativo que la de él. Por cuestiones de espacio, no puedo examinar esas sutilezas aquí.

41 Hay filósofos que rechazan el rechazo de Fine: Gorman 2005, Correia 2007, and Zalta 2006 (los últimos dos de hecho afirman que la esencia es reducible a la necesidad; ver también Fine 2007b). Sigo sin convencerme de esto, pero aquí no entraré en el debate.

42 Ver también Hale 1996 y Whittle 2010 para otras dudas acerca del proyecto esencialista reductivo. (Whittle argumenta contra el proyecto de reducir a la necesidad contra el “análisis metafísico”, que básicamente se reduce a un esencialismo.) Véase Shalkowski 1997 para discusión de la objeción de Hale.



se abajo para referencias).<sup>43</sup> Finalmente, como notaré donde sea apropiado, mucho de lo que diré se basa en constreñimientos que de manera natural se imponen sobre cualquier teoría reductiva de la modalidad. El esencialismo no es especial en este aspecto.

#### 24.1 presentando la propuesta reductivista

Fine argumenta que la noción de esencia debería tomarse como primitiva, pero que puede ser iluminada por la noción de *definición real*: la definición real de  $x$  especifica lo que es  $x$ , cuál es la *naturaleza* de  $x$ , cuál es la *esencia* de  $x$ .<sup>44</sup> La esencia de  $x$  es especificada objetivamente por todas aquellas proposiciones que son *verdaderas solamente en virtud de la naturaleza* (o *identidad*) de  $x$ .

Ahora bien, Fine sugiere así un proyecto reductivo (1994: 9):<sup>45</sup>

Lejos de ver a la esencia como un caso especial de la necesidad metafísica, deberíamos ver a la necesidad metafísica como un caso especial de la esencia. Pues cada clase de objetos [...] dará lugar a su propio dominio de verdades necesarias, las verdades que fluyen de la naturaleza de los objetos en cuestión. Las verdades metafísicamente necesarias pueden entonces identificarse con las proposiciones que son verdaderas en virtud de todos los objetos en absoluto.\*

En otras palabras: sea  $Def(x)$  el conjunto de proposiciones verdaderas solamente en virtud de la identidad de  $x$  (en la notación de Fine,  $Def(x) = \{p : \Box_x(p)\}$ ). Ahora, si hay proposiciones que sean verdaderas en virtud de más de un objeto pero no en virtud de cualquiera de ellos solamente—como, quizá, *que Marte y Antares son esencialmente distintos*, cuya verdad depende tanto de la naturaleza de Marte como de Antares, pero no en la naturaleza de Marte ni en la de Antares por separado—las acomodamos en  $Def^+(x, y, \dots)$ , que incluye las definiciones de  $x, y, \dots$  además de aquellas proposiciones que sean verdaderas en virtud de  $x, y, \dots$  tomados *colectivamente*.<sup>46</sup> Así, el esencialista reductivista puede ser tomado como afirmando que la clase de *todas* las proposiciones metafísicamente necesarias es  $\cup_{\mathbf{x} \in O} Def^+(\mathbf{x})$ , donde  $O$  es el dominio que contiene a todas las clases  $\mathbf{x}$  de entidades (concediendo la asunción no-trivial de que haya tal dominio).

43 Se suele afirmar que tal favorecimiento llega hasta Aristóteles y la tradición forjada por su metafísica. Esta afirmación acerca de orígenes históricos es plausible, pero aquí no tendré nada que decir acerca de su verdad o importancia—o acerca de Aristóteles, para el caso. (Aunque, por supuesto, si esta postura sobre la doctrina aristotélica es verdadera y mis argumentos son correctos, Aristóteles es refutado en este punto.)

44 Uno podría dudar de si la noción de definición objetual hace sentido en absoluto. O, menos fuertemente, uno podría dudar, como lo hace Robertson (2007: §2), si la noción hace sentido cuando se aplica a cualquier objeto en absoluto. Para lo que valga, yo comparto las dudas: no me es claro si objetos concretos e *históricos* como Sócrates *tienen* una definición no trivial—sobre todo debido a los problemas ya remarcados en la literatura sobre el esencialismo del origen (Mackie 2006, Robertson 2000; ver también Almog 2010). Por mor de la discusión, de cualquier manera, dejaré de lado estas cuestiones.

45 Véase también 2012 y Koslicki 2012 para desarrollos de este proyecto reductivo.

\* “[... F]ar from viewing essence as a special case of metaphysical necessity, we should view metaphysical necessity as a special case of essence. For each class of objects [...] will give rise to its own domain of necessary truths, the truths which flow from the nature of the objects in question. The metaphysically necessary truths can then be identified with the propositions which are true in virtue of the nature of all objects whatever.”

46 Con lo que  $Def^+(x, y, \dots) = Def(x) \cup Def(y) \cup \dots \cup \{p : \Box_{x,y}(p)\} \cup \{p : \Box_{x,y,\dots}(p)\} \cup \dots$  (Véase See Correia 2006 para un refinamiento de la postura de Fine que toma estas complicaciones en cuenta de manera más natural.)

Veamos cómo esto encaja en TEORÍA REDUCTIVA, cuando sustituimos la variable  $R$  por *esencias*: los hechos sobre lo que es esencial, o lo que “fluye” de la esencia de los objetos, son hechos acerca de esencias, así que se satisface R2. Y la afirmación de que estos son necesarios puede tomarse como la afirmación de que las esencias de los objetos son las que sirven como hacedores de verdad para las proposiciones sobre tales objetos; por lo que esto va dirigido hacia R3. ¿Qué tal R1? Los hechos esencialistas ¿necesitan alguna ideología modal para caracterizarse? Pospondré esta pregunta para la siguiente sección—donde defenderé que debe responderse afirmativamente.

Recordemos que la esencia de  $x$  es especificada por todas aquellas proposiciones que son *verdaderas solamente en virtud de la naturaleza de  $x$* . Así que si tenemos un entendimiento suficientemente claro de lo que es que una proposición sea verdadera *en virtud de la identidad de un objeto*, hemos allanado el camino hacia una teoría reductiva y esencialista de la modalidad. ¿Qué significa este “en virtud de”? Aquí, es natural voltear hacia la teoría de modelos para la lógica esencialista. Esto *no es* asumir que uno podría, o debería, “leer” un *análisis* de la noción en la semántica—la asunción, más bien, es que la semántica debería ayudarnos a constreñir la interpretación conceptual que se pretende para el marco esencialista. Tal interpretación debería simplemente *ajustarse* a la estructura formal de la teoría de modelos.<sup>47</sup>

En su artículo de 2000, Fine define la verdad en un mundo de una proposición con la forma  $\Box_F(A)$ , que intuitivamente significa “ $A$  es verdad en virtud de la naturaleza de algunos (o todos) los objetos  $F$ ” como (p. 548):

(VERDAD ESENCIAL)  $w \models \Box_F(A)$  sii (a)  $[A]^M \subseteq c(F_w)$  y (b)  $v \models A$  siempre que  $I_v \supseteq F_w$ .

Desempaquemos esta definición.

Los modelos para la lógica de la esencia  $E5$  son cúadruplas  $M = \langle W, I, \succcurlyeq, \varphi \rangle$ , donde  $W$  es un conjunto no vacío (intuitivamente: el conjunto de mundos), e  $I$  es una función mapeando cada  $w \in W$  a un conjunto  $I_w$ , el conjunto de individuos en  $w$  (*no* necesariamente el conjunto de individuos existiendo en  $w$ :  $I_w$  podría contener no sólo aquellos objetos que existen en  $w$ , sino también aquellos que son meramente “posibles desde el punto de vista de  $w$ ”, véase Fine 2000: 543). Ahora bien,  $\succcurlyeq$  es la relación de dependencia del modelo: una relación transitiva y reflexiva sobre  $\cup_{w \in W} I_w$  (el conjunto de individuos en cualquier mundo); además, cada mundo está cerrado con respecto a  $\succcurlyeq$  (es decir, si  $a$  es un individuo en  $w$  y  $a \succcurlyeq b$ , entonces  $b$  también es un individuo en  $w$ ). Finalmente,  $\varphi$  es la función de interpretación, llevando cada constante a un individuo en algún mundo, cada predicado rígido (intuitivamente, predicados que se definen en términos de objetos particulares, con la forma:  $x$  es ya sea  $a$  o  $b$  o...) a su extensión, y cada par de un predicado (no rígido) de  $n$  lugares y un mundo  $w$  a un conjunto de  $n$ -tuplas tomado de  $I_w$ .

Ahora bien,  $[A]^M$  es el *contenido objetual de  $A$  de acuerdo a  $M$* : donde cada  $a_i$  es una constante ocurriendo en  $A$  y cada  $P_j$  un predicado rígido ocurriendo en  $A$ ,  $[A]^M = \{\varphi(a_1), \dots, \varphi(a_m)\} \cup \varphi(P_1) \cup \dots \cup \varphi(P_n)$ . Con lo que el contenido objetual de  $A$  (con respecto al modelo relevante) es el conjunto de todas las cosas que son ya sea referidas por una constante en  $A$ , o que no caen bajo la extensión de un predicado rígido en  $A$ . Finalmente,  $c(F_w)$  es la clausura de la extensión de  $F$  en  $w$  con respecto a la relación de dependencia:  $c(F_w) = \{b: a \succcurlyeq b \text{ para algún } a \in F_w\}$ , es decir,  $c(F_w)$  es el conjunto de todas las cosas que dependen en alguna cosa que es  $F$  en  $w$ .

Digamos (estipulativamente) que  $x$  es una “base” de  $y$  cuando  $y$  depende de  $x$  (*i.e.*, cuando  $y \succcurlyeq$

47 Agradezco a Kit Fine el haberme hecho clara la necesidad de esta aclaración.

$x$ ). Así, en resumen, VERDAD ESENCIAL se reduce a:  $\Box_F(A)$  es verdad en un  $w$  del modelo  $M$  sii los objetos mencionados en  $A$  (de acuerdo a  $M$ ) son todos una “base” de algún  $F$  en  $w$ , y cualquier mundo  $v$  que contenga al menos ésas entidades  $F$  de  $w$  también hará verdad a  $A$ . En otras palabras,  $A$  es verdad en virtud de los  $F$ s si es verdad en todo mundo conteniendo esos  $F$ s que dependen los objetos acerca de los que es  $A$ —ellos “fuerzan” a  $A$  a ser verdad.

¿Cómo es que esto valida la asimetría entre Sócrates y su unitario,  $\{S\}$ ? Arriba vimos que, intuitivamente, le es esencial a  $\{S\}$  el que Sócrates  $\in \{S\}$ , pero eso no le es esencial a Sócrates. Es decir,  $\Box_{\{S\}}(\text{Sócrates} \in \{S\})$  y  $\neg\Box_{\text{Sócrates}}(\text{Sócrates} \in \{S\})$ . Ahora bien, por VERDAD ESENCIAL:

(A)  $\Box_{\text{Sócrates}}(\text{Sócrates} \in \{S\})$  sii (i)  $\{\text{Sócrates}, \{S\}\} \subseteq c(s) = \{S\}$ ; y (ii) cualquier  $v$  tal que Sócrates  $\in I_v$   $v \models \text{Sócrates} \in \{S\}$ .

(B)  $\Box_{\{S\}}(\text{Sócrates} \in \{S\})$  sii (i)  $\{\text{Sócrates}, \{S\}\} \subseteq c(\{S\}) = \{\text{Sócrates}, \{S\}\}$ ; y (ii) cualquier  $v$  tal que  $\{S\} \in I_v$   $v \models \text{Sócrates} \in \{S\}$ .

Ahora bien, tanto (A)(ii) como (B)(ii) son verdaderas (pues Sócrates  $\in I_v$  sii  $\{S\} \in I_v$ , y  $v \models \text{Sócrates} \in \{S\}$  para todo mundo  $v$  en que Sócrates y  $\{S\}$  existen). Sin embargo, (A)(i) es falsa, y por ello  $\neg\Box_{\text{Sócrates}}(\text{Sócrates} \in \{S\})$ : con lo que no le es esencial a Sócrates el pertenecer a su unitario. (Aunque (B)(i) es verdadera, con lo que le es esencial al unitario de Sócrates el tener a Sócrates como miembro).

Ahora parece que el único concepto involucrado en VERDAD ESENCIAL que permanece sin explicar es el de *dependencia*. Fine define a la dependencia en términos de la esencialidad (1995: 275): *ser verdad en virtud de la identidad de* se define en términos de la esencia, y por ello la dependencia ontológica se define en términos de la esencialidad; pero, como vimos arriba, la esencialidad se define en términos de la dependencia ontológica. Sería injusto presionar una acusación de circularidad aquí: nunca se nos prometió una definición reductiva de esencia—la esencia es el punto de inicio, el punto primitivo.

## 24.2 evaluando la propuesta reductivista

Ya podemos evaluar la postura reductiva de nuestro esencialista. Primero que nada, preguntémonos: ¿que significa decir que  $x \in I_w$ ? Fine estipuló que

La presencia de un objeto en un mundo no se toma como garantizando su existencia, sino solamente su posibilidad. Así que cada mundo se tomará como involucrando su propia ‘visión’ acerca de qué objetos son posibles y cuáles no. (2000: 543, mi énfasis)\*

Tomado al pie de la letra (o, como dicen en inglés, “*at face value*”), esto implica que la condición es  $x \in I_w$  *per se* modal: es satisfecha por un objeto sólo si es *posible* desde “el punto de vista” de  $w$ . Así, si queremos satisfacer el requisito R1 de TEORÍA REDUCTIVA, ojalá que encontremos una manera de frasear tal condición sin terminología modal sin reducir.

Una manera de proceder sin ninguna ideología modal es, por supuesto, identificar a  $I_w$  (los individuos en  $w$ ) no con el conjunto de objetos *posibles* en  $w$ , sino con el conjunto de objetos *existiendo* en  $w$ . Y de hecho Fine toma esa opción (2000: 550, mi énfasis):

\* “the presence of an object in a world is not taken to guarantee its existence *but merely its possibility*. Thus each world will be taken to embody its own ‘view’ of which objects are possible and which are not.”

Hay una interpretación de la semántica formal en la que cada dominio  $I_w$  se toma como consistiendo en los objetos que existen en  $w$ . Un enunciado  $\Box_F A$ , entonces, es verdadero en un mundo  $w$  sii cada objeto mencionado en  $A$  depende para su existencia en algo que es  $F$  y  $A$  es verdad en cada mundo que contiene a los objetos que existen en  $F$ . La prueba [presentada posteriormente en el artículo] establecerá, entonces, que el sistema es completo bajo esta interpretación *modal-existencial* de  $\Box_F$ .\*\*

Aquí hay que notar *dos cosas*. Primero, que fue *contra* la teoría modal-existencial de los enunciados esencialistas que Fine tan efectivamente argumentó en su artículo de 1994. Tal postura es asumida en su artículo de 2000 por razones puramente técnicas—para simplificar las pruebas. Pero sería un error leer una teoría metafísica de la esencia en tal simplificación (como él mismo remarcó: 2000: 550).

Segundo, parece que *Fine mismo* malentendió su sistema, al decir que “ $\Box_F A$  ... es verdadero en un mundo  $w$  sii cada objeto mencionado en  $A$  depende para su existencia en algo que es  $F$  y  $A$  es verdad en cada mundo que contiene a los objetos que existen en  $F$ .” En realidad, como vimos arriba, VERDAD ESENCIAL NOS dice otra cosa:  $\Box_F A$  es verdadera en un mundo  $w$  sii todo objeto mencionado en  $A$  es una “base” para algún  $F$  de  $w$ , además de que  $A$  sea verdadera en todo mundo  $v$  cuyo dominio  $I_v$  contenga a los  $F$ s de  $w$ . En breve, en la cita Fine nos dice que los objetos mencionados en  $A$  dependen en los  $F$ s, mientras que VERDAD ESENCIAL NOS dice que los  $F$ s dependen de los objetos mencionados en  $A$ .<sup>48</sup>

Esto es relevante para los problemas filosóficos implicados por la interpretación modal-existencial del formalismo. Pues inmediatamente después de la cita de arriba, Fine dice:

Sin embargo, sería un serio error el suponer que la noción de esencia puede, por ello, ser analizada en términos de la modalidad y la existencia. Dos nociones pueden tener la misma lógica, incluso la misma semántica formal, y aún así ser intuitivamente muy distintas. De hecho, bajo la interpretación propuesta, cada existente existirá esencialmente, pero aún así este no es un resultado que, intuitivamente, queremos. O, de nuevo, es difícil ver bajo qué sentido de necesidad la existencia de  $\{2\}$  implicaría que  $2 \in \{2\}$  y aún así la existencia de 2 no implique que  $2 \in \{2\}$ .\*\*\*\*

Como vimos al final de la sección anterior, el esencialismo basado en VERDAD ESENCIAL bastaba para recuperar la asimetría intuitiva entre Sócrates y su unitario respecto a la oración “Sócrates  $\in \{S\}$ ”.<sup>49</sup> También basta para recuperar la asimetría respecto a la oración “ $2 \in \{2\}$ ” entre 2 y  $\{2\}$ ,

\*\* “[...T]here is an interpretation of the formal semantics in which each domain  $I_w$  is taken to consist of the objects that exist at  $w$ . A modal statement  $\Box_F A$  is then true at a world  $w$  iff each object mentioned in  $A$  depends for its existence on something that  $F$ 's and  $A$  is true in each world that contains the objects that exist in  $F$ . The proof to follow will then establish that the system is complete under this modal-existential interpretation of  $\Box_F$ .”

48 En una versión anterior de este trabajo, yo mismo fui presa del malentendido que Fine comete aquí. Debo a Alessandro Torza el haberme hecho notar esto.

\*\*\*\* However, it would be a serious mistake to suppose that the notion of essence can therefore be analyzed in terms of modality and existence. Two notions can have the same logic, even the same formal semantics, and yet be intuitively quite different. Indeed, under the proposed interpretation, each existent will essentially exist and yet this is not a result which, intuitively, we want. Or again, it is hard to see under what sense of necessity the existence of  $\{2\}$  would entail that  $2 \in \{2\}$  and yet the existence of 2 not entail that  $2 \in \{2\}$ .

49 Debo este punto también a Alessandro Torza.

debido a que VERDAD ESENCIAL incluye, crucialmente, la noción de dependencia ontológica,  $\succcurlyeq$  (al incluir los conjuntos clausura, que se definen en términos de ella). Y la dependencia ontológica es una noción hiper-intensional, en el sentido en que puede ser verdad que  $a \succcurlyeq b$  y que  $\neg(b \succcurlyeq a)$ , aunque  $a$  y  $b$  existan en exactamente los mismos mundos (como sucede con 2 y {2}).

Sin embargo, todavía no estamos libres de problemas. Pues VERDAD ESENCIAL y la interpretación modal-existencial del formalismo (*i.e.*, aquella donde cada dominio  $I_w$  es el conjunto de los objetos que *existen* en  $w$ ) implican, como bien nota Fine, que *Sócrates es esencialmente existente*: para cualquier  $w$ ,  $w \models \Box_{\text{Sócrates}}(\text{Existe}(\text{Sócrates}))$  sii (i) Sócrates depende de Sócrates, y (b) Existe(Sócrates) es verdad en todo mundo  $v$  donde Sócrates  $\in I_v$ . Debido a la interpretación de los dominios  $I_v$  como dominios de existencia, Sócrates es un existente esencial. Pero esto fue rechazado por el mismo Fine en su artículo de 1994 (p. 6).

Así que la interpretación modal-existencial del formalismo debería tomarse como un artefacto del modelo formal, no como una característica de la realidad metafísica subyacente que es modelada por la lógica esencialista. Pero entonces estamos de vuelta a donde empezamos: ¿cómo deberíamos entender la condición de que  $x \in I_w$ ?

Al abrir su artículo de 2000, Fine escribe (550, mi énfasis):

La idea básica detrás de la semántica es que un enunciado debería tomarse como verdadero en virtud de la naturaleza de ciertos objetos si es verdadero en todo mundo *compatible* con la naturaleza de esos objetos. Haremos la asunción simplificadora de que cada mundo es compatible con la naturaleza de todos y sólo aquellos objetos que contiene.<sup>\*\*\*, 50</sup>

Pero ¿qué significa que un mundo  $w$  sea *compatible* con la naturaleza de los objetos en  $x \in I_w$ ? El significado de “compatibilidad” que inmediatamente viene a la mente es uno *modal*: dos proposiciones son compatibles si *pueden* ser verdaderas conjuntamente; dos objetos son compatibles si *pueden* coexistir. Si hay alguna otra noción no-estipulativa, *no-modal* de compatibilidad, todavía no sé de ella. Además, la compatibilidad en cuestión es plausiblemente metafísica—una compatibilidad epistémica, o lingüística, infectaría de dependencia de la mente todo el aparato, llegando hasta la noción de esencia. Así que el esencialista debe encontrar una manera de reducir a la compatibilidad a la esencialidad, o de entenderla de una manera no modal.<sup>51</sup> Tal investiga-

---

\*\*\* “The basic idea behind the semantics is that a statement should be taken to be true in virtue of the nature of certain objects if it is true in any world *compatible* with the nature of those objects. We shall make the simplifying assumption that each world is compatible with the nature of all and only those objects that it contains.”

50 Claramente, la “asunción simplificadora” podría desecharse al introducir una relación de accesibilidad  $R$ —de tal manera que  $Rww$  sii  $w$  es compatible con la naturaleza de los objetos en  $I_v$ —con ello permitiéndonos trabajar con lógicas esencialistas análogos a lógicas modales diferentes de  $S5$ . Los puntos presionados abajo todavía podrían presionarse, obviamente.

51 Fine (en comunicación personal) consideró la idea de que la relación de compatibilidad entre un mundo (o su darse) y la naturaleza de los individuos que contiene podría ser una noción *puramente esencialista*. Pero no logro ver la motivación para esta idea—la noción de una *naturaleza* es plausiblemente esencialista; pero la noción de *compatibilidad* me parece ser claramente modal. Así que si estoy en lo correcto, la noción de *compatibilidad de naturalezas* combina nociones *tanto* esencialistas *como* modales.

Fuera de eso, aceptar la compatibilidad de naturalezas como una noción puramente esencialista haría al reductivismo esencialista demasiado “barato”. Pues por la misma razón, podríamos aceptar la compatibilidad de proposiciones como una noción puramente “mundista”, y por ello el argumento ampliamente aceptado de que el ersatzismo lingüístico no logra reducir a la modalidad porque asume que sus mundos son hechos de proposiciones compatibles (*cf.* §2.2, arriba), se vendría abajo. En general, no quedaría ninguna motivación para el requisito

ción no ha sido llevada a cabo por los esencialistas, así que aquí intentaré desarrollarla de la mejor manera en que pueda.

Queremos entender qué es que algo sea hecho verdad por las esencias. Nos es dicho que pensemos a las esencias como la definición real de un objeto. Esa definición real es dada por las proposiciones verdaderas en virtud de la naturaleza de un objeto. La estructura de la relación *ser verdad en virtud de la naturaleza de* es dada por VERDAD ESENCIAL, que se traduce en la idea de que una proposición de la forma  $\Box_F(A)$  es verdadera sii, dado que  $A$  es acerca de algunos objetos que dependen en algunas cosas,  $A$  es verdad en todo mundo  $w$  conteniendo estos en su dominio. El problema surgió porque tales dominios eran entendidos de la noción modal de *compatibilidad*, y la alternativa más obvia—entenderlos en términos de existencia—trajo de vuelta precisamente esos problemas de la postura tradicional contra la que Fine tan fuertemente arguyó.

Entonces tenemos el siguiente constreñimiento sobre una interpretación conceptual aceptable de la maquinaria lógica: no debemos entender a los dominios de los mundos como consistiendo en entidades *existiendo* en esos mundos, ni de entidades *metafísicamente posibles* relativo a esos mundos (o *metafísicamente compatibles* con la naturaleza de los objetos en esos mundos, o con la naturaleza de los mundos mismos). Tenemos que—¡de alguna manera!—encontrar un camino que no sea ni la llana existencia ni la posibilidad—del tipo metafísico, claro está. Si encontramos una manera plausible de hacerlo, podemos simplemente “ensartar” esa noción en la definición de los dominios de mundos, y así habremos logrado llevarnos a una interpretación conceptual que encaje en la estructura de la teoría de modelos, con ello rescatando al esencialismo reductivista.

Un camino posible es volverse anti-realista, y afirmar que la *concebibilidad*, la *imaginabilidad*, o algún otro tipo de característica conferida por la mente es lo que hace que un objeto pertenezca a un dominio de un mundo (Sidelle 1989; Sveinsdóttir 2008). Por ejemplo, uno podría afirmar que la condición  $x \in I_w$  es satisfecha por un objeto sólo si la existencia de ese objeto en  $w$  es *concebible*. Como dije arriba, esto infectaría de dependencia de la mente todo el aparato esencialista, llegando hasta la noción de esencia—y, para los esencialistas reductivistas, hasta la necesidad metafísica. Pero seguramente uno preferiría ser llevado hacia el anti-realismo de la necesidad *mediante argumentos*, no ser arrojado a este por problemas inherentes al esencialismo. Sea como sea esto, también está la preocupación de que ojalá que resulte que ser concebible *no necesite* interpretarse como siendo *metafísico-posiblemente concebido*, pues entonces hemos metido una noción metafísico-modal de contrabando (Bealer 2002: §1).

Supongamos que queremos resistirnos al anti-realismo. Entonces la manera más obvia de proceder es apelar a la modalidad *lógica*, pues en un sentido parece más *débil* que la existencia y la posibilidad metafísica: lo que es lógicamente posible, es muy común pensar, no necesita ser metafísicamente posible, ya no digamos existente. Además, para la mayoría de posturas, la modalidad lógica es *independiente* de la mente, y por ello no presenta riesgo alguno de esparcir el anti-realismo.<sup>52</sup>

---

plausible de que una teoría que busque reducir la modalidad no debería apelar a relaciones primitivas de compatibilidad.

52 Uno podría presionar aquí, pero yo no lo haré, la pregunta de cómo vamos a entender a la modalidad lógica. Fine mismo (1994) sugiere que la modalidad lógica es verdad en virtud de los “conceptos” lógicos; con ello, efectivamente, sugiriendo que la modalidad lógica también se reduce a la esencialidad. Pero estamos tratando de entender a la esencialidad, y apelar a la modalidad lógica en ese caso no serviría de nada.

La idea entonces sería que la condición  $x \in I_w$  significa que  $x$  es lógicamente compatible con cualquier  $y$  tal que  $y \in I_w$ . ¿Qué es esta noción de compatibilidad lógica objetual? Tomemos las definiciones reales de  $x$  y de  $y$ . Si no implican una contradicción lógica, entonces son compatibles lógico-objetualmente. Uno podría entonces pasar a asignar un objeto arbitrario a un mundo arbitrario  $w$ , y entonces “rellenar” a  $I_w$  con cualesquiera otros objetos que satisfagan el requisito de compatibilidad (en una construcción análoga a una construcción estilo Lindenbaum de conjuntos máximos y consistentes de fórmulas).

Una preocupación inicial acerca de la idea es esta. Intentamos entender a la compatibilidad como compatibilidad lógica, y para ello entendimos a la compatibilidad lógica de dos objetos como la compatibilidad lógica de sus definiciones reales. Sabemos que estas se toman como primitivas por los esencialistas. Pero aún así su estructura se acepta como necesitando algo de aclaración; podemos hacer un punto sobre esta base.

Tenemos dos acercamientos complementarios para iluminar la estructura de la verdad esencial. Primero, son las proposiciones que codifican la naturaleza de los objetos. Segundo, son las proposiciones que dan la interpretación intencional de la lógica de la esencia de Fine. Los acercamientos son complementarios porque se asume tanto que la interpretación intencional de la lógica de la esencia será dada en términos de las proposiciones que dan definiciones reales, como que las proposiciones que dan definiciones reales encajarán en la interpretación intencional de la lógica. (Compárese: una verdad necesaria es tanto una que es verdadera para toda manera en que el mundo pueda ser, como una que encaja en la interpretación intencional de la semántica kripkeana de la lógica modal.) Pero lo que nos trajo hasta aquí fue el problema de cómo entender la condición de que  $x \in I_w$ —que es un componente esencial de la estructura lógica de la interpretación intencional de la lógica, dada por VERDAD ESENCIAL. Así que nos hemos atrapado en un círculo muy apretado: intentamos entender a la teoría esencialista como la interpretación intencional de la lógica; para entender a la lógica intentamos entender la condición de que  $x \in I_w$ ; pero para hacer esto necesitamos entender a la teoría esencialista—en particular, a la noción de definición real.

Nótese que no estoy pidiendo que la esencialista abandone su primitivismo acerca de la esencia. Estaría completamente satisfecho con una explicación esencialista de la verdad esencial—una que nos dé la interpretación intencional de VERDAD ESENCIAL en términos puramente esencialistas. Pero he estado presionando el punto de que tal información no se ha dado. Por todo lo que se ha dicho, VERDAD ESENCIAL requiere ideología modal para entenderse.

Un segundo problema surge incluso condicional en la suposición de que tenemos un entendimiento suficiente, que no requiere ideología modal, de lo que hace que una verdad sea esencial. Es acerca de la base sobre la que la esencialista podría ofrecer su propuesta. ¿Por qué es que cualquier proposición metafísicamente posible ha de ser hecha verdadera por hechos acerca de esencias lógicamente compatibles? Necesitamos una garantía de que cualquier cosa que sea lógicamente compatible con verdades esenciales es, por ello, metafísicamente posible. (Compárese: necesitamos una garantía de que cualquier cosa que sea verdad en algún mundo posible es por ello metafísicamente posible. De aquí que los ersatzistas tengan que aceptar una noción *metafísica* de consistencia—a diferencia de una noción meramente lógica: sin ella, podrían haber proposiciones (o fórmulas u oraciones) verdaderas en algún mundo que aún así no fueran metafísicamente posibles (cf. §2.2, arriba)).

Ninguna garantía tal es dada automáticamente sólo por lo que ha sido ofrecido por las esencialistas. Nada hasta ahora previene que algunos objetos tengan esencias metafísicamente *imposibles*. El conjunto paradójico de Russell, o el Dios leibniziano que puede crear y destruir mundos posibles a voluntad, tienen definiciones reales claras. Bajo muchas posturas, aún así, ellos son metafísicamente posibles. Así, incluso si la definición real de algún  $x$  es lógicamente compatible con la verdad esencial de que *el Dios de Leibniz puede crear y destruir mundos posibles a voluntad*, eso no necesariamente implica que  $x$  es metafísico-posiblemente existente. Si no, entonces los hechos acerca de la esencia de  $x$  no son claramente apropiados para ser los hacedores de verdad de proposiciones metafísico-posiblemente verdaderas. El esencialista podría afirmar que su teoría nunca tomará un  $x$  así para dar cuenta de la verdad de cualquier tal proposición. Pero necesitamos una garantía de ello.

La solución es obvia, pero su costo es dejar al reductivismo de lado. La solución es que  $x$  es metafísicamente posible sii es objetual-lógicamente compatible con aquellos objetos que son, ellos mismos, metafísicamente posibles. Obtenemos, a lo más, una equivalencia interesante; pero nada cercano a una reducción.

Hay un tercer problema, relacionado con esto. Consideremos la definición real de Sócrates. Plausiblemente, incluye proposiciones como <Sócrates es humano> y <Sócrates es el hijo de Sofronisco y Fenarete>. Muy plausiblemente, como Fine notó, *no* incluye la proposición <Sócrates  $\in$  {Sócrates}>. Pero, por supuesto, la definición real de {Sócrates} *sí* incluye la proposición de que <Sócrates  $\in$  {Sócrates}>. Así que tenemos una asimetría aquí. La teoría esencialista que estamos considerando tiene los recursos para dar cuenta de ella: es porque es *lógicamente* consistente que un mundo incluya las proposiciones <Sócrates es humano> y <Sócrates es el hijo de Sofronisco y Fenarete>, pero no <Sócrates  $\in$  {Sócrates}>, que {Sócrates} no necesita existir en cada mundo donde Sócrates exista. Y es porque no es lógicamente consistente que un mundo incluya la proposición <Sócrates  $\in$  {Sócrates}> pero no a Sócrates, que Sócrates existirá en todo mundo donde {Sócrates} exista.

Pero las teorías matemáticas son metafísicamente necesarias. Así que la teoría de conjuntos con urelementos debe serlo también. Por ello, en todo mundo *metafísicamente* posible, si Sócrates existe, entonces {Sócrates} lo hace. Así que la asimetría a la que apuntamos en el párrafo anterior implica que hay hechos *esencialmente posibles* que son metafísicamente *imposibles*—como el que Sócrates exista pero {Sócrates} no.<sup>53</sup>

Así que lo que es posible “por todo lo que las esencias saben” (*i.e.*, los hechos que están determinados por la relación de compatibilidad lógica entre las esencias de los objetos) incluye tanto a lo que es metafísicamente posible como a lo que es metafísicamente *imposible*. Es decir, la esencialidad nos da un criterio *no suficiente* para la posibilidad metafísica (aunque parece más claro que *sí* da una condición necesaria): no todo lo que sea lógicamente compatible con las esencias es metafísicamente posible.

Pero entonces necesitamos un criterio para especificar qué cosas esencialmente posibles son

---

53 Con “hecho esencialmente posible” (cámbiese por *proposición* o por *verdad* o por una noción parecida, dependiendo de sus propios compromisos) simplemente me refiero a hechos que son lógicamente posibles relativo a las definiciones reales de los objetos— $h$  es esencialmente posible sii no se sigue una contradicción de la conjunción entre  $h$  y la conjunción de todas las definiciones reales. Debo el haber notado la necesidad de aclarar esto a Alessandro Torza.



*también* metafísicamente posibles: un criterio, *que no sea modal*, para discriminar a la posibilidad metafísica de la imposibilidad metafísica. No se ha ganado nada.

Recordemos que estábamos intentando entender cómo dar la interpretación intencional de VERDAD ESENCIAL. Para ello, intentamos encontrar un camino intermedio entre la existencia y la posibilidad metafísica—entre la Escila de la teoría modal-existencial y la Caribdis de una teoría no-reductiva de la necesidad. Hice mi mejor esfuerzo apelando a la modalidad lógica, pero encontré a tal idea insatisfactoria. Creo que ahora la carga de la prueba está del lado esencialista. Quizás la esencialista quiera mostrar que hay nociones puramente esencialistas que apuntan hacia tal camino intermedio. O quizá intentará de nuevo con la modalidad lógica. O quizá todavía otro camino pueda encontrarse.

### 24.3 un argumento general contra toda reducción esencialista

(o: no se puede sacar un conejo modal del sombrero esencialista)

¿Podríamos encontrar otra manera de reducir modalidad a esencia? Pues yo aconsejaría esperar sentados. Ahora digo por qué.

Si quieres reducir la necesidad a las esencias, ojalá que no *necesites* especificar tus esencias con ideología modal. Pero entonces seguramente necesitarás algo de ayuda extra para lograr adecuación material—ese balance perfecto entre dejar entrar a algunas imposibilidades y dejar salir a algunas posibilidades. Y muy poco probable que tal ayuda extra llegue alguna vez. Así que es muy poco probable que logres adecuación material sin ideología modal. Veamos.

No hay nada en el esencialismo que prohíba que los objetos imposibles tengan esencias (así, parte de la esencia del Dios de Leibniz es ser capaz de crear y aniquilar mundos posibles—pero tal entidad, bajo la mayoría de las posturas, es imposible).<sup>54</sup> Pero dejar fuera a las esencias de objetos imposibles es importante por las siguientes razones. El esencialista reductivista necesita encontrar algún tercer estatus metafísico que esté estrictamente entre la existencia y la posibilidad metafísica—encontrar un camino entre la teoría modal-existencial y el abandonar el reductivismo. Pero, si hay algo *útil* que todavía no sea la posibilidad metafísica y que no sea igual a la existencia, seguramente no es la imposibilidad metafísica, pues una teoría de la verdad esencial bajo la cual  $\langle p \rangle$  siendo verdad en virtud de los *Fs* permite que los *Fs* sean *impossibilia*, no logrará la adecuación material. Pues el reductivismo requiere que toda proposición modal verdadera sea hecha verdad por un hecho acerca de las esencias; pero seguramente nada metafísicamente imposible hace verdad a ninguna proposición acerca de lo que es posible. Y es por ello que el esencialista reductivista necesitará ayuda extra para especificar aquellas esencias tales que los hechos acerca de ellas serán los hacedores de verdad para las proposiciones modales.

Sólo el tiempo dirá si hay tal tipo de ideología, fuera de la esencialista, no definida por ninguna noción modal, y no idéntica al mero concepto de existencia. Pero yo no me quedaría a esperar—

<sup>54</sup> Por supuesto, lo que es imposible no podría existir. Pero (1) podría argumentarse que incluso las esencias de objetos no existentes, existen; y, más importantemente, (2) *qua* esencialista reductivista, tienes prohibido el simplemente apelar a la diferencia entre posibilidad e imposibilidad en tu marco teórico—pues tu postura busca *reducir* a la modalidad a ese marco, *no* juntarlo con el lenguaje de la metafísica modal.

si hay alguna, los metafísicos han estado ignorándola por mucho tiempo, y para su desgracia, pues ¡seguramente es tan fundamental y filosóficamente importante como los conceptos de *existencia*, de *esencia* y de *posibilidad*!

## 2.5. conclusión

Es hora de concluir este capítulo.

Hemos revisado una teoría reductiva—la esencialista—y una familia de teorías que, aunque se abstienen de reducir a la modalidad, la toman como parcialmente determinada por otra cosa o cosas—las teorías no-no. He dado argumentos contra ambas, y he lanzado la hipótesis de la fundamentalidad de la modalidad. En el siguiente capítulo, propondré un marco teórico desde el cual representar a la modalidad, de una manera que, me parece, da mejor cuenta de su fundamentalidad, y de las “piedras en el camino” para otras teorías que hemos hallado hasta ahora.

Una última nota sobre la cuestión de la esencialidad. Por todo lo que he argüido y todo lo que sé, la teoría modal-existencial de la esencialidad es incorrecta y deberíamos abandonarla: la teoría definicional de Fine está mucho mejor diseñada para llevar a cabo cualquier trabajo que queríamos que hiciera la teoría modal-existencial. Mi afirmación es, más bien, que la teoría definicional todavía involucra modalidad metafísica primitiva. *La esencia no ha de ser reducida a la modalidad, pero tampoco se ha de reducir a la modalidad a la esencia*. Por supuesto, una vez que hemos aceptado modalidad no-reducida en el intento esencialista reductivista, podemos volvernos o reductivistas *parciales* (enfrentando así los problemas a los que apunté en la sección pasada), o podemos volvernos no-reductivistas *absolutos*. Ambas opciones son consistentes con la aceptación del esencialismo.

Finalmente, hay que notar que el rechazo de la reducción de la modalidad a la esencialidad—y de la esencialidad a la modalidad—no implica que sean completamente independientes. Pues parece que lo que es esencial debería ser necesario—incluso si tal necesidad no es reducible a lo que es esencial. También parece completamente plausible que lo que es posible no debería conflictuar con lo que es esencial (asumiendo que los objetos *tienen* esencias), incluso si lo que es esencial no se reduce a dónde están los límites de la posibilidad. Hay una relación cercana y sistemática entre la modalidad y la esencialidad, incluso si tal relación no es una de reducción. Debido a los límites de espacio y alcance de este trabajo, no puedo intentar especificar cómo exactamente caracterizar la cercana relación entre esencia y modalidad—pero ciertamente tomo ello como un interesante proyecto.

### 3. Una teoría de la modalidad: *estructuralismo absolutista*

The virtual is not opposed to the real but to the actual. The virtual is fully real in so far as it is virtual [...] Indeed, the virtual must be defined as strictly a part of the real object — as though the object had one part of itself in the virtual into which it plunged as though into an objective dimension [...] The reality of the virtual consists of the differential elements and relations along with the singular points which correspond to them. *The reality of the virtual is structure.* We must avoid giving the elements and relations that form a structure an actuality which they do not have, and withdrawing from them a reality which they have.

—Gilles Deleuze, *Difference and Repetition*  
(Citado en Manuel DeLanda, *Intensive Science and Virtual Philosophy*)

#### 3.1. formulando afirmaciones sobre la estructura fundamental

**S**i la modalidad no necesita de nada para surgir, ¿cómo podemos caracterizarla? En el capítulo anterior argumenté contra las teorías no-no y una teoría reductivista. También asumí el objetivismo. Parece entonces que hemos librado el camino hacia la idea de que lo modal es un aspecto fundamental del mundo. Mi acercamiento aquí asume que uno debería tomar a la modalidad como *ese aspecto de la estructura metafísica que constriñe cómo las cosas podrían combinarse entre sí*. Pero ¿cómo vamos a modelar *constreñimientos*? Antes que eso, ¿cómo vamos a modelar *constreñimientos fundamentales*?

Podríamos tomar una teoría contemporánea sobre la fundamentalidad metafísica en general, y entonces encajar ahí a la modalidad, argumentando que las afirmaciones generales sobre la teoría (o al menos la mayoría de ellas) acerca de la estructura fundamental se aplican a la estructura modal. Por ejemplo, podríamos tomar el “operador estructural” de Sider (usaré ‘ $\Sigma$ ’ para denotarlo) y decir que  $\Sigma(\phi)$ —que el operador de posibilidad “recorta en las juntas” de la realidad (Sider 2012).

De acuerdo a Sider (2012, §6.3),  $\Sigma$  se une directamente a cualquier ítem lingüístico  $i$ , de cualquier categoría sintáctica, y devuelve una oración bien formada  $\Sigma(i)$ . Aunque ni el inglés ni el español tienen un tipo de expresión correspondiente, esto no es obviamente un problema para el acercamiento—es, al menos, nada obvio que las *teorías fundamentales* deberían escribirse con las reglas gramáticas de algún lenguaje natural particular, y tampoco es claro que la teorización sobre lo fundamental se vuelva ininteligible una vez que uno las deja de lado (después de todo, el sánscrito y el chino son inteligibles, incluso con sus gramáticas tan distintas de nuestra lengua).<sup>55</sup>

55 Sider explora algunas maneras en que podrías iluminar más a la sintaxis de  $\Sigma$  sin alejarnos demasiado de la sin-

¿Cuál es la carga semántica del operador? En particular, ¿cómo vamos a modelar lo que se requiere del mundo para que  $\Sigma(\diamond)$ ? En el marco sidereano,  $\Sigma(\diamond)$  es verdad sii hay al menos una verdad,  $S$ , involucrando ideología modal, tal que  $S$  no tiene una *semántica metafísica*. Es decir,  $\Sigma(\diamond)$  es verdad sii para al menos una verdad  $S$ , no hay un análisis de las condiciones de verdad de  $S$ , dado en términos de pura ideología que recorta en las juntas, y que no incorpore ninguna ideología modal (2012, §6.4).

Ahora bien, como sucede con el caso del acercamiento de Sider a los cuantificadores de primer orden, que sea verdad que  $\Sigma(\diamond)$  no necesita implicar que *el contenido semántico* de  $\diamond$  es una entidad fundamental (p. 91):

La semántica es, como las otras ciencias especiales, no fundamental. [...] Pero si la semántica no es fundamental, los hechos acerca de recortes en las juntas no pueden involucrar, fundamentalmente, entidades semánticas.

[...] Por razones semánticas, sería un error pensar a la estructura como concerniente a los ítems lingüísticos. [...] El hecho (si es que es un hecho) de los cuantificadores de primer orden recortan en las juntas no es un hecho acerca de los ítems lingüísticos “hay” y “para todos”. Es un hecho acerca del mundo—específicamente, acerca de su aspecto cuantificacional.\*

Desafortunadamente, Sider no dice mucho más que esto acerca de la estructura cuantificacional (aunque véanse §§7.7 y 10.6 de su libro para temas relacionados). Pero si asumimos tanto que la lógica clásica recorta en las juntas como que *ésa* afirmación no debería tomarse como un compromiso con la fundamentalidad de la estructura modelo-teórica de primer orden o de cualquier otra carga semántica propuesta para los cuantificadores, entonces, ¿cómo hemos de entender la fundamentalidad de la estructura cuantificacional?

Ahora esbozo una postura que podría resultar prometedora. Cuando queramos decir que una cierta pieza de ideología lógica  $i$  recorte en las juntas, una condición necesaria para hacer eso sería tomar la axiomatización de  $i$  como describiendo maneras en las que la estructura fundamental se comporta. Imaginemos que tomamos al condicional material  $\rightarrow$  como parte de nuestra ideología fundamental. Entonces podríamos decir que si el mundo incluye estructura al estilo proposicional o de hechos, y  $\alpha$ ,  $\beta$ , y  $\gamma$  son variables proposicionales (o de hechos) en el meta-lenguaje de nuestra teoría fundamental, el mundo es fundamentalmente tal que: siempre que  $\alpha \rightarrow \beta$  y el mundo es tal que  $\alpha$ , el mundo es tal que  $\beta$ ; el mundo es siempre tal que  $\alpha \rightarrow (\beta \rightarrow \alpha)$ ; y siempre que  $\alpha \rightarrow (\beta \rightarrow \gamma)$ ,  $(\alpha \rightarrow \beta) \rightarrow (\alpha \rightarrow \gamma)$ .<sup>56</sup> Hay que admitir que esto no encaja muy bien

---

taxis de los lenguajes naturales como el inglés o el español. Una que parece prometedora es tomarla como un operador unario de *oraciones*. Añadiríamos entonces a nuestro lenguaje, para cada categoría sintáctica, un conjunto de variables “dummy” que no pueden ligarse, y expresaríamos afirmaciones sobre fundamentalidad usándolas: con las variables “dummy” en negritas, por ejemplo,  $\Sigma(F\mathbf{x})$  dice que  $F$  recorta en las juntas y  $\Sigma(\Box\mathbf{p})$  dice lo mismo para el operador “caja”.

\* “Semantics is, like other special sciences, not fundamental. [...] But if semantics isn’t fundamental, the facts about carving at the joints can’t fundamentally involve semantic entities.

[...] For similar reasons, it would be a mistake to think of structure as concerning linguistic items. [...] The fact (if it is a fact) that the first-order quantifiers carve at the joints isn’t a fact about the linguistic items ‘there is’ and ‘for all’. It’s a fact about the world—specifically, its quantificational aspect.”

56 Como una nota secundaria, hay que ver cómo es que aceptar la fundamentalidad de una pieza de ideología no es una cuestión aislada: si queremos aceptar al condicional material  $\rightarrow$  en su interpretación estándar como na pieza de ideología fundamental, estaríamos comprometidos directamente con toda la estructura involucrada en su ca-

con el rechazo de Sider de la fundamentalidad de la metalógica (*ibid.*, §10.3). Pero si no podemos entender a la fundamentalidad de un ítem lógico en términos de el que su definición imponga constreñimientos sobre la estructura fundamental, me es muy poco claro cómo es que hemos de entender el que la lógica recorte en las juntas.

Usaré este acercamiento. Una vez que hemos tomado a  $\diamond$  como un operador que recorta en las juntas, es natural tomar a la estructura que se necesita para dar su definición semántica (a su vez dictada por la interpretación intencional de los axiomas relevantes) como recortante-en-las-juntas. Entonces, si la mejor teoría de las condiciones de verdad para enunciados de la forma  $\diamond\alpha$  han de darse en términos de un pluriverso de mundos posibles obedeciendo los axiomas de la teoría kripkeana de modelos, entonces la estructura fundamental del mundo incluye (como subestructura) la estructura de un marco kripkeano.<sup>57</sup>

Así que una vez que hemos aceptado la fundamentalidad de la modalidad (como el recortar-en-las-juntas de los operadores modales), queda por ver con qué tipo de estructura estamos comprometidos por tal aceptación. Lo que sigue es precisamente mi intento de esbozar qué tipo de estructura podría hacer el trabajo.

Antes de ello, notaré que la interpretación más directa de las estructuras formales que siguen es en términos de una *jerarquía metafísica*, estructurada por constructores como relaciones de dependencia metafísica o fundamentación, además de la estructura metafísica. Esta interpretación está en una fuerte tensión con la postura de Sider, quien rechaza el modelo jerárquico con sólo el nivel fundamental y su acercamiento basado en la “*metaphysical semantics*” (2012, caps. 5 y 8).

Sin embargo, el acercamiento desarrollado a continuación no necesita contradecir al acercamiento sidereano. Pues las estructuras desarrolladas a continuación también se pueden interpretar, algo forzosamente, como postulando sólo el nivel fundamental, mientras que la estructura jerárquica puede dejarse *para nuestras representaciones*. Es decir, tenemos una ontología plana, “*flat*”, y el orden de fundamentalidad ordena entonces no la ontología—*sino nuestras representaciones de ella*. La jerarquía sería entonces una *jerarquía de condiciones de verdad para lenguajes cada vez menos fundamentales* (cf. también Williams 2010). Así, los constreñimientos modales no actuarían desenvolviendo a la realidad desde el nivel fundamental, *sino a las representaciones posibles de ella*. Esto no necesariamente implica un anti-realismo de la modalidad, pues los constreñimientos modales todavía podrían estar presentes a nivel fundamental—y de hecho el absolutismo está comprometido con ello.

Estas son dos maneras divergentes en que se puede interpretar filosóficamente al formalismo, y muy probablemente haya más, igualmente interesantes. Aquí ya no podré tocar esa cuestión. Finalmente, y yendo, digamos, “desde el otro lado”, me gustaría notar que no creo que tomar a

---

racterización. (Pero ¿sólo en la axiomatización al estilo Hilbert? ¿Por qué no sistemas de secuentes a lo Gentzen? ¿Por qué no usar semántica algebraica? Estos son cosas donde un cambio en el enfoque parece traducirse en un cambio en los compromisos estructurales.) Notemos también que *no* he asumido que las nociones de *verdad* o de *interpretación* deberían aparecer en la caracterización relevante.

57 Así que la idea aquí es que los axiomas modales describen parcialmente a la estructura fundamental de la realidad. Este pensamiento está todavía algo indefinido—pues, ¿qué axiomas modales tenemos en mente aquí? Además, una vez que hemos fijado los axiomas de una lógica modal proposicional como nuestros axiomas fundamentales, ¿qué pasa con la lógica de primer, o incluso de alto, nivel? Estas son preguntas profundas que no puedo responder aquí.

la modalidad como absolutamente fundamental lo comprometa a uno al siguiente desarrollo. Es sólo una manera en que la postura puede desarrollarse—y, creo, una manera particularmente fructífera e interesante—pero seguramente no la única.

### 3.2. una teoría formal para el estructuralismo absolutista

Si es que voy por buen camino (el camino que lleva a la verdad), la modalidad—los hechos acerca de qué es posible y qué es necesario—no tiene que ver directamente con entidades en particular, sino con la estructura fundamental de la realidad—la estructura que la realidad tiene de manera *ontológicamente anterior* a la existencia de cualquier cosa en particular.

¿Qué es esta estructura fundamental de la realidad? De acuerdo a desarrollos recientes en la metafísica, la realidad tiene una cierta forma, una cierta *estructura* que, en el orden de la explicación metafísica, está fijada *antes*—en el orden explicativo—que la existencia y características de cualesquiera cosas que la realidad pueda contener. La realidad se despliega, en el orden ontológico, desde su nivel fundamental, para constituir la variedad de objetos, cualidades, relaciones, cantidades, hechos, etcétera, entre los cuales están los objetos que podemos experimentar.

Siguiendo estas ideas todavía muy generales, la idea es que hay *constreñimientos*—reglas, leyes, principios, limitantes—*modales*—que determinan lo que es posible, lo que es necesario, y lo que es de hecho (*actual*)—*fundamentales*—es decir, que no están, a su vez, determinados por otros principios gobernando a la realidad, o por otras estructuras, o por otras cualidades que pueda tener la realidad, o por otros objetos que puedan aparecer en esta. Los constreñimientos modales fundamentales determinan la posibilidad al determinar relaciones primitivas de *compatibilidad*. Generalizamos y hacemos explícitos, así, las relaciones de compatibilidad que necesitan las teorías no-no, y el esencialismo reductivista que revisamos en el capítulo 2; también las relaciones primitivas de compatibilidad que Wang (2013) ha argumentado que están implícitas en el combinatorialismo de Armstrong. Tomamos, entonces, una estructura modal primitiva—esa estructura que resulta de los constreñimientos modales fundamentales—que se hace explícita en relaciones primitivas de compatibilidad. Es por estas relaciones que hay posibilidad y necesidad en el mundo—no en el sentido en que la posibilidad y la necesidad se *reduzca* a ellas, pues las mismas relaciones ya son completamente modales, ni, por la misma razón, en el sentido en que la posibilidad o necesidad de un objeto (un hecho, la ejemplificación de una propiedad, la existencia de un objeto, la verdad de una proposición) se *fundamente parcialmente* en el aspecto no modal de tales relaciones—pues no hay tal aspecto.

La idea, entonces, es representar la *fundamentalidad* de los constreñimientos modales construyendo un modelo formal que (1) haga explícito que tales constreñimientos no se reducen a otro aspecto o componente del mundo, y (2) muestre cómo tales constreñimientos funcionan desde el nivel ontológico fundamental. El nivel ontológico fundamental, por ponerlo en una metáfora, se inscribe en una estructura metafísica que ya es, de suyo, modal.

Asumiré la existencia, en el nivel fundamental, de tres tipos de objetos que pueden poblar al mundo: el de *propiedades*, el de *constructores*, y el de *individuos*.<sup>58</sup> La existencia de tales tipos da

<sup>58</sup> Ver Armstrong 1980 y Lewis 1986, 1984, para argumentos en favor de la distinción entre fundamentalidad relativa para propiedades. Ver Bennett 2010, Fine 1991, Fine 2010, y Schaffer 2009 para argumentos a favor de cosas

un primer componente de la estructura fundamental: la estructura de los posibles tipos de entidades. Asumiré que tales tipos son disjuntos: ningún individuo es un constructor, ninguna cualidad es un individuo, etcétera;<sup>59</sup> también que al menos uno de los tres tipos no es vacío.<sup>60</sup>

Queremos que nuestro marco básico respete las tesis fundamentales del estructuralismo primitivista: en primera, claro, la idea central: la modalidad no es reducible a, y ni siquiera está en parte fundamentada en, cualquier otro aspecto del mundo: la modalidad es *absolutamente* fundamental, primitiva. En segunda, la idea más concreta (de espíritu estructuralista) de que las nociones metafísico-modales reflejan constreñimientos en la existencia de las cosas. Una tercera tesis es que los constreñimientos modales fundamentales actúan directamente sobre los componentes fundamentales de la realidad (y, ya asumiendo una ontología jerárquica, tales constreñimientos también funcionan en los niveles menos fundamentales).

Lo que ahora haré es construir un marco formal para estas tesis. Para dar algo de familiaridad con el acercamiento, lo introduciré con un “toy model”, un ejemplo simplificado.<sup>61</sup> El marco formal, que es una llana estructura matemática, resultará en un modelo intencional para un lenguaje que dejaré sin explorar en esta tesis. (Es como si primero tuviésemos ideas intuitivas acerca de la modalidad, luego pasáramos a formalizar esas ideas en un modelo intencional que ejemplifica una cierta estructura matemática (los marcos de Kripke), y dijéramos que tal formalización es el modelo intencional de un lenguaje que podríamos no especificar—pero que, ahora sabemos, es el lenguaje estándar de la lógica modal.) Sin embargo, al final de la construcción, comentaré sobre qué lenguaje parece ser necesario para describir la estructura resultante.

La idea básica es que tenemos la ontología fundamental,  $\lambda_0$ , que, como vimos arriba, se compone de tres tipos de objetos: individuos, propiedades y constructores. Representaré esos tipos de objetos con los conjuntos  $I_0$  (los individuos fundamentales),  $P_0$  (las propiedades fundamentales) y  $B_0$  (los constructores fundamentales). Así que:

$$\lambda_0 = \langle I_0, P_0, B_0 \rangle$$

La ontología fundamental,  $\lambda_0$ , se compone tanto de las entidades fundamentales como de la existencia de los tres posibles tipos de objetos—o *categorías ontológicas*.<sup>62</sup> (De nuevo, que sean tres es algo que puede modificarse, en desarrollos posteriores de este acercamiento.) Abajo se

que caen bajo lo que aquí he llamado “constructores” (Bennett les llama “builders”, Fine (1991) les llama “constructors”; Fine (2010) apunta a aplicaciones *generativas* de operaciones de partes (*parthood*), y Schaffer acepta a la fundamentación como imponiendo una estructura metafísica ordenada, análoga de manera relevante a lo que estudiaremos al postular constructores.) Finalmente, para una postura de acuerdo a la cual *hay exactamente un* individuo fundamental, ver Schaffer 2009, 2010; para una postura de acuerdo a la cual *todo* individuo es fundamental, ver Adams 1979; para una postura de acuerdo a la cual no hay *ningún* individuo fundamental, ver Dasgupta 2009.

59 Creo que estas asunciones pueden, coherentemente, acomodarse a un amplio rango de teorías acerca de la naturaleza de cada uno de los tipos: por ejemplo, una teoría de tropos podría decir que hay individuos que son cualidades, o una teoría del haz (*bundle*) podría decir que los individuos son haces de cualidades. Sin embargo, explorar las diferentes combinatorias sería una larga distracción del tema central.

60 ¿Qué pasa si la base fundamental está vacía, si no contiene objetos de ningún tipo? ¿Todavía habría distinciones modales en el mundo? Esta postura es compatible con una respuesta positiva—pues las leyes, o *constreñimientos*, que se aplican a nada todavía pueden ser leyes.

61 Estoy muy agradecido con Alessandro Torza por haber notado la conveniencia de definir primero un *toy model* para aclarar el funcionamiento del marco formal que propondré.

62 Este acercamiento está muy influido por el acercamiento de Fine (1991) a las “*constructional ontologies*”, ver pp. 286-287.

notará que es útil reunir todas las entidades fundamentales en un solo conjunto, al que llamaré  $\mathbf{E}_0$ . Por definición:

$$\mathbf{E}_0 = \mathbf{I}_0 \cup \mathbf{P}_0 \cup \mathbf{B}_0,$$

que son entonces nuestras *entidades* fundamentales.

Veamos un ejemplo extremadamente simplificado. Supongamos que hubiese sólo dos individuos fundamentales, que llamaríamos Juan y Martha (que podrían ser, digamos, dos partículas fundamentales en una mecánica newtoniana). Supongamos, además, que hubiese también sólo dos propiedades fundamentales, Amarillo y Azul (y asumimos que podrían ser ejemplificadas por nuestros individuos fundamentales). Finalmente, supongamos que hubiese sólo tres constructores fundamentales: + (suma mereológica de individuos), & (fusión de propiedades), y  $\Pi$  (instanciación, que construye hechos a partir de cualidades y objetos). Así, tendríamos:

$$\lambda_0 = \langle \mathbf{I}_0, \mathbf{P}_0, \mathbf{B}_0 \rangle = \langle \{\text{Juan, Martha}\}, \{\text{Amarillo, Azul}\}, \{+, \&, \Pi\} \rangle$$

Con lo que:

$$\mathbf{E}_0 = \{\text{Juan, Martha, Amarillo, Azul, +, \&, \Pi}\} = \mathbf{I}_0 \cup \mathbf{P}_0 \cup \mathbf{B}_0$$

Ahora buscamos que esta base fundamental se despliegue para individuar objetos derivados, en un “proceso” limitado por los constreñimientos modales. Para ello, vamos a diseñar una estructura que represente los posibles despliegues individuativos desde la base fundamental—limitados por los constreñimientos. Tal estructura será un *árbol* (en el sentido de la teoría matemática de grafos) cuya raíz será el nivel fundamental, de donde “nacerán” las ramas, que representarán los despliegues posibles—con lo que cada rama será un mundo posible, y cada nivel en cada rama representará un nivel ontológico de cada mundo posible. (Si no asumimos que el nivel fundamental existe necesariamente, simplemente construimos diferentes árboles, con diferentes niveles fundamentales como raíces; el pluriverso de mundos posibles será entonces un *bosque*—un conjunto de árboles—determinado por los niveles fundamentales y los constreñimientos modales.)

Para ello, la raíz del *árbol modal* “contendrá” únicamente al nivel fundamental. Esto nos da una primera manera en que actúan los constreñimientos modales: asignando entidades que son metafísicamente compatibles en cada nivel. Vamos a necesitar un conjunto de puntos para representar niveles ontológicos. A tal conjunto lo llamaré  $N$  (abajo veremos cómo obtener su orden natural). Así, en el nivel 0 (el nivel ontológico fundamental), tendríamos:

$$d_0: N \rightarrow \mathbf{E}_0,$$

tal que, para exactamente un  $n_0 \in N$ :

$$d_0(n_0) = \mathbf{E}_0, \quad d_0 \text{ está indefinida en todo otro punto de } N.$$

Es decir, por lo visto arriba, las entidades fundamentales son asignadas como dominio—o, como lo llamaré para distinguirlo de otros dominios introducidos abajo, como *dominio de coexistencia*—al nivel fundamental:

$$d_0(n_0) = \{\text{Juan, Martha, Amarillo, Azul, +, \&, \Pi}\}$$

Ahora es relevante una segunda manera de actuar de los constreñimientos modales: al codificar la naturaleza modal de los constructores. Por ejemplo, la suma mereológica de individuos, +, *no puede* operar con un individuo y una propiedad (o una propiedad y un constructor), sino que toma dos individuos. Los constreñimientos modales, así, actúan de una segunda manera seleccionando a los objetos que son *compatibles-para-construcción* (ya no compatibles *respecto a su existencia*), en el sentido en que un constructor puede tomarlos para individuar nuevos objetos.



Para ver eso, tomemos a la suma mereológica,  $+$ . A este constructor lo representamos como una operación, que toma dos individuos e individua uno nuevo—la suma de aquellos. Así, por ejemplo, la suma de Juan y Martha, Juan+Martha, sería un tercer individuo, compuesto de Juan y de Martha, que llamaremos Josefina. Es decir: Juan+María = Josefina. Por lo que:

$$+: \{ \text{Juan, María, ...} \} \times \{ \text{Juan, María, ...} \} \rightarrow \{ \text{Josefina, ...} \}$$

En particular,

$$+ \langle \text{Juan, María} \rangle = \text{Josefina}$$

Josefina, al ser un nuevo objeto individuado en términos de Juan y María, es *derivado* de estos—Juan y María son *más fundamentales* que Josefina. Josefina, entonces, será parte de un nuevo nivel ontológico—ya no de  $\lambda_0$ .<sup>63</sup>

Pero  $+$  no puede operar con Juan y Amarillo: Juan+Amarillo no está definido; la suma mereológica de individuos, *necesariamente*, individua nuevos individuos *a partir de otros*. Los constreñimientos modales van a reflejar este tipo de hechos.

De manera análoga,  $\&$  individua nuevas propiedades a partir de otras dos. Así, por ejemplo, Azul $\&$ Amarillo = Verde. Con lo que:

$$\&: \{ \text{Azul, Amarillo, ...} \} \times \{ \text{Azul, Amarillo, ...} \} \rightarrow \{ \text{Verde, ...} \},$$

y:

$$\& \langle \text{Azul, Amarillo} \rangle = \text{Verde}$$

La acción de los constreñimientos modales aquí actuará seleccionando sólo propiedades como compatibles para ser tomadas por  $\&$  para individuar nuevas propiedades.

De igual manera, la instanciación ( $\Pi$ ) actúa tomando un individuo y una cualidad, creando con ello un hecho (o estado de cosas posible): Juan $\Pi$ Azul = el hecho de que Juan es Azul. Usando los corchetes '[' y ']' para representar hechos con la notación estándar de la lógica, tenemos:

$$\Pi: \{ \text{Juan, Azul, ...} \} \times \{ \text{Juan, Azul, ...} \} \rightarrow \{ [\text{Azul}(\text{Juan})], \dots \}$$

Y, en particular:

$$\Pi \langle \text{Juan, Azul} \rangle = [\text{Azul}(\text{Juan})]$$

En resumen: los constreñimientos modales actúan seleccionando no sólo entidades que son *compatibles* entre sí (que pueden coexistir), sino también entidades que son *co-constructibles* (que pueden, mediante la acción de los constructores, individuar nuevos objetos). Estos nuevos objetos—Josefina, Verde, el hecho de que Juan es Azul—son *derivados* de los objetos co-constructibles.

Así, definimos:

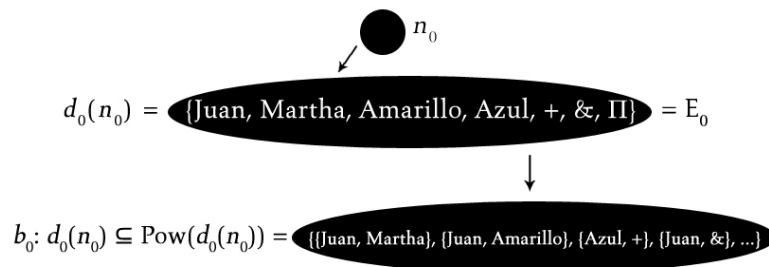
$$b_0: d_0(n_0) \subseteq \text{Pot}(d_0(n_0))$$

Si el *dominio de existencia* (dado por  $d_0$ ) del nivel fundamental  $n_0$  era {Juan, Martha, Amarillo, Azul,  $+$ ,  $\&$ ,  $\Pi$ }, el *dominio de construcción* ( $b_0$ ) de  $n_0$  será un conjunto que contiene conjuntos de entidades de  $n_0$ —estos conjuntos, a su vez, contendrán entidades que son co-constructibles: que pueden ser conjuntamente tomadas por los constructores para individuar nuevos objetos.

Hasta ahora, hemos visto cómo los constreñimientos modales determinan (1) qué objetos pueden existir con cuáles otros, y (2) cómo la individuación de nuevos objetos está modalmente limitada, a través de la limitación modal de la acción de los constructores. Así, hasta ahora, la

63 Este no es el caso general, pues podría suceder que existen todos (*wholes*) derivados de sus partes. (Por ejemplo, quizá un montoncito de arena depende de cada uno de los granos que lo componen). Hay que recordar que los ejemplos aquí son usados con un propósito meramente ilustrativo.

imagen que resultaría podría diagramarse así:



Pero esto resultará, como veremos ahora, en diferentes *posibles* despliegues desde el nivel fundamental hasta el nivel derivado.

Para ello, asumimos que  $B_0$  no está vacío. Cada uno de sus elementos es una operación  $m$ -ádica que toma  $m$  objetos co-constructibles e individua con ellos un nuevo objeto, uno *derivado*. Es decir (usaré  $\beta$  como variable para constructores) :

$$\beta^m: X^m \rightarrow E_1, \quad \text{para algún } X \in b_0(d_0(n_0)),$$

donde  $E_1$  es el conjunto de entidades individuadas (construidas) a partir de las entidades del nivel fundamental, que agrupamos en el conjunto  $E_0$ . Así, como vimos arriba, la suma mereológica + partía del conjunto  $\{\text{Juan, María}\}^2$  y nos devolvía a Josefina—un objeto derivado, agrupado ahora en  $E_1$ .

Lo siguiente es reunir todas las entidades individuadas por los constructores fundamentales en el conjunto  $E_1$ , lo cual hacemos mediante la ecuación:

$$B_0(n_0) = \bigcup_{\beta \in B_0} \beta^m(X^m) = E_1 \text{ para algún } X \in b_0(d_0(n_0))$$

Es decir, uniendo en un sólo conjunto a todos los objetos individuados por cada constructor fundamental a partir del dominio de construcción del nivel  $n_0$ , obtenemos a exactamente los objetos del primer nivel (si quisiéramos dejar lugar para un tipo fuerte de emergencia, en la ecuación de arriba diríamos que  $B_0(n_0) \subset E_1$ ). Al conjunto  $B_0(n_0)$  lo llamamos *la imagen construida de  $n_0$* .

Las entidades en  $E_1$ , a su vez, serán ya sea individuos, propiedades o nuevos constructores (recordemos que arriba asumimos que era parte de la estructura de la realidad que sólo contenía tales tipos de objetos; pero un desarrollo posterior de estas ideas podría quitar estas categorías y poner otras en su lugar). Es decir, estamos asumiendo que  $E_1$  se descompone en tres conjuntos disjuntos:

$$E_1 = I_1 \cup P_1 \cup B_1$$

Formalmente hablando, esta simplemente es una partición arbitraria en tres subconjuntos de  $E_1$ . Pero, *metafísicamente* hablando (en el modelo intencional), lo que determina esta partición es la naturaleza metafísica de los objetos individuados y agrupados en  $E_1$ : si un objeto individuado (a partir de otros en  $E_0$ ) es una propiedad, caerá en  $P_1$ , etcétera.

Ahora bien, así como  $d_0$  le asignaba al nivel fundamental  $n_0$  un dominio de cosas que eran compatibles para coexistir, en el siguiente nivel—el nivel 1— $d_1$  asignará también dominios de cosas compatibles-respecto-a-coexistencia. Pero aquí ya hay diferentes maneras en que se pueden distribuir los objetos. Quizá haya objetos derivados que no puedan coexistir entre sí, pero sí con otros. (Lo cual, metafóricamente hablando, nos diría que el germen de la incompatibilidad estaba ya en los objetos fundamentales). Esto nos da la primera ramificación del árbol modal: di-

ferentes posibles distribuciones de objetos derivados, fijadas relativo a cuáles de ellos son compatibles o incompatibles entre sí. Para ello, representamos los constreñimientos modales en este primer nivel como:

$d_1: N \rightarrow \text{Pot}(\mathbf{E}_1)$ , donde  $d_1$  es una función parcial, y tal que  $d_1(n_0)$  está indefinido.

Intuitivamente, los niveles a los que  $d_1$  les asigna entidades derivadas (*i.e.*, miembros de  $\mathbf{E}_1$ ) son diferentes despliegues posibles desde el nivel fundamental. El conjunto  $d_1(n_i)$  de entidades de  $\mathbf{E}_1$  será un conjunto de entidades compatibles entre sí; mientras que  $d_1(n_i)$  será un conjunto de entidades tales que, en conjunto, son incompatibles con las entidades en  $d_1(n_j)$  tomadas también en conjunto (para  $i \neq j$ ).

Que  $d_1$  sea una función parcial representa el hecho de que sólo estará definida para niveles derivados del nivel fundamental (y no para el nivel fundamental), pues justamente le asigna a estos entidades individuadas a partir de las entidades del nivel 0.

Ahora podemos ir trazando la estructura grafo-teórica de los árboles modales. Por ejemplo, definimos el conjunto de hijos del nodo raíz ( $n_0$ ) como:

$\text{hijos}(n_0) = \{v \in N : d_1(v) = X, \text{ para algún } X \text{ tal que: } X \cap \mathbf{B}_0(n_0) \neq \emptyset\}$

Es decir, los hijos del nodo  $n_0$  son aquellos niveles-mundo cuyo dominio de coexistencia contiene objetos tomados de la imagen construida de  $n_0$  (*i.e.*, de objetos individuados a partir de objetos en el nivel fundamental  $n_0$ ).<sup>64</sup>

Como anteriormente, los constreñimientos modales ahora “filtrarán” objetos que no sólo son compatibles-respecto-a-existencia, sino también respecto-a-construcción:

$b_1: d_1(n) \subseteq \text{Pot}(d_1(n))$ , para cada  $n$  para el cual  $d_1$  esté definida.

Y como arriba, cada constructor  $m$ -ádico (ahora de  $\mathbf{B}_1$ ) tomará  $m$  objetos compatibles-para-construcción e individualará con ellos un objeto derivado (de nivel 2):

$\beta^m: X^m \rightarrow \mathbf{E}_2$ , para algún  $X \in b_1(d_1(n_1))$ ,

Y volvemos a agrupar a todas las entidades individuadas por los constructores en  $\mathbf{B}_1$  a partir de entidades tomadas de un mundo-nivel  $n$  (de nivel 1) en el conjunto de entidades de nivel 2:

$\mathbf{B}_1(n) = \bigcup_{\beta \in \mathbf{B}_1} \beta^m(X^m) \subseteq \mathbf{E}_2$  para algún  $X \in b_1(d_1(n))$ ,  $n$  para el cual  $d_1$  esté definida.

Esto, notemos, ya no ocurre en un único mundo-nivel, sino en cada mundo-nivel  $n$  de nivel 2. Y como anteriormente, podemos definir la estructura grafo-teórica del árbol modal al definir el conjunto de hijos de cada nodo (*i.e.*, cada mundo-nivel). Ahora lo hacemos en el caso general:

$\text{hijos}(v_n) = \{w : d_{n+1}(w) = X, \text{ para algún } X \text{ tal que: } X \cap \mathbf{B}_n(v_n) \neq \emptyset\}$

Los hijos de un mundo-nivel  $v_n$ , que está en un nivel ontológico  $n$ , son aquellos mundo-niveles cuyo dominio de coexistencia es un conjunto que contiene objetos de la imagen construida de  $v_n$ : los objetos que coexisten en los hijos de  $v_n$  fueron individuados a partir de los constructores, tomando objetos desde  $v_n$  (*i.e.*, desde su dominio de construcción).

Ahora podemos dar una definición rigurosa de los mundos a un nivel, que casa con la estructura grafo-teórica de los árboles modales:

$(N \uparrow 0) = \{n_0\}$ ,

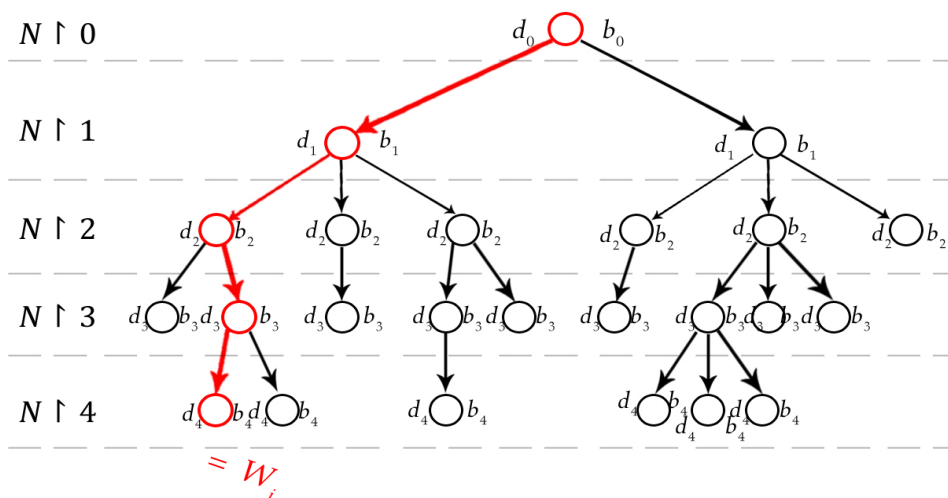
$(N \uparrow 1) = \text{hijos}(N \uparrow 0) = \text{hijos}(n_0)$

<sup>64</sup> Requiero solamente una intersección no vacía, en vez de pedir subconjuntos (*i.e.*, que todo elemento de  $X$  haya sido tomado de  $\mathbf{B}_0$ ), de nuevo, para dejar abierta la posibilidad del emergentismo. Si hay emergentismo, entonces en el primer nivel podrían surgir cosas que no se individualan a partir de los objetos fundamentales, y por ello no están en la imagen construida de  $n_0$ .

$(N \uparrow 2) = \text{hijos}(N \uparrow 1) = \text{hijos}(v) \cup \text{hijos}(w) \cup \dots$  para  $v, w, \dots \in \text{hijos}(n_0)$ ;  
en general,

$$(N \uparrow n) = \text{hijos}(N \uparrow n-1), \text{ para } n > 0.$$

(En el lenguaje de la teoría de grafos,  $(N \uparrow n)$  es simplemente el conjunto de *nodos de profundidad*  $n$ .) Esta estructura grafo-teórica da cuenta de la dependencia ontológica de los niveles derivados sobre los niveles más fundamentales. La representación es altamente perspicua: es desde la raíz del árbol, el nivel ontológico fundamental  $n_0$ , que todos los demás objetos son individuados, en construcciones limitadas por principios modales que dictan qué objetos pueden coexistir y qué objetos pueden dar lugar a nuevos objetos.<sup>65</sup> Con lo que la imagen ahora podría diagramarse así:



Habiendo definido los mundos a un nivel  $n$ , podemos dar la definición general de los dominios de coexistencia y de construcción, así como de las imágenes construidas.

*Dominios de coexistencia:*  $d_n: (N \uparrow n) \rightarrow \text{Pot}(\mathbf{E}_n)$   
(también podríamos haber dicho:  $d_n: N \rightarrow \text{Pot}(\mathbf{E}_n)$ , una función parcial indefinida para cualesquiera  $v \in N$  para los que alguna  $d_{m < n}$  esté definida)

*Dominios de construcción:*  $b_n: d_n(v) \subseteq \text{Pot}(d_n(v))$  para  $v \in N$  donde está definida  $d_n$ .

*Imágenes construidas:*  $\mathbf{B}_n(v) = \bigcup_{\beta \in \mathbf{B}_1} \beta^m(X^m) \subseteq \mathbf{E}_{n+1}$   
para algún  $X \in b_n(d_n(v))$ ,  $v$  para el cual  $d_n$  esté definida.

*Entidades del nivel n:*  $\mathbf{E}_n = \mathbf{I}_n \cup \mathbf{P}_n \cup \mathbf{B}_n$

*Ontología del nivel n:*  $\lambda_n = \langle \mathbf{I}_n, \mathbf{P}_n, \mathbf{B}_n \rangle$

Finalmente, definimos la función modal,  $\mathbf{F}$ , como la composición de los constreñimientos modales en cada nivel:

*Función modal:*  $\mathbf{F} = (b_n \circ d_n) \circ (b_{n-1} \circ d_{n-1}) \circ \dots \circ (b_1 \circ d_1) \circ (b_0 \circ d_0)$

La idea intuitiva detrás de  $\mathbf{F}$  es que esta codifica las “leyes”, principios o constreñimientos que determinan qué objetos son compatibles con cuáles (tanto respecto a coexistencia como respecto

<sup>65</sup> Por supuesto, estoy asumiendo el fundamentalismo ontológico: la idea de que *hay* un nivel fundamental, y que la dependencia nunca entra en un ciclo. Utilizando herramientas grafo-teóricas, es posible modelar tesis contrarias al fundamentalismo, como algún tipo de dependencia cíclica (“circular”).

a co-construcción).<sup>66</sup> Por ello, la acción de **F** es crucial para darle forma al espacio modal—representado aquí por el árbol modal. Así, **F** da estructura al espacio de las posibilidades—al pluriverso de mundos.

Finalmente, definimos a nuestros marcos jerárquicos **H** como las estructuras del tipo:

$$\mathbf{H} = \langle \lambda_0, \mathbf{E}, N \rangle,$$

Fuera de la aplicación metafísica que he hecho de estas estructuras, podríamos definir, en términos puramente matemáticos, a un marco jerárquico **H** como una tupla:

$$\mathbf{H} = \langle \mathbf{E}, \mathbf{B}, \mathbf{F}, N \rangle,$$

con:

**E**: un conjunto no vacío,

**B**  $\subseteq$  **E**: un conjunto de operaciones definidas sobre **E** y bajo las cuales **E** está cerrado,

**N**: un árbol etiquetado (*labeled tree*) cuyos nodos forman un conjunto no vacío disjuncto de **E**,

**F**: la composición de las funciones  $(b_n \circ d_n) \circ \dots \circ (b_0 \circ d_0)$ , definidas arriba.

En estas estructuras abstractas, la acción de **F** determinaría qué subconjunto de **E** cuenta como el conjunto de entidades fundamentales (a saber, aquellas que **F** le asigna a la raíz del árbol). En general, la acción de **F**, llevando del conjunto de entidades al árbol modal, determina los niveles ontológicos. **N** sería un árbol etiquetado y **F** la función de etiquetado correspondiente: a cada nodo de **N**, **F** lo etiqueta con un conjunto, que a su vez contiene otros dos (los que aquí llamamos su *dominio de coexistencia* y su *dominio de construcción*.) El dominio de coexistencia de cada nodo es un subconjunto de los resultados de las operaciones sobre el dominio de construcción del nodo padre. Todo esto siempre constreñido por **F**.

La raíz del árbol modal sería  $n_0$ , y sus ramas tendrían la forma:

$$\langle n_0, n_{1i}, n_{2ij}, n_{3ijk}, \dots \rangle,$$

con  $n_{1i}$  siendo el  $i$ -ésimo hijo de  $n_0$  (asumiendo un orden izquierda-derecha estándar sobre los hijos de cada nodo),  $n_{2ij}$  siendo el  $j$ -ésimo hijo de  $n_{1i}$ , y así por el estilo. Un *mundo posible* es entonces una secuencia de la forma:

$$\langle n_0, x \in \text{hijos}(n_0), y \in \text{hijos}(x), z \in \text{hijos}(y), \dots \rangle$$

Cada una de estas secuencias representa la manera en que nuestro mundo podría resultar estar constituido desde su nivel fundamental (compartido por todos los mundos del árbol modal) hasta cualquier nivel (contable) que llegue a alcanzar—*siempre constreñido por la función modal*.

Voy a finalizar este capítulo mostrando varias aplicaciones filosóficas de los marcos jerárquicos, que denotaré con **H**.

### 3.3. algunas aplicaciones de los marcos jerárquicos, y una comparación

En primera, cada marco jerárquico automáticamente determina un marco kripkeano para la lógica modal proposicional,  $\langle W, R \rangle$ , con el conjunto de mundos,  $W$ , definido, como arriba, como el conjunto de las ramas del árbol modal (o, si nos decidimos por tomar al nivel fundamental como

---

<sup>66</sup> He usado comillas para hablar de *leyes* de una manera menos comprometedora; pero en realidad yo me tomo la idea bastante en serio: ver §3.3, abajo.

contingente y entonces tomamos un pluriverso constituido por un bosque de árboles modales,  $W$  sería el conjunto de las ramas de cada uno de tales árboles). Además,  $\mathbf{H}$  determina un marco kripkeano para la lógica modal de primer orden estándar (con dominios variantes):  $\langle W, R, I, Q \rangle$ , donde  $I = \mathbf{I}$  (el conjunto de individuos tanto derivados como fundamentales, *i.e.*,  $\bigcup_{0 \leq i} \mathbf{I}_i$ ) y  $D$  es una función que le asigna a cada  $w \in W$  un conjunto  $D(w) \subseteq \mathbf{I}$  de individuos que se representan como existiendo en  $w$ , donde  $D(w)$  está determinado por la acción de  $\mathbf{F}$  y de  $\mathbf{B}$  en la manera esbozada arriba.<sup>67</sup>

Una vez que  $\mathbf{H}$  determina marcos kripkeanos para la lógica modal proposicional y la lógica modal cuantificacional, podemos simplemente copiar las cláusulas de verdad estándar. Por ejemplo, las cláusulas cruciales para los operadores modales:

$$w \models_a \Box \alpha \text{ sii para todo } u \in W \text{ tal que } R(w, u): u \models_a \alpha$$

$$w \models_a \Diamond \alpha \text{ sii para algún } u \in W \text{ tal que } R(w, u): u \models_a \alpha$$

Así, como suele suceder, *la teoría sucesora incluye a la anterior como un caso especial*. En este caso, *los marcos jerárquicos incluyen a los marcos kripkeanos como sub-estructura*.

Ahora tengo que dar cuenta de una primera y obvia nota sobre la terminología de “mundos posibles” que he estado usando. En contraste con el ersatzismo en un uso ontológico-fundador, la verdad modal no se representa aquí como fundada en la verdad-en-un-mundo, en primera porque *los mundos mismos son determinados por la función modal*. Por ello, al final del día, *lo que determina a la verdad modal es la función modal misma*—y, como la función modal es una representación de los constreñimientos modales, *verdad-de-acuerdo-a-la-función-modal resulta en verdad-de-acuerdo-a-los-constreñimientos-modales*, que no es una base de fundamentación, sino la verdad modal misma.

La interpretación más directa del formalismo, entonces, nos invita a poner todo el peso ontológico de la modalidad sobre los hombros de la función modal.<sup>68</sup> Con lo que es la función modal la que debería ser el objeto del estudio metafísico: los constreñimientos que representa son lo que determinan qué posibilidades hay, dada una base fundamental. Podríamos volvernos ficcionalistas, reductivistas, agnósticos, ... etcétera, acerca de los mundos posibles, dejando intacta la realidad objetiva y básica de la estructura modal—*pues nada metafísicamente importante depende*

67 Más definiciones sobre  $\mathbf{H}$  y  $\mathbf{B}$  nos darán una función de interpretación “natural” para los predicados de un lenguaje modal de primer orden. Si, tomando en serio el ejemplo de arriba, reservamos algún  $\pi^n \in \mathbf{B}_n$ , para cada  $n$ ,  $i$ , y lo tomamos como el “operador de instanciación”, construyendo estados de cosas a partir de relaciones  $n$ -ádicas y  $n$  individuos, entonces  $\pi^n \langle i_1, \dots, i_n, P \rangle$  podría tomarse como diciendo que los individuos  $i_1, \dots, i_n$  están relacionados por  $P$ ; y entonces, la función de interpretación “natural”  $\varphi_w(F)$ , para  $F$  un símbolo relacional  $n$ -ádico seleccionado para representar a  $P$  y  $w$  un mundo, reflejaría al operador de predicación por:

$$\varphi_w(F) = \{ \langle i_1, \dots, i_n \rangle : \langle i_1, \dots, i_n \rangle \in D(w)^n \text{ y } \pi^n(i_1, \dots, i_n, P) \in \mathbf{B}_{m+1}(w), \text{ para } w \in (W \cdot m) \}.$$

Esto podría llevar a algunos a preocuparse sobre regresos estilo Bradley para la noción de *constructor*. Desafortunadamente, aunque esta es una cuestión importante, debo dejarla de lado en este trabajo; pero véase Bennett 2010 para una solución a una preocupación parecida respecto a la noción de *fundamentación*; ver también Schlieder 2004.

68 Una réplica confundida que podría hacerse aquí, apuntaría al hecho de que las funciones dependen de sus dominios y codominios, pues su reducción conjunto-teórica es un conjunto de pares de *inputs* y *outputs*, y los conjuntos son entidades paradigmáticamente dependientes. El problema con esta objeción potencial es que reifica, de manera equivocada, a los dispositivos representacionales—la dependencia de  $\mathbf{F}$ , *qua* función, en  $W$  es lo que Kaplan llamó un *artefacto* del modelo: “When we construct a model of something, we must distinguish those features of the model which represent features of that which we model, from those features which are intrinsic to the model and play no representational role. The latter are *artifacts of the model*” (1975: 722.)

de qué sean los mundos posibles. Estos son usados con una mera función representacional, y no se pretende darle a este aparato un uso ontológico.

La estructura adicional nos da la posibilidad de añadir dispositivos expresivos extra a nuestro lenguaje modal. Por ejemplo, un marco jerárquico  $\mathbf{H}$  tiene la estructura suficiente para dar modelos de la lógica de la esencia de Fine (Fine 2000; cf. §2.4.1, arriba). Es decir, los modelos de Fine también son una *sub-estructura* de los marcos jerárquicos.

Para la condición de verdad crucial, VERDAD ESENCIAL, sólo necesitamos un buen “sustituto” de la relación  $\geq$  de Fine, en el sentido en que  $x \geq y$  sii  $xRy$ , donde  $R$  es nuestro “sustituto”—una relación constructible dentro de nuestro marco formal.

Recordemos que, según la define Fine,  $\geq$  es simplemente tal que:

- es una relación transitiva y reflexiva en el conjunto de individuos en cualquier mundo;
- cada mundo está cerrado con respecto a  $\geq$ .

Definimos tal relación en nuestro marco con:

$x$  depende<sup>0</sup> de  $y$  sii  $x = y$ ,

$x$  depende<sup>1</sup> de  $y$  sii  $\beta i_1, \dots, y, \dots, i_n = x$  (para algún  $\beta \in \mathbf{B}$  y cada  $i \in \mathbf{E}$ ),

$x$  depende<sup>2</sup> de  $y$  sii  $\beta i_1, \dots, z, \dots, i_n = x$ , y  $\beta'(j_1, \dots, y, \dots, j_m) = z$ ,

en general,

$x$  depende <sup>$n$</sup>  de  $y$  sii para algún  $z$ ,  $z$  depende <sup>$n-1$</sup>  de  $y$ , y  $x = \beta i_1, \dots, z, \dots, i_n$ .

$x$  depende<sup>+</sup> de  $y$  sii para algún  $n$ ,  $x$  depende <sup>$n$</sup>  de  $y$ .<sup>69</sup>

Para que la relación *depende<sup>+</sup> de* sea un buen sustituto de  $\geq$  no sólo formalmente, sino *en el modelo intencional* de la lógica de Fine (en el sentido en que  $x$  dependa ontológicamente de  $y$  sii  $x$  depende<sup>+</sup> de  $y$ ), simplemente añadimos (consistentemente con el anti-reductivismo de la esencia a la necesidad) ideología esencialista:  *$x$  depende<sup>+</sup> de  $y$  sii  $y$  ocurre en la definición real de  $x$* . Con esto, podemos dar cuenta de la asimetría entre Sócrates y {Sócrates}, de la misma manera en que Fine lo hace (§2.4.1, arriba).

También tenemos modelos para la lógica pura del fundamento de Fine (la “*pure logic of ground*”, ver Fine 2012a). Tomemos a los operadores  $n$ -ádicos de “instanciación” de la nota 67. (Uso comillas porque aunque se podrían interpretar así asumiendo una metafísica de universales, la interpretación podría variar si las propiedades fundamentales resultaran ser, por ejemplo, tropos.) Tales operadores,  $\pi^n$ , construyen estados de cosas a partir de  $n-1$  entidades y una relación (del orden y aridad apropiadas). Si tenemos otra familia de constructores,  $\{\theta \in \mathbf{B}_i\}$ , los *operadores de fusión* que construyen hechos compuestos (en  $\mathbf{E}_i$ ) a partir de hechos atómicos (en  $\mathbf{E}_{i-1}$ ), podemos definir a un marco factual generalizado, a lo Fine, como:

$\langle F, V, \Xi \rangle$ ,

donde:

$F = \bigcup_i (\Pi_i(\mathbf{E}_i))$ ,

$\Pi_i = \pi_i^1 \circ \pi_i^2 \circ \dots$ ,

$\Xi = \bigcup_i (\theta(\mathbf{E}_i^2))$ ,

$V = \{v \subseteq F: \text{la fusión de cada subconjunto } v' \text{ de } v \in v, v' \neq \emptyset\}$ .

En general, las fusiones  $n$ -ádicas son:

69 Por supuesto, esto sólo nos da la estructura formal de la lógica de Fine. Como presioné en §2.3.2, para interpretarla, además de ideología esencialista, necesitamos ideología puramente *modal*.

$$(\dots((f_1 \theta f_2) \theta f_3) \theta \dots) \theta f_n.$$

Ahora bien, la fusión de los conjuntos de hechos  $F_1.F_2. \dots$  es el conjunto de las fusiones de los hechos dentro de cada conjunto, *i.e.*,

$$F_1.F_2. \dots = \{f_1 \theta f_2 \theta \dots : f_1 \in F_1, f_2 \in F_2, \dots\}.$$

Finalmente,

el conjunto de los conjuntos de hechos  $\{F_1, F_2, \dots\}$  es un *fundamento débil completo* para el conjunto  $G$  de hechos si  $F_1.F_2. \dots \subseteq G$ .

Esta definición básica nos permite emular las cláusulas de verdad que Fine da para su lógica.

Finalmente, el marco tiene aplicaciones fuera de la “metafísica pura”. Por ejemplo, también contiene la suficiente estructura para fundamentar la propuesta neo-logicista en filosofía de las matemáticas de Linnebo (2009). Esta propuesta resuelve un importante problema para el neologicismo (el *problema de la mala compañía*) postulando una estructura bien fundada para los “procesos” de individuación de las entidades matemáticas. Veamos con un poco de más detalle (*cf.* §§1-7 de Linnebo 2009).

La idea central del neologicismo es obtener objetos matemáticos a partir de abstracción sobre “conceptos” (propiedades y relaciones). Tales principios tienen la forma:

$$(\text{ABS}(\alpha) = \text{ABS}(\beta)) \leftrightarrow (\alpha \sim \beta),$$

donde  $\alpha$  y  $\beta$  son variables de algún tipo,  $\sim$  es una relación de equivalencia sobre las entidades sobre las que corren esas variables, y  $\text{ABS}(\alpha)$  es el abstracto de  $\alpha$  con respecto a  $\sim$ .<sup>70</sup> Un problema fundamental es que no todo principio de abstracción es consistente.<sup>71</sup> Así, el neologicista debe mostrar cuáles principios de abstracción son consistentes, filosóficamente motivados, y suficientemente poderosos como para fundamentar a las matemáticas.

La propuesta de Linnebo es tomar a la individuación de objetos matemáticos como un “proceso” bien fundado, realizado por etapas, de tal manera que los principios de abstracción sólo permitan individuar objetos con base en objetos ya individuados. Este proceso está constreñido por cinco *Principios de Individuación*:

- (1) Las entidades se introducen sucesivamente mediante una serie bien ordenada de etapas.
- (2) La introducción de una entidad consiste en la especificación de su condición de identidad.
- (3) Toda entidad mantiene su condición de identidad a través del resto del proceso.
- (4) *Fundamentación*. La condición de identidad de una entidad sólo puede presuponer entidades individuadas antes de esta.
- (5) *Libertad Futura*. Ninguna entidad puede ser individuada si restringe la libertad futura para individuar nuevas entidades de acuerdo a los Principios de Individuación.

Estos principios definen una relación de accesibilidad entre las etapas de individuación, que será: (a) reflexiva, (b) anti-simétrica, (c) transitiva, (d) dirigida, y (e) bien fundada. Con esto, podemos dar marcos kripkeanos para un lenguaje modal—que obedecerá, debido a los constreñimientos

<sup>70</sup> El ejemplo clásico es el famoso Principio de Hume:

$$\#F = \#G \leftrightarrow F \approx G,$$

que nos dice: el número de cosas que son  $F$  = el número de cosas que son  $G$ , sii hay una función uno-a-uno de los  $F$ s con los  $G$ s ( $F \approx G$ ).

<sup>71</sup> El ejemplo clásico es la famosa, pero desafortunadamente inconsistente, Ley Básica V de Frege:

$$\text{extensión}(F) = \text{extensión}(G) \leftrightarrow \forall x(Fx \leftrightarrow Gx),$$

que permite la derivación de la paradoja de Russell.



sobre la accesibilidad, los axiomas de la lógica  $S4.2$ .

Linnebo propone construir modelos para esta lógica, que además satisfagan una variedad de axiomas que constriñen la manera en que los objetos se abstraen de otros, además de las maneras en que se individuán conceptos. En particular, Linnebo se inclina por condiciones de comprensión que restringen la individuación de conceptos a aquellos procesos que toman objetos que hayan sido individuados previamente. Así, Linnebo postula el *principio de fundamentación*, que requiere a las condiciones que individuán conceptos, que estén determinadas por entidades ya individuadas.

Con las condiciones fundamentadas Linnebo construye modelos para la clase de teorías que incluyan los axiomas mencionados. Primero, define *dominios de abstracción*, conjuntos que especifican los *tipos* y su extensión. Sobre estos dominios, define *reportes de individuación*, que son conjuntos especificando qué entidades han sido individuadas, y cómo. Por ejemplo, si un reporte contiene a  $x$ , esto representa el hecho de que  $x$  ha sido individuado; y si contiene a una tripleta de la forma  $\langle \mathbf{R}, F, x \rangle$  esto representa que el abstracto de  $F$  bajo  $\mathbf{R}$  es  $x$  (p. 383). Linnebo pone dos requisitos sobre los reportes de individuación: (i) cuando un reporte represente a una cosa como habiendo sido individuada, debe especificar cómo fue individuada, y (ii) cuando represente a una cosa como el abstracto de un concepto bajo una relación, debe representar al concepto y a la relación como ya individuados.

El siguiente paso es utilizar los reportes de individuación para interpretar a las fórmulas de nuestro lenguaje de la manera obvia. Por ejemplo, " $\text{Abst}_{\mathbf{R}}(F, x)$ " será satisfecha por un reporte  $I$  sii  $I$  contiene a la tripleta  $\langle \mathbf{S}, G, y \rangle$  (donde  $\mathbf{S}$  es la denotación de " $\mathbf{R}$ ",  $G$  la de " $F$ " y  $y$  la de " $x$ "). Donde  $[[\varphi(x)]]_I$  es el conjunto de objetos del dominio  $D$  que satisfacen la condición  $\varphi$  cuando esta se interpreta en  $I$ , e  $I \leq J$  dice que *el reporte  $J$  extiende al reporte  $I$* , es decir, que todas las entidades individuadas según  $I$  están individuadas también de acuerdo a  $J$  (y en la misma forma),<sup>72</sup> podemos decir que una condición  $\varphi(x)$  *está determinada en  $D$  por  $I$*  cuando:

$$\forall J (I \leq J \rightarrow [[\varphi(x)]]_I = [[\varphi(x)]]_J),$$

es decir, cuando todo reporte de individuación que extiende a  $I$  le asigna a la condición  $\varphi(x)$  la misma extensión que  $I$ .

Con esto, Linnebo probará algunos lemas que muestran que las condiciones que están determinadas por un reporte son adecuadas para individuar conceptos, además de un teorema según el cual se pueden construir secuencias de reportes de individuación sobre un dominio de abstracción infinito. Lo importante de tales secuencias es que (1) en el límite, todas las entidades del dominio se representan como individuadas, y (2) con ellas se puede construir un modelo de Kripke para la teoría de la abstracción sin condiciones de comprensión.

Ahora por fin podemos ver cómo el acercamiento de Linnebo resulta ser una aplicación de los marcos jerárquicos que aquí definí. Primero, tenemos marcos jerárquicos en el sentido de la estructura abstracta que definí arriba. Como vimos, cada mundo posible de estos marcos será una secuencia con la forma:

$$\langle n_0, x \in \text{hijos}(n_0), y \in \text{hijos}(x), z \in \text{hijos}(y), \dots \rangle$$

Donde, para todo  $n \in$  el árbol  $N$ ,  $x \in \text{hijos}(n)$  sii el dominio de coexistencia de  $x$  contiene objetos de la imagen construida de  $n$ —es decir, hay objetos que coexisten en  $x$  que fueron construi-

<sup>72</sup> En particular, si  $I \leq J$ , entonces todo concepto que según  $I$  se aplica a ciertas entidades, se aplicará a exactamente esas mismas entidades según  $J$ .

dos a partir de objetos que coexistían en  $n$ .

La idea ahora es tomar cada uno de estos mundos como “procesos” de individuación de objetos matemáticos, donde lo que arriba llamé “constructores” en esta aplicación serán operaciones de abstracción o de individuación de conceptos, por ejemplo. Estos mundos nos darán un árbol modal en el sentido definido arriba. Este árbol tendrá etapas de individuación (los segmentos de cada rama), ordenadas por un orden reflexivo, anti-simétrico, transitivo, dirigido, y bien fundado (en el nivel  $n_0$ , la raíz). Así, nuestra función modal toma los cinco principios de individuación de Linnebo como constreñimientos modales.

A su vez, los reportes de individuación de Linnebo simplemente toman elementos de los dominios de coexistencia de cada etapa. Con esto, podemos construir modelos para las teorías de la abstracción en exactamente la misma manera en que Linnebo lo hace.

Hemos visto, entonces, que los marcos jerárquicos son una estructura con la que podemos construir modelos de Kripke para la lógica modal, la lógica de la esencia, la lógica de la fundamentación, y las teorías neo-logicistas sobre la individuación de objetos matemáticos. Me parece que esto nos da razones de peso para pensar que el aparato de marcos jerárquicos es un marco formal, por derecho propio, apropiado para explorar la cuestión de la *individuación de objetos contruidos a partir de otros, cuando esta es constreñida por leyes o principios*. Esto le da un aspecto interesantemente general, y le gana, creo, un lugar en el “*toolbox*” de la metafísica.

Una posibilidad todavía abierta y bastante interesante es que la estructura arbórea que postula este acercamiento nos permita describir y explorar aspectos de los pluriversos determinados por los marcos jerárquicos de manera sistemática: por ejemplo, en la ciencia teórica de la computación se están explorando algoritmos para medir distancias entre árboles (Bille 2005). Quizá esto nos dé alguna esperanza de basar en ello una métrica de *distancia entre mundos* basada en algo más que nuestros juicios subjetivos de parecidos; pero aquí todavía falta bastante por explorar.

Como se ve, he preferido explorar de manera somera más el poder teórico-filosófico de los marcos jerárquicos que sus características lógico-matemáticas (digamos, al construir un lenguaje formal con relaciones de completud y corrección respecto a esta maquinaria). Pero ahora doy una nota sobre este aspecto. La estructura resultante—los *árboles modales* resultantes de la acción de **H** y **B**—puede llegar a ser muy similar a las estructuras semánticas de la lógica de tiempo ramificante. De hecho, no es difícil llegar a la hipótesis de que el lenguaje de la lógica de árboles computacionales (CTL), o uno muy cercano (quizá, por ejemplo, el cálculo  $\mu$ ), servirá para describir a la estructura resultante en los marcos jerárquicos. Pero dejo estas cuestiones para un lugar más apropiado donde explorar aspectos más profundamente lógicos de este acercamiento.

No quiero dejar de notar que, según mi postura, *los mundos posibles son sólo dispositivos representacionales*. Ellos no son necesarios para fundamentar a la modalidad—por la simple razón de que ellos mismos están fundamentados en ella. Como vimos arriba, el único trabajo teórico para el que son invocados los mundos posibles, es para *representar* despliegues posibles desde el nivel ontológico fundamental. Si los mundos posibles son sólo dispositivos representacionales, no necesitamos hacerles espacio en nuestra clasificación ontológica fundamental. Todavía más, *ellos pueden ser casi cualquier cosa que queramos*, dado que *nada ontológicamente sustantivo depende de su naturaleza intrínseca*. Lo que lleva todo el peso ontológico son los constreñimientos modales en juego, codificados en la función modal.

Antes de pasar a la última parte de este capítulo (con la que también cerramos la parte metafí-

sica de esta tesis), podría servir hacer una breve comparación de mi versión del absolutismo con otras posturas acerca de la modalidad que no he tocado arriba.

Primero que todo, están las que tomo como parientes cercanos de mi postura: la de DeRosset, la de Maudlin, la de Roy, y la de Wang (DeRosset 2012; Maudlin 2007, caps. 1, 2; Roy 1993; Wang 2013).

Bajo la postura de DeRosset (2012), los mundos posibles no son la fuente de la modalidad; en lugar de eso, son modelos de Kripke interpretados de tal manera que representen explícitamente a los hechos modales. No es claro cómo DeRosset entiende a tales hechos, pero la idea parece ser que tenemos medios independientes de los mundos posibles de llegar a conocer tales hechos modales, que después codificamos en el lenguaje natural, que a su vez usamos para interpretar a los modelos. Creo que algo parecido puede decirse de mi postura. Los mundos posibles, en cuanto medios representacionales, se usan para modelar hechos modales. Sin embargo, la maquinaria que diseñé es, como vimos arriba, bastante más general y permite estudiar las cuestiones interrelacionadas de la individuación, la fundamentalidad, y la modalidad.

Bajo la postura de Maudlin, las leyes naturales son primitivas. Tengo una enorme simpatía por el programa naturalista de Maudlin, y creo que mi postura es en principio compatible con ella. Por ejemplo, algo que no defendí en este trabajo, pero que me parece altamente atractivo, es tomar a lo que he venido llamado “constreñimientos modales” como simplemente las leyes de la naturaleza. (Por supuesto, muchos filósofos creen que se puede mostrar que las leyes de la naturaleza son metafísicamente contingentes mediante argumentos de concebibilidad, pero soy altamente escéptico de los intentos de extraer conclusiones metafísicas desde lo que podemos imaginar.) Así, el despliegue del mundo desde su aspecto fundamental está guiado y limitado—por las leyes. La maquinaria formal que desarrollé anteriormente está diseñada para ser compatible con un acercamiento naturalista de este tipo; pero *no* necesita de él. Otras cosas, como las “leyes metafísicas” para quien crea en ellas, podrían codificarse en la función modal *F*.

Bajo la postura de Roy, la modalidad *no* está fundamentada en los mundos posibles, sino que surge cuando las estructuras de las propiedades—que resultan ser dadas por sus análisis—se dejan fijas. La modalidad es aquello que fija las relaciones de (in)compatibilidad o implicación entre tales estructuras—es decir, lo que es posible o necesario está fijado por el análisis de las propiedades, que, en este caso, daría un ejemplo de lo que he venido llamando “constreñimientos modales”. Mi postura es una generalización de esta idea; así que el acercamiento de Roy puede resultar como aplicación del mío.

Bajo la postura de Wang (que ella llama *primitivismo de incompatibilidad*), la idea es ser combinatorialistas al estilo de Armstrong, pero aceptando que los hechos acerca de qué propiedades son *compatibles* con cuáles otras son primitivos, y determinan los demás hechos modales. De nuevo, mi postura generaliza esta idea, de tal manera que los hechos primitivos de compatibilidad podrían verse como un caso de los constreñimientos modales primitivos de mi postura.

Aunque iluminadoras, si nos enfrascáramos en una “disputa interna”, argüiría que mi versión particular del primitivismo (el estructuralismo absolutista) tiene algunas ventajas: es más explícito cuando representan los constreñimientos modales fundamentales, tiene interacciones interesantes con la metafísica fundamental, y al menos podemos ver cómo llegaría a tener conexiones interesantes con estructuras profundamente estudiadas en la ciencia de la computación (la teoría de los árboles y la CTL). De hecho, creo que mi postura podría verse como abstrayendo

algo más severamente que las anteriores: podríamos tomar a los “constreñimientos modales” como simplemente leyes *naturales*, o leyes “metafísicas” que determinan, ya sea hechos de incompatibilidad entre propiedades (o incluso individuos), o a los mismos mundos posibles.

La postura presentada aquí contrasta más claramente con otros dos tipos de posturas. El primer tipo es, obviamente, cualquier postura *reductivista*, como el *realismo modal* de Lewis (1986) y el esencialismo de Fine (que revisamos en esta tesis: §2.4).

En esta tesis no he dado ningún argumento contra la reducción lewisiana. Aunque el realismo lewisiano ha estado bajo el microscopio conceptual desde su nacimiento hace casi medio siglo, para muchos filósofos es todavía una opción teórica abierta. Por mi parte, el sistema de Lewis me parece, visto muy en abstracto, sospechoso en dos puntos importantes (en temas de detalle habrá otros acuerdos o desacuerdos): en primera, su postulado central (la *superveniencia humeana*) tiene consecuencias altamente cuestionables desde el punto de vista de la física contemporánea (Maudlin 2007, cap. 2). En segunda, su idea de “pagar con la moneda de la ontología” para lograr “una mejora en lo que Quine llama ideología” (1986: 4) me parece altamente debatible en el contexto de su teoría: ¿por qué deberíamos considerar natural expandir la ontología con sus mundos para evitar introducir ideología modal primitiva? ¿En realidad nos es menos clara la ideología modal primitiva, que la existencia de sus mundos? ¿Por qué es un menor costo teórico una ontología tan expandida, que una ideología con un primitivo más?

Bueno, estas preguntas no sirven para refutar a Lewis, eso me queda claro. De hecho, no he ofrecido *argumentos* contra Lewis (sólo he citado a Maudlin—y en secciones anteriores a Lycan y Divers-Melia, quienes dudan de la reducción lewisiana de la modalidad). No lo haré en este trabajo. De hecho, lo que me interesa aquí no es una reacción *de frente* a la, digamos, *pintoresca* ontología de Lewis; sino buscar una imagen alternativa de la modalidad que iguale, o al menos rivalice cercanamente, a la imagen de Lewis con respecto a sus beneficios teóricos. Por todo lo que puedo ver, una vez que hemos notado los problemas con las posturas no-no y con el reduccionismo esencialista, el primitivismo acerca de la modalidad es un contendiente serio—del lado *realista*.

Pues el segundo caso de contraste es, por supuesto, con cualquier clase de postura *anti-realista*. En esta tesis simplemente restringí la discusión al campo *realista*, con una motivación desde el realismo científico, como hice claro al inicio del capítulo 2.

Teniendo pendiente el desarrollo de la dialéctica, entonces, la elección a tomarse en este punto entre el anti-realismo y el realismo tendrá que ser motivada por razones algo holistas—como qué tan bien cada tipo de teoría recupera intuiciones básicas acerca de la modalidad, o interactúa con teorías en otras áreas de la metafísica, o teorías empíricas. Como señalé al inicio del capítulo, parece que el realismo científico puede dar un mejor fundamento al realismo de la modalidad. Esta sería otra área interesante para investigar.

### 34. estructuralismo absolutista y las esencias

El último punto a tocar en la parte metafísica de esta tesis es, por supuesto, la relación entre el estructuralismo absolutista y las esencias. Como vimos arriba, nuestra mejor teoría disponible

acerca de la esencialidad (la teoría definicional de Fine) es incapaz de reducir la modalidad a la esencialidad. Fine también arguyó que la esencialidad era irreductible a la modalidad. Como los argumentos de Fine no dependían de la existencia de mundos posibles, o de cualesquiera otras entidades de alguna teoría no-no, sino sólo de la existencia de diferencias y propiedades modales, parece claro que se aplican también a *cualquier* intento de reducir la esencialidad a la modalidad.

Pero ¿queda alguna relación interesante entre el esencialismo y la modalidad? En esta última sección argüiré que sí y delinearé maneras en que podríamos teorizar esta relación (de nuevo y tristemente, por razones de espacio y enfoque tendré que dejar un tratamiento más profundo de estas cuestiones para futuros trabajos).

*La estructura modal del mundo determina la distribución de esencias* (si las hay, por supuesto). En la pintoresca pero iluminativa metáfora clásica, pensemos que Dios tiene frente a sí las esencias de *toda* cosa. Entonces Dios primero “fija” la estructura modal del mundo, y luego “inserta” esas esencias en los lugares y con las relaciones que esa estructura permite. La modalidad determina qué es compatible, e incompatible, con qué; en particular, determina cuáles esencias (¡si es que las hay!) pueden coexistir con otras.

Esto se anunciaba ya desde el argumento que di contra la reducción fineana, cuando argumenté que tal reducción no puede lograrse porque presupone hechos de *compatibilidad* que *no* son reductibles a hechos esencialistas—y, sobre todo, que son hechos *modales*. La idea es que los hechos modales *no* son lo suficientemente finos al discriminar entre lo que es *esencialmente necesario* y lo que es *accidentalmente necesario*. Decimos que para un  $x$  es *esencialmente necesario* que  $p$  cuando la esencia de  $x$  y la estructura modal en conjunto fuerzan a  $p$ ; mientras que a  $x$  le es *accidentalmente necesario* que  $p$  cuando es sólo gracias a hechos modales (determinados, por supuesto, por la estructura modal) que  $p$  y, además, esos hechos y  $p$  son compatibles con la existencia de  $x$ .<sup>73</sup>

Esto lleva naturalmente a la idea, algo vaga, de que, aunque la modalidad y la esencialidad no son reductibles una a la otra, la intuición detrás de los intentos reductivistas (ya fueran los de la concepción estándar de la esencialidad, o de la teoría de Fine) tenía algo de correcto: que ambas esencialidad y modalidad son estructuras muy emparentadas entre sí, que ambas “trabajan en conjunto”. No me queda claro cómo precisar esta idea, pero parece ir por buen camino.

Con esto cierro el capítulo; también concluyo así la parte metafísica de la tesis. En la mitad que sigue, me embarcaré en una investigación sobre la cuestión de la posibilidad del conocimiento modal y de esencias, así como su estructura. Poco de lo que diga entonces supone más que el realismo de la modalidad—la mayoría ni siquiera supone el primitivismo modal.

---

73 Esto está en una analogía muy cercana a las nociones fineanas de *esencia constitutiva* y *esencia consecucional* (Fine 1997); estas, como nota Fine, corresponden más o menos a las nociones tradicionales de esencia y *propia*.

II. EPISTEMOLOGÍA:  
CONOCIMIENTO, FIABILIDAD Y REPRESENTACIÓN

#### 4. Contra el escepticismo del conocimiento modal y de esencias

[...] *philosophy-hating philosophers* (a common breed) [...]  
—T. Williamson, *The Philosophy of Philosophy*

Dejando de lado al escepticismo acerca del mundo externo, sé que estoy sentado en esta silla. De manera parecida, dejando de lado al escepticismo acerca de la modalidad, sé que esta silla podría haberse puesto medio metro más a la derecha. Dejando de lado los escepticismos, sé tanto que los solteros son hombres que no se han casado, como que es *necesario* que, si  $x$  es un soltero,  $x$  es un hombre que no se ha casado. ¿Hay alguna diferencia radical entre estos dos tipos de conocimiento? A primera vista, no lo parece. Parece que tengo ambos tipos de conocimiento, y ambos pueden llegar a ser importantes en mi vida práctica y/o intelectual.

Es cierto que sé que estoy sentado en esta silla sobre bases perceptuales: veo y siento mi estar sentado en ella; y no veo ni siento (al menos no kinestésicamente, como con mi estar sentado) que la silla esté posiblemente a medio metro más a la derecha. Pero, he visto algo parecido a ello: la silla estando *de hecho* medio metro más a la derecha. Así que sé que la realidad permanece coherente cuando esta silla se acomoda a medio metro más a la derecha—lo cual parece ser, plausiblemente, mi razón para fundamentar *ése* hecho como una posibilidad.

Así que tenemos una manera de conocer posibilidades: conociendo hechos. Lo que es real también es, por supuesto, posible; así que si sé que algo ha sido el caso o es el caso (o será el caso en el futuro), por ello estoy justificado en creer que sé que *ése* algo no sólo es real, sino también posible.

Hay más. Supongamos que Bonnie, una de mis perritas, tuvo un accidente hace unos meses. Salió de la casa de manera furtiva, y un coche la atropelló accidentalmente (afortunadamente, ella ahora está bien). No parece muy controvertido afirmar que sé que Bonnie *podría* no haber sufrido tal accidente (falta, por supuesto, saber cómo es que sé eso). Eso es conocimiento de *meras* posibilidades: posibilidades que no son y no serán reales. Esta vez no infiero posibilidad de mi conocimiento de la realidad solamente, ya sea de hechos presentes, pasados o futuros. ¿Cómo entonces se supone que puedo lograr tal conocimiento?

Un problema todavía más enrevesado es cómo obtenemos—si es que lo obtenemos—conocimiento de posibilidades que están menos directamente “unidas a la experiencia”. Pues probablemente sea teóricamente importante distinguir entre dos tipos de conocimiento: *a priori* y *a posteriori*. Tradicionalmente y sin entrar en detalles,  $S$  sabe *a priori* que  $p$  sii  $S$  sabe que  $p$  y la justificación de su creencia no depende de factores empíricos, fuera de aquellos factores que fueron necesarios para entender que  $p$ ; mientras que  $S$  sabe *a posteriori* que  $p$  sii  $s$  sabe que  $p$  y no lo sabe *a priori*. Entonces una pregunta que surge es ¿cómo podríamos conocer, de manera *a priori*, proposiciones meramente posibles o necesarias? Sabemos  $\langle 5 + 7 = 12 \rangle$ , sabemos  $\langle$  todos los subgrupos de un grupo abeliano libre son, a su vez, abelianos y libres  $\rangle$ , y sabemos que esas proposiciones de las matemáticas son *necesarias*. Pero, plausiblemente, nunca hemos visto números o grupos abelianos. Entonces, ¿cómo sabemos que tales hechos, conocidos *a priori*, son

*necesarios?*

Notemos que el conocimiento de hechos necesarios necesita el conocimiento de las posibilidades, pues hay una equivalencia extensional: aquello que es necesario es aquello que es el caso en toda posibilidad. Si tenemos razones para sostener que un hecho no se da en una posibilidad, tenemos razones para creer que el hecho no es necesario. Entonces, si tuviésemos una teoría satisfactoria de cómo podemos obtener conocimiento de posibilidades, quizá podríamos extenderla al conocimiento de necesidades: saber que  $\langle p \rangle$  es necesaria se reduce a saber que  $\langle p \rangle$  es verdad en toda posibilidad. Por supuesto, todavía faltaría explicar cómo sabemos que algo es verdad en *toda* posibilidad, pero esto, una vez habiendo explicado el conocimiento de posibilidades, se reduciría a explicar el conocimiento de hechos *totales*—hechos acerca de que tales o cuales objetos son *todos* los objetos de un tipo específico. Pero ésa es una tarea epistemológica que no es distintivamente modal.<sup>74</sup>

La pregunta es entonces cómo podemos conocer posibilidades—cómo podemos, como diré en adelante, tener *conocimiento modal*—fuera de sólo inferirlas de hechos reales o de variaciones mínimas en los hechos reales. Esto, claro, presupone que *podemos* conocer posibilidades.

En este capítulo revisaré dos propuestas escépticas, de acuerdo a las cuales tenemos mucho menos conocimiento de las posibilidades (y, con ello, de las necesidades) de lo que muchos filósofos creen. Primero consideraré las motivaciones que me parece que tenemos para cargar la balanza, de entrada, a favor del anti-escepticismo modal—de tal manera que el escéptico modal sea la persona con la carga de la prueba. Después consideraré críticamente dos propuestas escépticas—una de Peter van Inwagen y otra de Robert Nozick. Aunque, como argumentaré, no acepto la letra de sus propuestas, creo que en su “espíritu” hay motivaciones interesantes—que usaré para extraer tres problemas que acotan el espacio de epistemologías modales aceptables.

Una pregunta que surge es si hay alguna analogía teóricamente interesante entre el caso del conocimiento modal y el caso del conocimiento de las esencias. Pues, si aceptamos con Kit Fine que la esencialidad no se reduce a la necesidad (cf. §2.4.1, arriba), y aceptamos mi argumento que concluye que la necesidad no se reduce a la esencialidad (§§2.4.2., 2.4.3), entonces esto parece implicar que la epistemología de la necesidad y la epistemología de la esencialidad divergen—la manera en que expliquemos el conocimiento de posibilidades no será la misma que la manera en que expliquemos el conocimiento de las esencias de las cosas.

Encuentro este argumento plausible; pero no creo que sea lo único que haya que decir sobre la epistemología de la esencialidad. Pues incluso Fine acepta que la esencialidad es necesaria—es decir, que si a  $x$  le es esencial ser  $E$ , entonces es *necesario* que, si  $x$  existe,  $x$  es  $E$ . La necesidad metafísica es, pues, una condición necesaria para la esencialidad. Recordemos, además, que la propuesta de Fine consiste en asumir que la esencialidad es primitiva, pero puede entenderse bajo el modelo de la definición real: aquello que le es esencial a  $x$  es aquello que define lo que  $x$  es (por supuesto, asumiendo el primitivismo esencialista, el sentido de este “ser” de  $x$  no es otro que aquello que  $x$  es *por naturaleza*, o *esencialmente*). Así, una primera idea sobre la forma de una epistemología de la esencialidad sería la siguiente:  $S$  sabe que  $y$  es esencialmente  $E$  sii (i)  $S$  sabe

---

<sup>74</sup> Agradezco a Melahuac Hernández y a Areli Figueroa sus comentarios aquí, que me llevaron a aclarar este párrafo.



que necesariamente  $E(y)$  (aquí apelaríamos a la  $X$  de arriba), y (ii)  $S$  sabe que  $E$  es parte de la *definición real* de  $x$ .

Por supuesto, la carga explicativa se pasa del lado de cómo podemos saber que una propiedad define a un objeto (véase el capítulo 6, abajo). Pero, por ahora, esto nos indica que un escepticismo modal afectaría también a la tesis de que conocemos esencias. Pues si no tenemos el suficiente conocimiento modal, parecería entonces que hay casos en los que no podemos conocer las suficientes posibilidades para afirmar que algo es necesariamente  $E$ , y, por ello, que ése algo es esencialmente  $E$ .

#### 4.1. motivaciones para el anti-escepticismo modal

Hay que notar que nuestra supuesta posesión de conocimiento modal no es sólo un supuesto hecho acerca de la cotidianidad—como mi supuesto conocimiento modal acerca de mi perrita—sino también un hecho central *para la filosofía*. Muchas teorías filosóficas, de variante complejidad, presuponen que podemos tener conocimiento de hechos modales. Esto se argumenta de dos maneras.

En primera, muchas teorías filosóficas—especialmente las metafísicas—buscan afirmar hechos modales acerca del mundo. Es, de acuerdo a la platonista, necesario que existan universales; es contingente, de acuerdo al dualista, que la mente esté unida a la materia. Incluso cuando la modalidad es más débil que la metafísica, siempre puede entenderse como una *restricción* de la modalidad metafísica: por ejemplo, el epistemólogo dirá que es *esencial a nuestro conocimiento* el que sea un estado fáctico (*i.e.*, que si se lo tiene respecto a una proposición  $p$ , ésta sea el caso), y por lo tanto que es necesario que, si existe tal estado, este sea fáctico. Incluso tesis filosóficas que no hablen de la “esencia” de tal o cual fenómeno pueden tener compromiso con el conocimiento de hechos modales. Una filósofa del lenguaje podría afirmar que es un hecho contingente acerca del lenguaje el que sus usuarios cometan errores de ejecución. Pero entonces, sean cuales sean las razones para ello, esto equivaldría a verse comprometida con la idea de que tales errores *no necesitan* suceder para que exista tal o cual lenguaje natural. Así que, si la filosofía—como se la entiende tradicionalmente—ha de ser posible, debemos decir todavía mucho sobre qué es y cómo es el conocimiento modal.

De todo esto podemos ver que, aunque parece que tenemos conocimiento modal de algunos tipos—del tipo, por ejemplo, de mi conocimiento de que mi silla podría haber sido movida medio metro a la derecha—y aunque los filósofos asuman que tienen mucho conocimiento modal, nos falta una teoría adecuada acerca de cómo eso puede ser el caso. Sin ella, el escepticismo es un peligro latente.

Así, pues, tenemos dos motivaciones para ser, de entrada, anti-escepticos sobre el conocimiento modal: (i) en la vida cotidiana parecemos tener tal conocimiento, y (ii) una vasta cantidad de teorías filosóficas lo suponen. Además, para evitar el escepticismo acerca de ese supuesto conocimiento, requerimos una teoría semántica (y, con ello, de una ontológica) y una teoría epistemológica, y además requerimos que se condigan. He explorado el primer requisito en la primera

parte de la tesis; esta segunda parte se dedica al segundo y tercer requisitos.

Pero ¿qué tal si resultara que no tenemos tal conocimiento? ¿Por qué debería eso importarnos? Algunos filósofos se verían inclinados a razonar así:

Si la filosofía presupone conocimiento modal, entonces, dado que *no* tenemos tal conocimiento, ¡peor para ella! Eso muestra que o debemos dejar de hacer filosofía, o debemos dejar de hacerla con asunciones de conocimiento modal.

Creo que esta es una manera errada de ver las cosas. En la vida cotidiana, el valor del conocimiento de posibilidades no es trivial. Tomemos el caso de mi perrita Bonnie. Después de su accidente, *baso mis acciones posteriores en parte sobre ese aprendizaje*: como sé que si no *hubiera* escapado eso no *hubiera* pasado, me prevengo de tales y cuales maneras, para ya no dejarla escapar. El conocimiento modal puede servir para *guiar acción futura*—como todo otro conocimiento.

Al inicio del capítulo 3 noté algo que la filosofía de la ciencia contemporánea ha venido a reconocer: que la física postula *espacios de posibilidad*, y de una manera tal que es implausible suponer que tales espacios pueden eliminarse de tales teorías sin pérdida de poder explicativo. Dado esto, en conjunto con el hecho de que la física es el paradigma de una ciencia exitosa, podríamos obtener la base para un argumento a favor de la cognoscibilidad de las posibilidades—a saber, al hacer ciencia, cuando esta postula espacios de posibilidad. Aunque este potencial argumento me parece muy atractivo, sé que podría no resultar suficiente. Pues un escéptico podría todavía replicar que eso sólo muestra que un área altamente especializada—la física matemática—puede conocer posibilidades; pero que todavía queda por ver si eso es posible fuera de ella.

Así que es importante notar que, según nos dice la evidencia de la psicología, una buena parte de nuestra vida (diaria y teórica) sería radicalmente distinta (*si es que* aún existiera) si no echara mano de nuestras creencias modales. Dado esto, suponer que no tenemos conocimiento modal es ir contra la evidencia científica—así, el conocimiento modal gana credenciales naturalistas. Veamos.

La literatura sobre el pensamiento modal en psicología se originó en el trabajo clásico de Kahneman y Tversky (1982; 1986). Es importante notar que se ha descubierto que nuestra imaginación está constreñida de tal manera que las maneras más inmediatas en que imaginamos “mundos posibles” se guían bajo regularidades—la imaginación no necesariamente es un facultad desaforada, sin guía alguna (Byrne 2002, 2005: cap. 3; cf. Williamson 2007: cap. 5). Según algunas hipótesis, el pensamiento sobre posibilidades contrafácticas es necesario para la atribución de intencionalidad (Nanay 2010a, 2010b). Las creencias modales—que en psicología se llama “razonamiento contrafáctico”—también intervienen en el aprendizaje a partir de la experiencia, tanto individual como grupal (Segura & Morris 2005). Además, el razonamiento contrafáctico está fuertemente involucrado en la *evaluación* de nuestra experiencia: los contrafácticos “*upward*” (las creencias de que un mejor “mundo posible” podría haber resultado), sirven para aprender de los errores; mientras que los “*downward*” (creencias de que un *peor* “mundo posible” podría haber resultado) tiene efectos emocionales, y nos llevan a sentirnos mejor (Markman & McMullen 2005). También sabemos (Decety & Stevens 2009) que antes de ejecutar una acción, “planeamos” mentalmente el movimiento, partiendo de nuestra posición actual (informados por la

propiocepción) y diseñamos un plan mental de cómo ese estado puede ser transformado por el cuerpo en el estado deseado después del movimiento. Estas representaciones mentales son una simulación del sistema sensorimotor, y representan trayectorias *posibles*—así, el conocimiento modal está involucrado en la base de las acciones. La evidencia también indica que la simulación más fina de cursos de acción inmediata rastrea la línea entre expertos y novatos—donde tal simulación es de acciones *posibles*, aunque regularmente ejecutadas inmediatamente—de tal manera que los expertos han “automatizado” su conocimiento práctico, lo que lo cierra relativamente a la introspección (Beilock & Lyons 2009). Además, el razonamiento contrafáctico está crucialmente involucrado en las creencias causales (Spellman *et al.*, 2005) y, por supuesto, en la imaginación y la creatividad (Wong *et al.* 2009).

Una cuestión empírica todavía abierta y profundamente interesante es si la imaginación de escenarios posibles guarda alguna relación no trivial con otras facultades psicológicas—incluso si comparten una red funcional en la arquitectura cerebral. Buckner y Carroll (2006) argumentan que esa es una hipótesis plausible: dada la evidencia de las ciencias cognitivas, parece existir un sistema cerebral (involucrando a los lóbulos frontal y temporal-parietal medial) que subyace a varios mecanismos psicológicos, relacionando a la memoria episódica con la imaginación de posibles episodios futuros (conocida como “prospección”, *prospection*), la teoría de la mente (la capacidad de concebir los estados mentales de otras personas), y la facultad de navegación (la capacidad de guiarse en el entorno). La idea es que la memoria *construye* la representación interna de perspectivas alternativas (p. 55):

Pensar acerca del futuro, recordar episódicamente, concebir la perspectiva de otros (teoría de la mente) y la navegación ponen a trabajar esta red [cerebral], lo que sugiere que se basan de manera similar en modos internos de cognición y en sistemas cerebrales que permiten la percepción de puntos de vista [*vantage points*] alternativos. Quizá estas habilidades, tradicionalmente consideradas como distintas, son mejor vistas como parte de una clase más amplia de función que permite formas flexibles de auto-proyección. En esta postura, la auto-proyección se basa cercanamente sobre los sistemas de memoria porque la experiencia pasada sirve como el fundamento sobre el cual las perspectivas alternativas y los futuros concebidos se construyen. [...]

Sospechamos que el valor adaptativo de la memoria episódica no es solamente su habilidad de permitir la reconstrucción mental del pasado, sino en su contribución al construir modelos mentales—simulaciones—de lo que podría pasar en seguida u otras perspectivas del ambiente inmediato, como lo que otros están pensando.\*

---

\* “Thinking about the future, episodic remembering, conceiving the perspective of others (theory of mind) and navigation engage this network, which suggests that they share similar reliance on internal modes of cognition and on brain systems that enable perception of alternative vantage points. Perhaps these abilities, traditionally considered as distinct, are best understood as part of a larger class of function that enables flexible forms of self-projection. By this view, self-projection relies closely on memory systems because past experience serves as the foundation on which alternative perspectives and conceived futures are built. [...] We suspect the adaptive value of episodic memory is not solely in its ability to afford mental reconstruction of the past but rather in its contribution to building mental models—simulations—of what might happen next or other perspectives on the immediate environment, such as what others are thinking.”

(*cf.* Suddendorf et al. 2009; Szpunar & McDermott 2009.) Esto, creo, apunta a que nuestras capacidades de imaginación de posibilidades tienen un claro valor adaptativo (Suddendorf *et al.* 2009; Williamson 2010). No es de sorprender, entonces, que al menos desde los 2 años, los niños puedan ejercer pensamiento contrafáctico (Byrne 2002), y la pregunta de si es una capacidad innata sigue abierta. Sin embargo, es controversial el si tal capacidad es poseída por animales no humanos (Suddendorf *et al.* 2009). Lo que sí sabemos es que es compartida a través de las culturas (Gilovich et al., 2003), y puede haber juicios contrafácticos—imaginativos—incluso en lenguajes sin marcadores específicos, como el chino (Au 1983)—lo que nos lleva a pensar que estamos ante una capacidad que no presupone la capacidad lingüística. Incluso hay quien ha propuesto que el lenguaje *presupone* la capacidad imaginativa: según la hipótesis de Suddendorf y sus colegas (2009), el lenguaje evolucionó para expresar los hechos guardados en la memoria episódica, y contruidos por la imaginación de hechos futuros posibles. Todo esto podría llevarnos a preguntar, creo, si tendremos una facultad diseñada para imaginar posibilidades: un “módulo para la modalidad” (*cf.* §4.3, abajo).<sup>75</sup>

Pero no solamente tenemos evidencia desde las ciencias cognitivas. También puede argumentarse que, sin cierta medida de conocimiento modal, no podríamos utilizar el lenguaje para comunicarnos—*al menos* si una importante teoría sobre la aseveración, la de Stalnaker (1978), va por buen camino. Para Stalnaker, una conversación se da con un *contexto* de trasfondo—donde tal contexto es la información que cada una de las participantes en la conversación están asumiendo. Esta información a su vez se modela como un conjunto de *mundos posibles*—aquellos que son *compatibles* con las aseveraciones dichas por una u otra participante y que han sido aceptadas para propósitos comunicativos. Cuando una de las participantes asevera una oración y las demás no la rechazan, el contenido semántico (la información transmitida) de esta se agrega al contexto, y se cuenta como presuposición—mientras que los mundos posibles donde ocurre algo incompatible con el contenido de esa aseveración dejan de presuponerse.

Si el acercamiento stalnakeriano a la comunicación va por buen camino, entonces los usuarios de un lenguaje hablado tenemos cierta medida no-trivial de conocimiento modal. Pues al menos podemos detectar fiablemente que tal o cual información transmitida por el hablante es *incompatible* con tales o cuales posibilidades.

Dado todo esto, parece suficientemente plausible asumir que el conocimiento modal está fuertemente imbricado en buena parte de, si no es que *todo*, nuestro conocimiento—aunque resultara que de lo que sabemos no dependiera directamente de nuestro conocimiento modal, aún así puede ser que dependa de otras cosas que sabemos y que, a su vez, sí lo hacen. Así que si no pudiésemos adquirir conocimiento modal, no podríamos tener una enorme parte del conocimiento que ahora tenemos—quizá, incluso, no podríamos tener nada de él.

Quien desee asumir lo contrario tiene la siguiente carga de la prueba: debe, *primero*, darnos alguna pista de cómo podríamos *destejer* el conocimiento modal de nuestro conocimiento no-modal, en el sentido de darnos al menos algo cercano a reglas que nos permitieran eliminar co-

---

<sup>75</sup> Contrario a lo que Williamson (2007, p. 142) dice (“As for the psychological study of the processes underlying our assessment of counterfactual conditionals, it remains in a surprisingly undeveloped state ...”), esto debería indicar que el estudio del pensamiento contrafáctico de hecho tiene bastante datos que aportar.

nocimiento que dependiera esencialmente de conocimiento modal sin con ello eliminar conocimiento que no haga; *segundo*, debe mostrarnos que, aún si pudiésemos llevar esa primera tarea acabo, todavía podríamos encontrarnos en una situación epistémica saludable—o, al menos, no mucho peor que la que ahora tenemos.

Dado todo esto, el siguiente condicional está favorecido por la evidencia disponible:

(COND1) Si no tuviésemos una cantidad considerable de conocimiento *modal*, no tendríamos una cantidad considerable de conocimiento *no-modal*.

Este condicional se traduce fácilmente en el siguiente:

(COND2) Si el escepticismo *modal* global es el caso, entonces muy probablemente el *escepticismo global* es el caso.

Ahora bien, en esta tesis voy a asumir que poseemos una cantidad considerable de conocimiento.<sup>76</sup> Es decir, asumiré:

(¬GLOBAL) El escepticismo global no es el caso.

Haciendo una simple inferencia desde COND2 y ¬GLOBAL, obtenemos:

(¬MODALGLOBAL) El escepticismo modal *global* es muy probablemente falso.

Así que, si queremos evitar ya sea el escepticismo global o la penosa tarea arriba aludida (o *ambos*), es mejor que evitemos el *escepticismo modal global*: es mejor que aceptemos que *poseemos una cantidad sustantiva de conocimiento modal*, es decir, ¬MODALGLOBAL.

Con lo que la tarea de brindar una epistemología modal *aceptable* debe evitar, de principio, un escepticismo global. Esto, por supuesto, deja abierta la posibilidad de que tengamos un escepticismo modal *local*: por todo lo que hasta ahora sabemos, cierto sector de nuestro conocimiento modal podría ser desechable debido a razones escépticas. Cuáles sean esas razones y qué sector podrían ser ese: he aquí el tema de este capítulo.

*¿Qué tanto podemos saber?* Ese “podemos” importa tres cuestiones muy relacionadas: aquello que podamos saber será fijado, en primera, por cómo sea el mundo; en segunda, por el alcance de nuestras capacidades cognitivas y, en tercera, por cómo esas capacidades y el mundo interactúen bajo los estándares epistémico-normativos que dan lugar al conocimiento. Así que una manera de fijar los límites de nuestro conocimiento es fijando ya sea cómo es el mundo, cuál es el alcance de nuestras capacidades, o cómo el mundo y nuestras capacidades pueden relacionarse de tal manera que obedezcan los estándares que regulan nuestro conocimiento.

En la primera parte de esta tesis describí lo que me parece que es un acercamiento aceptable al aspecto modal del mundo—a su, como le llamo, *estructura modal*. Así que nos queda investigar qué capacidades cognitivas se ponen en juego en la obtención del conocimiento modal y cómo es que ellas logran relacionarse, en concordancia con los principios de la normatividad epistémica, con la realidad modal, de tal manera que podamos tener conocimiento modal. Si alguna de las tres cosas—el mundo, nuestra mente, la relación epistémica entre ambos—falla, el conocimiento que creíamos poseer sobre esa base no existirá.

En buena parte, los escepticismos que revisaremos aquí son motivados por la duda de que po-

---

<sup>76</sup> Una versión anterior de esta tesis incluía una discusión sobre el escepticismo del mundo externo, pero los detalles de tal discusión la hacían excesivamente larga y algo fuera de lugar respecto al tema principal de esta tesis. Por ello he decidido dejar al anti-escepticismo en el nivel de una asunción no argumentada aquí.

seamos las capacidades cognitivas que requiere cierto sector de conocimiento modal que creemos poseer. Es decir: parten del hecho de que nuestras capacidades cognitivas no tienen el alcance o potencial suficiente como para lograr que tengamos un sector importante del conocimiento modal que hemos venido asumiendo que tenemos. Esto presupone al menos tres tareas:

- (ESCEPTICISMO1) Separar un reino de hechos modales (donde el caso límite, prohibido por  $\neg$ MODALGLOBAL, es tomar a *todos* los hechos modales como tal reino);
- (ESCEPTICISMO2) Especificar cuáles son los requisitos epistémicos (*i.e.*, tanto la posesión de las capacidades cognitivas y la satisfacción de los estándares epistémicos) necesarios para conocer tal reino;
- (ESCEPTICISMO3) Argüir que los agentes cognitivos como nosotros no cumplimos con tales requisitos.

La tarea que el escéptico modal tiene en ESCEPTICISMO1 es central justo por esa razón: a riesgo de llegar al escepticismo global, el escéptico modal tiene que aislar una clase “parasitaria” de pseudo-conocimiento modal que no sea esencial para nuestra vida cognitiva. Por esta razón, escepticismos modales radicales que nieguen de entrada nuestra posesión de cualquier conocimiento modal—o incluso la mera *significatividad* del discurso sobre hechos modales, como la postura de Quine (1953a,b), serán descartados de inicio en este trabajo.<sup>77</sup>

Supongamos que el escéptico modal logra llevar a cabo las tareas ESCEPTICISMO1-3. Parece adecuado requerirle todavía que nos diga por qué creíamos poseer tal conocimiento—qué era ese factor ilusorio que nos hacía creer que teníamos cierto acceso a cómo es el mundo, acceso que en realidad no poseemos. Esto sugiere que un escepticismo modal no sólo debería complacerse en llevar a cabo las tres tareas arriba mencionadas, sino que también debería:

- (ESCEPTICISMO4) Dar una explicación de por qué, en primer lugar, creíamos poseer conocimiento modal que de hecho nunca poseímos, pero con el cual contábamos en varios sectores de nuestra vida cognitiva.

Ahora bien, *no* quiero afirmar que un escéptico modal quede refutado si no cumple con esta tarea. Ciertamente, esta tarea es menos central que las otras tres. Pero sí que llevarla a cabo hace la postura del escéptico más plausible, pues no hace que nuestras firmemente arraigadas creencias de poseer conocimiento modal sean un misterio, o parezcan como completamente irracionales. Cierta tipo de *principio de caridad* puede invocarse aquí: *no atribuyas irracionalidad cuando no es necesario*. El escéptico que cumple con la tarea ESCEPTICISMO4 busca cumplir este principio, y en esa medida su posición gana fuerza dialéctica. Un escéptico que no se interese la tarea ESCEPTICISMO4 no necesariamente afirma que los metafísicos modales y otra gente que asume la existencia de conocimiento modal es estúpida, por supuesto; pero sí que nos deja con la duda de por qué gente *no estúpida* podría haber confiado *tanto* en una creencia—que poseemos conocimiento modal—que resultó ser falsa, y descubrible como falsa tras pura reflexión.

Como veremos, la opción preferida de los escepticismos que revisaremos (el de Nozick y el de van Inwagen) es negar que conozcamos un cierto reino de posibilidades que juegan un papel

---

<sup>77</sup> No quiero dejar de notar la deliciosa ironía histórica de que haya sido la ciencia cognitiva—uno de cuyos ancestros es la psicología conductual, favorita de Quine—la que haya brindado evidencia empírica de que el discurso modal *tiene* que tener sentido, pues nuestra cognición depende de él de maneras no triviales.

central en la investigación filosófica. Como veremos, los escepticismos de Nozick y van Inwagen parecen trabajar bajo la asunción de que negar el conocimiento de muchas posibilidades filosóficamente relevantes tiene pocas consecuencias en otros sectores de nuestra vida cognitiva. Hay que reconocer que la motivación para estos escepticismos no es para nada descabellada: como en muchos sectores de la filosofía podemos encontrar apelaciones a posibilidades tan (al menos *prima facie*) lejanas de nuestra realidad, uno muy fácilmente comienza a preguntarse si en realidad tenemos acceso cognitivo a esas supuestas posibilidades (veremos ejemplos en la siguiente sección).

Sin embargo, no toda posibilidad filosóficamente importante es una posibilidad remota. Muchos experimentos mentales para motivar una teoría filosófica o un contraejemplo a una teoría filosófica apelan a posibilidades no sólo muy cercanas a nuestra realidad, sino, incluso, a veces reales: por ejemplo, es fácil encontrar casos reales que cumplan con la estructura relevante de las posibilidades imaginadas por Gettier (1963) o Goldman (1976). Así que necesitamos un criterio más exacto para llevar a cabo la tarea ESCEPTICISMO1.

Pero aún hecho esto, el escéptico modal todavía se enfrentará con las tareas, altamente no-triviales, ESCEPTICISMO2-4, que piden especificar cómo nuestras capacidades cognitivas fallan en satisfacer los estándares epistémicos necesarios para tener acceso cognitivo al reino modal y por qué creíamos poseer tal acceso.

Aunque terminaré rechazando las dos posturas escépticas aquí consideradas (que son, por todo lo que sé, las mejor articuladas), veremos que podemos rescatar varias de las intuiciones que parecen molestar al escéptico modal, sin con ello comprometernos con su postura. Veremos también que el escéptico tiene bastante razón en presionar ciertos problemas que delimitan lo que deberíamos contar como una solución epistemológica satisfactoria, pero que, de nuevo, tales problemas no son irresolubles. Hay esperanza para nuestro conocimiento de las esencias y la necesidad.

## 4.2. contra van Inwagen

Tengo que confesar que el “Modal Epistemology” de Peter van Inwagen (1998) me parece oscuro en varios puntos—y debido a esto, muchas veces tendré que citar donde con un texto más claro podría simplemente reconstruir. Aquí haré mi mejor esfuerzo interpretativo, e intentaré exponer cómo es que van Inwagen pretende llevar a cabo las tareas ESCEPTICISMO1-4.

Por lo que respecta a ESCEPTICISMO1, van Inwagen acepta que tenemos cierto conocimiento modal, que usamos en la vida cotidiana, en la ciencia e incluso en la filosofía, pero rechaza que tengamos *todo* el conocimiento modal que los filósofos asumen que tenemos (es decir, su postura es compatible con  $\neg$ MODALGLOBAL). Definiré a su versión de escepticismo modal como la conjunción de dos tesis (*cf.* 1998, p. 69):

- (a) Tenemos *cierta* cantidad de conocimiento modal, pero
- (b) Usando nuestras capacidades (y sin ayuda externa), no *podemos* poseer *todo* el cono-

cimiento modal que el metafísico afirma tener.<sup>78</sup>

Así como está formulado, la postura de van Inwagen todavía no lleva a cabo la tarea ESCEPTICISMO1: separar un reino específico de hechos modales, al que después se nos niegue acceso epistémico. Necesitamos más detalle. ¿Con qué regla efectuamos una *partición* en el reino de las proposiciones modales, en la teoría de van Inwagen, que nos de dos sub-reinos: aquel de las proposiciones que podemos conocer y el de las que no podemos poseer?<sup>79</sup>

Una primera pista son los ejemplos con los que van Inwagen abre su artículo, los argumentos que él llama "*possibility arguments*" o *argumentos de posibilidad* (67-8):

(A1) Es posible que haya un ser perfecto (un ser que tenga esencialmente todas las perfecciones).

(A2) La existencia necesaria es una perfección. Por lo tanto,

(AC) Hay un ser perfecto.<sup>80</sup>

(B1) Es posible que yo exista y no exista nada material.

(B2) Cualquier cosa material es esencialmente material. Por lo tanto,

(BC) Yo no soy una cosa material.

(C1) Es posible que existan vastas cantidades de sufrimiento para las cuales no hay explicación.

(C2) Si existe un ser omnisciente, omnipotente, y moralmente perfecto, no pueden existir también vastas cantidades de sufrimiento para las cuales no hay explicación. Por lo tanto,

(C3) Es imposible que haya un ser esencialmente omnisciente, omnipotente, y moralmente perfecto.

van Inwagen nota que las premisas más controvertidas en los anteriores argumentos son A1, B1, y C1: las premisas en las que se afirma la existencia de cierta posibilidad. Estas son, además, ejemplos del tipo de proposiciones que el escepticismo de van Inwagen niega que podamos conocer. Sin embargo, esto todavía no resuelve ESCEPTICISMO1: ¿qué es lo que A1, B1, y C1 tienen *en común*, que las hace incognoscibles? No puede ser simplemente que afirmen la existencia de posibilidades, o que sean posibilidades filosóficamente interesantes, pues las posibilidades afirmadas en los casos de Gettier y de Goldman son *filosóficamente interesantes* pero, presumiblemente, cognoscibles.

La manera en que van Inwagen busca responder a esta cuestión parece basarse en su *analogía con la distancia* (p. 246):

En mi postura, muchos de nuestros juicios modales son análogos a juicios de distancia

78 La restricción al uso de *nuestras* (i.e., de nuestra especie) capacidades cognitivas deja abierta la posibilidad de que llegáramos a conocer proposiciones que fueran inaccesibles para nuestras capacidades, mediante el testimonio de agentes con capacidades diferentes (van Inwagen 1998: 245, n. 4).

79 Aquí ignoro algunos detalles que dependerán de nuestra teoría de las proposiciones: por ejemplo, quizá podemos conocer una proposición modal bajo un "modo de presentación" particular pero no bajo otro.

80 Alessandro Torza me ha hecho notar que no es claro que este argumento sea válido. Aunque el argumento es puramente ilustrativo de parte de van Inwagen, hay que notar que el argumento requiere de asunciones extra para probar su conclusión. Sin embargo, no me detendré en este punto.



hechos a ojo. [...] Tales juicios no son, por supuesto, infalibles, pero en un amplio rango de circunstancias pueden ser muy exactos. Hay, de cualquier manera, circunstancias—circunstancias remotas respecto a las cuestiones prácticas de la vida diaria—en las que no son exactas en absoluto.\*

Esto parece sugerir que el criterio de van Inwagen para distinguir las proposiciones modales cognoscibles de las incognoscibles, a la luz de la analogía propuesta, es:

(CRITERIO DE LEJANÍA)

No estamos justificados en juzgar si una proposición  $\langle p \rangle$  es posiblemente verdadera (o si una cosa  $x$  es posiblemente existente, o si una propiedad  $P$  es posiblemente ejemplificada, o si un concepto  $c$  tiene posiblemente una extensión), si el mundo  $w$  en el que  $\langle p \rangle$  es verdadera (existe, etc.) está *demasiado lejos* (dada una métrica de distancia entre mundos posibles) del mundo real.

Por supuesto, CRITERIO DE LEJANÍA no tiene suficiente contenido hasta que uno haya especificado una métrica de distancia entre mundos posibles—pero la idea intuitiva es clara:  $w$  está “muy lejos” del mundo real (*actual*)  $a$  si *muchos* hechos son diferentes en  $w$ , relativo a  $a$ . El problema con la idea intuitiva es que “muy lejos” hereda la vaguedad de “muchos”, y no tenemos una regla clara y constante para precisificar tal idea (cf. Lewis 1979, 1986: §1.2). Aún así, queremos que la regla no nos dé un resultado “todo-o-nada”: por ejemplo, no nos serviría de mucho postular que todo mundo *imposible* está a distancia 1 de  $a$ , todo mundo *posible* está a distancia 0 de  $a$ , y eso es todo. Quizá van Inwagen, por ejemplo, desearía afirmar que los mundos llenos de egos cartesianos y sin ningún objeto material, deben contar como “lejanos” dada la métrica que debemos buscar; mientras que un mundo que difiera de  $a$  sólo en que la camisa que elegí vestir hoy es verde en lugar de azul, debe contar como “muy cercano” de  $a$ . Pero ¿qué pasa con mundos donde sólo hay un ego cartesiano (asumiendo que en  $a$  no hay uno sólo) y fuera de eso no se cambia a  $a$  en nada? El cambio es relativamente mínimo respecto a la lista de entidades que tal mundo contiene y que  $a$  no contiene; pero quizá el cambio se deba a cambios en las leyes fundamentales—lo cual, probablemente, debería contar como un cambio que sólo se da en los mundos “lejanos”.

Esto indica que el problema de hacer una teoría sistemática y general sobre la distancia entre los mundos posibles es altamente no-trivial. Lo cual, a su vez, hace que CRITERIO DE LEJANÍA tenga que quedarse sin un contenido exacto, esperando que haya más desarrollos en la metafísica modal.

Aún así, por mor de la discusión voy a asumir que hay una tesis cercana a CRITERIO DE LEJANÍA que puede defenderse. Ciertamente, es atractivo pensar que, entre más cotidiana sea una posibilidad, mejor seremos al juzgar sobre ella—en los casos extremos, somos relativamente buenos juzgando posibilidades que suceden todos los días en el mundo real, mientras que quizá las posibilidades que nunca han sucedido pero que no sean tan distintas—como, por ejemplo, el sol de

---

\* “In my view, many of our modal judgments are analogous to judgments of distance made by eye. [...] Such judgments are not, of course, infallible, but in a wide range of circumstances can be pretty accurate. There are, however, circumstances—circumstances remote from the practical business of everyday life—in which they are not accurate at all.”

nuestro sistema solar colapsando—nos cuesten un poco más de trabajo; pero ciertamente la idea misma que yo podría existir sin ningún objeto material en el universo, es bastante más difícil de aceptar. Así que CRITERIO DE LEJANÍA es una postura razonable—que no obvia.

Que CRITERIO DE LEJANÍA, o algo parecido a esa tesis, se está asumiendo en el texto de van Inwagen, recibe más evidencia cuando consideramos una de las interpretaciones: la de Geirsson (2005). En pp. 281-2, Geirsson expone uno de los tres argumentos que, según él, van Inwagen da: el *argumento analógico*:

- (G1) Los juicios visuales son fiables en las circunstancias que involucran las circunstancias prácticas de la vida diaria.
- (G2) Los juicios visuales no son fiables en las circunstancias muy lejanas a las circunstancias prácticas de la vida diaria.
- (G3) En su fiabilidad, los juicios modales son como juicios visuales. Así,
- (G4) Los juicios modales son fiables en las circunstancias que involucran las circunstancias práctica de la vida diaria, y no son fiables en las circunstancias muy lejanas a las circunstancias prácticas de la vida diaria.\*

La respuesta de Geirsson a este argumento, en breve, es que ignora que las capacidades que nos permiten ser más fiables en un dominio de juicios son tales que muchas veces pueden entrenarse. Por ejemplo, los juicios matemáticos de una persona no-matemática serán poco fiables cuando se alejen de la aritmética y quizás la geometría que se utiliza en la vida diaria; pero eso es obviamente distinto cuando nos enfrentamos con un medalla Fields, por ejemplo. Así, de acuerdo a Geirsson (p. 282-3), plausiblemente, las personas entrenadas en la filosofía durante años, pueden llegar a mejorar sus capacidades y así convertirse en evaluadores fiables de posibilidades: tanto como un medalla Fields es mucho más fiable haciendo juicios matemáticos alejados de la vida ordinaria que otras personas, quizá un premio Schock de filosofía sea más fiable que otras personas al hacer juicios filosóficos alejados de la vida ordinaria.

Creo que la respuesta de Geirsson no es suficiente para bloquear el argumento. Siguiendo con la analogía, hay preguntas abiertas en matemáticas que ningún medalla Fields, que sepamos, ha podido responder. Así, quizá hay posibilidades tan remotas—asumiendo que de hecho sean *posibilidades*—que ningún filósofo(a) puede considerar justificadamente. Ciertamente, el talento se puede entrenar y con ello aumentar la fiabilidad—pero todavía no sabemos cuánto entrenamiento sería necesario para *saber* si podría o no haber egos cartesianos. Esto me lleva a pensar que el siguiente es uno de los problemas *para toda epistemología modal* (como CRITERIO DE LEJANÍA, asume la existencia de una métrica entre mundos apropiada):

(PRIMER PROBLEMA)

¿Qué tan *lejos* puede estar un mundo posible, si es que hemos de tener creencias justificadas acerca de él?

---

\* “1. Visual judgments are reliable in circumstances that involve the practical circumstances of everyday life.  
 2. Visual judgments are not reliable in circumstances that are far removed from the practical circumstances of everyday life.  
 3. In their reliability, modal judgments are like visual judgments. So,  
 4. Modal judgments are reliable in circumstances that involve the practical circumstances of everyday life, and they are not reliable in circumstances that are far removed from the circumstances of everyday life.”

Así, en resumen (y por lo que puedo entender del artículo de van Inwagen), una manera de entender su respuesta a ESCEPTICISMO1 (la tarea de seleccionar un reino de hechos modales tal que no podamos tener acceso epistémico a ellos) se basa en una tesis al menos muy similar a CRITERIO DE LEJANÍA: aquellos hechos modales que están *bastante lejanos* de la realidad, *a*, dada una métrica de distancia entre mundos, son los hechos modales que no podemos conocer. Como dije arriba, ciertamente esta tesis tiene plausibilidad *prima facie*. Pero, como dije arriba, todavía necesitamos especificar (1) una métrica entre mundos posibles que sea adecuada, y (2) resolver la vaguedad en la noción de “lejanía suficiente” que se utiliza cuando se dice que las posibilidades incognoscibles son aquellas que están suficientemente lejanas.

Esto todavía deja sin solución a las tareas ESCEPTICISMO2-4: ¿qué requisitos epistémicos, en general, gobiernan nuestra vida cognitiva y son tales que, bajo ellos, nuestras capacidades cognitivas nos permiten conocer posibilidades cercanas pero no posibilidades lejanas? Y ¿por qué creíamos poseer capacidades que también nos permitieran conocer posibilidades lejanas?

El requisito epistémico que van Inwagen parece estar asumiendo es algo parecido a esto:

(CRITERIO DE FIABILIDAD)

*S* está justificado en creer que *p* sólo si la creencia de *S* de que *p* es fiable, donde la noción de *fiabilidad* en cuestión es la noción del fiabilismo de proceso (Goldman 1979). Extraigo este criterio de la cita de arriba, donde van Inwagen hace la analogía ente los juicios de posibilidad y los juicios de distancia: parece presuponerse que, relativo a cierta área, nuestros juicios sobre la distancia entre objetos son *fiabiles*, pero dejan de serlo cuando esa distancia es “suficientemente grande”; de manera análoga, nuestros juicios sobre posibilidades serían fiables, cuando esas posibilidades son “suficientemente cercanas”, pero dejan de serlo, cuando la distancia entre ellas y la realidad es “suficientemente grande”. Esto también haría sentido de la respuesta de Geirsson, cuya tesis es que nuestras capacidades para hacer juicios sobre posibilidades pueden entrenarse y, con ello, ser más fiables.

Me parece que CRITERIO DE FIABILIDAD, por sí mismo (fuera de la interpretación de van Inwagen), es una hipótesis plausible. Ciertamente, hay una importante tradición en la epistemología contemporánea que asume que, si se muestra que cierta clase de nuestros juicios no son fiables, eso es evidencia al menos *prima facie* de que no estamos justificados para creerlos, y por ello no nos brindan conocimiento. Ahora, también es claro que la teoría fiabilista no es universalmente aceptada, incluso cuando se la toma como enunciando solamente que la fiabilidad es *necesaria* para la justificación (o el conocimiento).

Hablaré más sobre el fiabilismo en el siguiente capítulo. Por el momento, quiero subrayar que, aunque el fiabilismo no es universalmente aceptado, sí es una de las más importantes teorías de la justificación—lo cual hace natural el pensar que un criterio fiabilista también debería aplicarse al caso del conocimiento modal. Ciertamente, los casos caso del conocimiento no-modal y del modal *no* parecen tener una diferencia relevante que impida tal generalización.

Ahora bien, ¿por qué cree van Inwagen que nuestras creencias sobre posibilidades lejanas no son suficientemente fiables? Es decir, ¿cuál es su respuesta para la tarea ESCEPTICISMO3?

Escribe van Inwagen (pp. 73-4):

Aunque no niego que tengamos algún conocimiento modal, tomo a mucho de este co-

nocimiento como misterioso. Algunos enunciados modales, he dicho, los conocemos mediante razonamiento desde lo que he llamado conocimiento modal “básico”—enunciados modales simples y obvios cuya verdad estamos de alguna en posición de conocer —, junto con algunos hechos acerca de cómo se construye el mundo. Pero ¿cómo comenzamos en este razonamiento? ¿Cómo sabemos que los enunciados modales “simples y obvios” son verdaderos? ¿Cuál es el fundamento del conocimiento modal básico? No sé cómo responder a estas preguntas.

No quiero implicar que todo nuestro conocimiento modal es o conocimiento modal “básico” u obtenido mediante deducción lógica o matemática de conocimiento modal “básico” y “hechos acerca de cómo está hecho el mundo”. [...]

Las preguntas que no sé cómo responder—hay dos—corresponden a proposiciones cuyos valores de verdad no pueden ser descubiertas por la reflexión sobre la lógica y los significados de las palabras o por el razonamiento matemático. Primero, ¿cómo podemos saber (o encontrar) si una proposición es posible cuando no sabemos que es verdadera—es decir, cuando o sabemos que es falsa o no sabemos si es falsa o verdadera? Segundo, ¿cómo podemos saber (o encontrar) si una proposición que sabemos que es verdadera también es necesaria?\*

Aquí, van Inwagen parece estar afirmando que tenemos conocimiento de ciertos hechos modales “básicos”, “obvios y simples”—aquellos que, bajo la hipótesis interpretativa de que está presuponiendo una tesis como CRITERIO DE LEJANÍA, se dan en mundos posibles suficientemente cercanos. Estos los conoceríamos mediante mecanismos cognitivos todavía a explicar. Además, estas fuentes de conocimiento modal “básico” podrían ser suplementadas por nuestras habilidades de razonamiento (sin comprometernos con que *todo* nuestro conocimiento modal deba poder ser explicado en estos dos pasos). Finalmente, estaría la afirmación de que estas capacidades no son suficientemente fiables (bajo la hipótesis de que se presupone CRITERIO DE FIABILIDAD) cuando se trata de juicios de posibilidad como las tesis A1, B1 y C1 de los argumentos de posibilidad arriba mencionados.

Lo que no me queda claro es que van Inwagen nos haya dado alguna base razonable para convencernos de que nuestras capacidades cognitivas *no* dan para conocer los juicios de posibilidad de los tres argumentos de arriba. Es decir (bajo las dos hipótesis interpretativas), no me queda

---

\* “Although I do not doubt that we have some modal knowledge, I regard much of this knowledge as mysterious. Some modal statements, I have said, we know by reasoning from what I have called “basic” modal knowledge – simple, obvious modal statements whose truth we are somehow in a position to know –, together with some facts about how the world is constructed. But how do we get started in this reasoning? How do we know the “simple, obvious” modal statements to be true? What is the ground of “basic” modal knowledge? I do not know how to answer these questions.

I do not mean to imply that all our modal knowledge is either “basic” modal knowledge or obtained by logical or mathematical deduction from basic modal knowledge and “facts about how the world is put together.” [...]

The questions I do not know how to answer – there are two – pertain to propositions whose truth-values cannot be discovered by reflection on logic and the meanings of worlds or by mathematical reasoning. First, how can we know that (or find out whether) a proposition is possible when we do not know that it is true – that is when we either know that it is false or do not know whether it is true or false? Secondly, how can we know that (or find out whether) a proposition that we know to be true is also necessary?”

claro que van Inwagen nos haya dado un argumento para pensar que no poseemos mecanismos cognitivos suficientemente fiables para creer proposiciones que sucedan en mundos posibles lejanos (como, presumiblemente, mundos en donde hay sólo egos cartesianos y no objetos físicos). Hay dos críticas.

En primera, como Geirsson hace notar (2005: 284), no se ha argumentado que las capacidades que los filósofos utilizan al hacer juicios sobre posibilidades lejanas sean distintas de las que utilizamos al hacer juicios sobre posibilidades cercanas, sólo que mejor entrenadas. Utilizando de nuevo la analogía entre matemáticas y filósofos, podríamos pensar que, así como una matemática utiliza las mismas capacidades que todo otro ser humano tiene para hacer geometría o aritmética, pero más desarrolladas gracias a (por ejemplo) un talento innato y un entrenamiento suficiente, quizá una filósofa utiliza las mismas capacidades que otras personas utilizan al imaginar posibilidades cercanas, sólo que pulidas por el ejercicio filosófico de considerar posibilidades remotas.

En segundo lugar, incluso suponiendo que utilizásemos capacidades diferentes, ¿por qué esto implicaría que las otras no son fiables? Supongamos que la ciencia cognitiva descubriera que los matemáticos profesionales utilizan áreas distintas del cerebro para hacer matemáticas que las que utiliza la gente con pocas habilidades matemáticas. ¿Mostraría esto que hay que dejar de creer en los resultados de la comunidad de matemáticos profesionales? Difícilmente. De hecho, creo que mostraría que esas *otras* capacidades son incluso más fiables (relativo al campo de las matemáticas) que las comunes. Por analogía, no es claro que, si descubriésemos que al considerar posibilidades remotas utilizamos capacidades diferentes de las que usamos al considerar posibilidades cercanas, eso mostraría que esas *otras* capacidades—las puestas en marcha por el ejercicio filosófico—no son fiables.

Lo que esto sí parece dejar claro es que hay otro problema que toda epistemología modal completa debería resolver:

(SEGUNDO PROBLEMA)

¿Qué habilidades cognitivas están involucradas en el logro de conocimiento modal? Los juicios filosóficos acerca de posibilidades remotas, ¿usan las mismas habilidades que los juicios sobre habilidades cercanas?

Finalmente, sobre la tarea ESCEPTICISMO<sup>4</sup>—la de responder por qué creíamos tener justificación para ciertas creencias modales para las cuales, de hecho, no tenemos justificación—van Inwagen afirma lo siguiente:

Muchos filósofos están muy confiados en que pueden hacer juicios modales filosóficamente interesantes acerca de preocupaciones muy lejanas de la vida cotidiana, y son impacientes con cualquiera que rete su afirmación de que tienen esta habilidad. [...]

[Un filósofo] no es consciente de que las creencias modales que expresa o presupone cuando él dice: “tendríamos más espacio si hubiéramos movido la mesa contra la pared” (es decir, que era posible que la mesa estuviera contra la pared) y las creencias modales que confiadamente expresa en sus escritos sobre teología filosófica tienen fuentes muy diferentes. [...] Las primeras tienen su fuente en nuestros poderes ordinarios de “modalización” [...]; las segundas tienen su fuente en su socialización profesional, en

lo que “sus compañeros le dejarán decir”.\*

La idea aquí es muy clara: los filósofos creen tener justificación en hacer juicios modales sobre posibilidades remotas porque así es como se ha asumido en su ambiente profesional. ¿Qué tan plausible es esto?

No creo que la idea sea descabellada. Siglos de historia de la filosofía nos muestran que han habido asunciones que durante cierta época se tomaban como obvias, o al menos poco discutibles—y que han resultado ser o falsas o al menos mucho menos obvias de lo que se creía antes. Los ejemplos abundan—“Aristóteles está básicamente en lo correcto acerca de casi todo”; “La distinción entre cosas extensas y cosas pensantes es exclusiva y exhaustiva”; “Kant está básicamente en lo correcto acerca de casi todo”—y un extenso “etcétera”. Dado esto, no es descabellado pensar que una de las asunciones de la filosofía analítica contemporánea—*que podemos conocer posibilidades remotas*—resultara ser mucho menos obvia de lo que se creía—o falsa.

Un problema con este primer punto es que básicamente lo mismo puede defenderse—o se ha defendido—sobre casi cualquier otra tesis (filosófica o no). Pero eso, en y por sí mismo, no implica nada sobre el estatus justificatorio de la tesis en cuestión.

La cuestión de en qué condiciones la existencia de una postura (argumentada, por supuesto) contra una tesis *T* disminuye el grado justificatorio de nuestra creencia en *T* es bastante compleja. Depende, por supuesto, del poder de los argumentos contra *T*, y también de la evidencia disponible a favor de *T*—y *no* quiero comprometerme con la idea de que *nunca* dependa de cuestiones extra-teóricas, como la cuestión de cuál “paradigma” es el dominante, y cuáles figuras intelectuales son importantes en el tiempo relevante.

Pero no necesitamos resolver esta complejidad aquí. Pues, por supuesto, la tarea ESCEPTICISMO4—la de decirnos por qué creíamos poseer justificación para creencias modales que de hecho no están justificadas—depende de que tal justificación *de hecho no exista*. Si resultase que de hecho podemos tener justificación para posibilidades lejanas—como las afirmadas en los argumentos modales—, entonces tal explicación cae por la borda de la irrelevancia. Y, como hemos visto, los argumentos ofrecidos por van Inwagen hasta el momento son insuficientes. Eso no muestra, claro está, que su escepticismo modal sea falso—sólo que las razones dadas para creerlo son insuficientes, y por ello que todavía hay lugar para argumentar que lo es.

Cerraré esta sección discutiendo un tercer problema para la epistemología modal traído a la luz por van Inwagen.

Una manera en que se podría pensar que logramos obtener conocimiento modal, y que es examinada brevemente por van Inwagen, utiliza ideas de Stephen Yablo (1993). De acuerdo a Yablo, yo puedo juzgar justificadamente (*prima facie*) que *p* es posible si, hablando en términos gene-

---

\* “Many philosophers [...] are very confident that they can make philosophically interesting modal judgments about concerns remote from everyday life, and they are impatient with anyone who challenges their claim to have this ability. [...]

[A philosopher] is unaware that the modal beliefs he expresses or presupposes when he says, “We’d have had more room if we’d moved the table up against the wall,” (e.g., that it was possible for the table to be up against the wall) and the modal beliefs he gives such confident expression to in his writings on philosophical theology have quite different sources. [...] The former have their source in our ordinary human powers of “modalization” [...]; the latter have their source in his professional socialization, in “what his peers will let him get away with saying.””

rales, puedo imaginarme un mundo posible en el que  $p$  es verdadera. Así que la habilidad involucrada en el conocimiento modal es la habilidad de concebir, que es la habilidad de imaginar mundos verificadores. Escribe van Inwagen (p. 78):

Aunque en cierto sentido es trivial que el aseverar la posibilidad de  $p$  es comprometerse con la posibilidad de una realidad coherente y completa de la cual  $p$  es una parte integral, la examinación de los intentos de los filósofos para justificar sus convicciones modales muestran que esta trivialidad es raramente, si alguna vez en absoluto, un factor operativo en estos intentos. Una o un filósofo dirá confiadamente que una vaca (naturalmente) púrpura es posible, pero él o ella no le habrá dedicado ningún pensamiento a la pregunta de si hay un pigmento púrpura químicamente posible tal que la codificación para las estructuras que serían responsables para su producción y su correcta ubicación en la piel de una vaca podría ser insertado coherentemente en cualquier ADN que realmente fuera ADN de vaca—o incluso ADN de “cosa-parecida-a-la-vaca-excepto-por-el-color”. [...] Pero si una filósofa no ha intentado hacer algo como esto, entonces esa filósofa no ha, en ningún sentido útil, intentado imaginar un mundo posible en el que hay vacas naturalmente púrpuras. Por lo tanto, si la tesis general de Yablo [que imaginar un mundo posible en el que  $p$  es verdadera brinda justificación *prima facie* para creer que  $p$  es posible] es correcta, y si yo estoy en lo correcto en mi aseveración de que en el estado presente del conocimiento nadie es capaz de imaginar un mundo posible en el que haya vacas naturalmente púrpuras, se sigue que (si no hay otra fuente de creencias modales *prima facie* justificadas que la que Yablo ofrece [...]), nadie está ni siquiera *prima facie* justificado en creer que las vacas naturalmente púrpuras son posibles.\*

Poco después, parece tener la misma presuposición en la p. 79:

¿Podemos imaginar un mundo en donde haya hierro transparente? No a menos que nuestras imaginaciones sucedan a un nivel de detalle estructural comparable al de las imaginaciones de los físicos de la materia condensada que intenta explicar, digamos, el fenómeno de la superconductividad.\*

La presuposición en cuestión parece ser que podemos imaginar una posibilidad en donde  $p$  sólo

\* Although it is in a sense trivial that to assert the possibility of  $p$  is to commit oneself to the possibility of a whole, coherent reality of which the truth of  $p$  is an integral part, examination of the attempts of philosophers to justify their modal convictions shows that this triviality is rarely if ever an operative factor in these attempts. A philosopher will confidently say that a (naturally) purple cow is possible, but he or she will not in fact have devoted any thought to the question whether there is a chemically possible purple pigment such that the coding for the structures that would be responsible for its production and its proper placement in a cow's coat could be coherently inserted into any DNA that was really cow DNA—or even “cow-like-thing-but-for-color” DNA. [...] But if a philosopher has not attempted to do something like this, then that philosopher has not, in any useful sense, attempted to imagine a possible world in which there are naturally purple cows. Therefore, if Yablo's general thesis is right, and if I am right in my assertion that in the present state of knowledge no one is able to imagine a possible world in which there are naturally purple cows, it follows that [...] no one is even *prima facie* justified in believing that naturally purple cows are possible.”

\* “Can we imagine a world in which there is transparent iron? Not unless our imaginings take place at a level of structural detail comparable to that of the imaginings of condensed-matter physicists who are trying to explain, say, the phenomenon of superconductivity.”

si imaginamos tal posibilidad con “un nivel suficiente de detalle estructural”. Pero ¿qué es que haya hechos que cuenten como descritos con “un nivel suficiente de detalle estructural”?

Creo que es plausible que la idea de van Inwagen sea la siguiente. Algunos hechos se *fundamentan* en otros—o, de manera más comprometedora, se *reducen* a otros—, en el sentido en que algunos hechos *explican la existencia y características intrínsecas* de otros hechos (cf. §2.1, arriba). Que una especie tenga tal ADN (y, quizá, tal historia evolutiva) *explica* que sea una vaca y que tenga tal tono de piel o de pelo; que haya tal o cual configuración microfísica *explica* que haya hierro transparente. La idea es entonces que no podemos imaginar una posibilidad *p* a menos que imaginemos sus *fundamentos*, en este sentido metafísico, a un nivel “suficiente” de detalle:

(CRITERIO DE DETALLE)

*S* está *prima facie* justificado en juzgar si *p* es posible sólo si *S* imagina un mundo donde los fundamentos de *p* son el caso, e imagina tales fundamentos con el suficiente detalle.

Un problema inmediato con CRITERIO DE DETALLE es que “suficiente detalle” es una expresión vaga; dejemos ese problema de lado por el momento. Un segundo problema es que varios metafísicos asumen que hay un *nivel metafísico fundamental*—aquel que consiste de todos aquellos hechos que fundamentan a todos los demás, pero que *no son fundamentados por nada* (e.g. Schaffer 2009; cf. §§3.1-3.2, arriba). Este segundo problema se evade fácilmente: CRITERIO DE DETALLE\* aplicaría CRITERIO DE DETALLE para los *hechos derivados*, mientras que diría que, para *hechos fundamentales*, la condición necesaria en cuestión es simplemente la imaginación de esos hechos con el grado relevante de detalle.

Un problema más interesante es que CRITERIO DE DETALLE exige demasiado.<sup>81</sup> Esto se ve de dos maneras distintas.

*Primera manera:* Incluso para las posibilidades que el escepticismo de van Inwagen (que asume ¬MODALGLOBAL) está comprometido a aceptar que conocemos, CRITERIO DE DETALLE exige mucho más de lo que, pre-teóricamente, es necesario (cf. Geirsson 2005: 285-8). Tomemos el caso de mi conocimiento de que Bonnie podría no haberse accidentado. Pre-teóricamente, sólo necesito imaginar que cierro la puerta a tiempo, en lugar de dejarla salir; entonces el auto que de hecho la atropelló pasaría de largo sin mayor consecuencia para mi perrita. Pero CRITERIO DE DETALLE exige que me imagine los hechos que fundamentarían esos hechos que me he imaginado: ciertos hechos sobre la trayectoria de Bonnie, mis intenciones al cerrar la puerta y el éxito al cerrarla, la dinámica del auto que de hecho la atropelló, y la interacción entre estas tres cosas—al menos.

Uno podría decir que, al imaginarme que Bonnie no es atropellada de la manera en que de hecho lo fue, *de hecho* ya estoy satisfaciendo, aunque implícitamente, CRITERIO DE DETALLE: mi imaginación en cuestión ya imagina ciertos hechos sobre trayectorias, intenciones y dinámicas—sólo que *yo* no conceptualizo tales hechos en tales términos—simplemente *me imagino* tales hechos. El problema es que, si esto es plausible, lo es también para el hierro transparente y las vacas con pelaje púrpura: resultaría que, al imaginarme vacas púrpuras, asumiendo que me imagino mundos biológicamente posibles, *ya* me imagino, automáticamente, hechos sobre cierto ADN (y aná-

81 O sus precisificaciones, o CRITERIO DE DETALLE\* y sus precisificaciones; en adelante me concentraré en la tesis más simple, dado que plausiblemente este problema se extiende a las demás.



logamente para el hierro).<sup>82</sup>

Una respuesta que elimina estas dificultades (sugerida por Hawke 2011) toma a CRITERIO DE DETALLE como restringido a la proposiciones modales que *no* son las que van Inwagen llamó “básicas”—es decir, aquellas que están suficientemente *lejos* de nuestro mundo real. Es con las proposiciones lejanas con las que se requiere que imaginemos con detalle a sus fundamentos, si es que hemos de tener conocimiento de ellas. Para las básicas, poseemos justificación *prima facie*.<sup>83</sup>

El problema es que esto entra en conflicto con CRITERIO DE LEJANÍA, que afirma que de hecho nuestras creencias sobre *ese* tipo de proposiciones *no tienen* justificación. Deberíamos, entonces, suavizar CRITERIO DE LEJANÍA, y afirmar que de ellas no tenemos justificación “por *default*”—más movimientos teóricos que simplemente el imaginarlas se requieren; en particular, como dice CRITERIO DE DETALLE, se requieren imaginar sus fundamentos. Pero esto no cierra, como se pretendía con el original CRITERIO DE LEJANÍA, la posibilidad en principio de que lleguemos a tener justificación para las proposiciones modales “lejanas”, como las de los tres argumentos de posibilidad.

Al menos, por supuesto, no lo cierra en la medida en que sea posible imaginar los fundamentos de hechos suficientemente lejanos. Una hipótesis es que van Inwagen podría haber supuesto que sus dos criterios se dieran soporte mutuamente.<sup>84</sup> Una primera opción es que, cuando un hecho es muy “lejano”, no podamos tener creencias justificadas sobre él—*debido a que* no puedo imaginar sus fundamentos con suficiente detalle. O quizá, no puedo imaginar sus fundamentos con suficiente detalle—*debido a que* el hecho es muy “lejano”. Sin embargo, no es claro que pueda haber tal soporte mutuo. Podemos imaginar hechos “lejanos” con bastante detalle—como cuando los científicos se preguntan qué pasaría si las constantes en las leyes fundamentales cambiaran. Así, no siempre podemos apelar a la “lejanía” para explicar la falta de detalle. De manera converso, no siempre podemos apelar a la falta de detalle para explicar la “lejanía”, pues nada impide que haya hechos muy cercanos que no podemos imaginar con mucho detalle—porque simplemente estén fuera del alcance de nuestras capacidades cognitivas.<sup>85</sup>

---

82 He asumido que me imaginaba una circunstancia *biológicamente posible*. Quizá incluso podría relajar esta asunción: si una filósofa o un biólogo me presionan y me dicen que la existencia de tal ADN no es compatible con las leyes de la biología de nuestro mundo, podría aceptar su objeción y decir que simplemente me imagino una circunstancia posible *relativamente a otras asunciones más generales* (digamos, quizá sólo a las leyes de la física fundamental), incluso si esas asunciones no fijan los hechos biológicos posibles de la misma manera en que lo hacen las leyes de la biología. Pero lo mismo pasa con el caso de Bonnie: aunque parece que es físicamente posible, eso podría resultar no ser el caso—ciertamente, todavía no poseemos un inventario completo de las leyes de la física. Así que ese inventario final, que no poseemos, podría resultar decirnos que la posibilidad de Bonnie no siendo atropellada es físicamente imposible.

Una pregunta profunda que no puedo tocar aquí es si contaríamos como *vacas* a aquellas cosas de pelaje púrpura que no obedecieran las leyes de la biología: uno podría pensar que los objetos son “formados” o “estructurados” por las leyes—de tal manera que nada es una vaca si no obedece las leyes de la biología y de la zoología (o leyes *suficientemente parecidas* a estas). Esto sería un caso análogo a la “schmasa” que Fine considera en (Fine 2002).

83 Esta justificación a su vez es analizada por Hawke en términos de dos principios: que lo real es posible y que dos cosas parecidas tendrán posibilidades parecidas (p. 360).

84 Debo esta posibilidad a Lourdes Valdivia, Anaid Ochoa y Melahuac Hernández, quienes me la sugirieron en discusión.

85 Hawke de hecho ofrece un “argumento inductivo” (pp. 363-4) para la afirmación de que de hecho nadie ha ima-

De cualquier manera, hay una *segunda manera* de mostrar que CRITERIO DE DETALLE es muy fuerte: sea  $p$  un hecho que me imagino y que por ello afirmo que es posible (podemos restringirnos a hechos “no-básicos”, o “remotos”). Entonces CRITERIO DE DETALLE me exige imaginarme los fundamentos de  $p$  con “suficiente” detalle. Sean tales fundamentos,  $q$ . Por una nueva aplicación de CRITERIO DE DETALLE, debo imaginarme los fundamentos de  $q$  con “suficiente” detalle. Sean  $r$  esos fundamentos—CRITERIO DE DETALLE me exigirá, de nuevo, imaginar *sus* fundamentos, y así hasta llegar a los hechos fundamentales del mundo que yo afirmo estar imaginando. Pero parece implausible requerir que, para imaginar cualquier posibilidad “remota”, tenga que imaginar todas las leyes fundamentales de un mundo posible, los hechos que obedecen esas leyes fundamentales, y así “hasta arriba”. Quizá no necesito imaginar tales cosas para los casos (si los hay) en que una posibilidad remota se fundamenta, en algún nivel, sobre un hecho “no-remoto”, para el cual van Inwagen (bajo la interpretación de Hawke) pueda aceptar que tengo justificación “por *default*”. Entonces eso de todos modos haría que los hechos remotos cuyos fundamentos son todos remotos, necesitaran de esta imaginación “hasta abajo”—hasta los hechos fundamentales. No me queda claro que esto tenga plausibilidad.

Creo que esto pone considerable presión sobre CRITERIO DE DETALLE (y las modificaciones mencionadas). Esto no significa que no haya una idea plausible en la cercanía. Ciertamente ciertos hechos son más “relevantes” para afirmar que imaginamos ciertas posibilidades que otros—y ciertamente es plausible requerir que imaginemos tales hechos “relevantes” con el “suficiente” detalle para que podamos afirmar que *de hecho* estamos imaginando *ésa* posibilidad que queremos imaginar. En el caso del hierro transparente, el hecho del hombre dando su discurso no parece ser “relevante” en el sentido en cuestión; el hecho de que ciertas partículas físicas tengan tal configuración sí parece serlo. Pero esto *no* es requerir algo tan fuerte como CRITERIO DE DETALLE—la pregunta es, justamente, qué es:

(TERCER PROBLEMA)

Al imaginar, concebir, ... etc., cierto hecho  $p$ , y con ello creer que  $\langle p \rangle$  es posible ¿qué otros hechos, y con cuánto detalle, necesitamos imaginar, concebir, ... etc., si es que nuestra creencia de que  $\langle p \rangle$  es posible ha de estar justificada?

Concluyo: van Inwagen *no* nos ha dado razones suficientes para ser escépticos modales en su sentido. Esto no muestra, claro está, que nuestro supuesto conocimiento modal y de esencias no sea problemático en absoluto—ciertamente hay retos que una epistemología modal debe resolver. Lo que creo que sí muestra, es que tales retos, al menos en el estado presente de la dialéctica, no muestran que *no tengamos* conocimiento modal, incluso el requerido para la metafísica. Esa hipótesis todavía está abierta a discusión, esperando una epistemología que la vindique—o que muestre que sí debemos ser escépticos.

Para cerrar esta discusión, repito los tres problemas que he extraído:

---

ginado posibilidades remotas con el detalle requerido; su argumento analiza el caso de los zombis chalmersianos y los fantasmas cartesianos, intentando mostrar que ninguno se ha imaginado suficientemente bien. Sospecho que su argumentación tiene un hueco importante al ignorar el argumento del vacío explicativo, pero las consecuencias de esto caen fuera de mis intereses aquí.

(PRIMER PROBLEMA)

¿Qué tan *lejos* puede estar un mundo posible, si es que hemos de tener creencias justificadas acerca de él?

(SEGUNDO PROBLEMA)

¿Qué habilidades cognitivas están involucradas en el logro de conocimiento modal? Los juicios filosóficos acerca de posibilidades remotas, ¿usan las mismas habilidades que los juicios sobre habilidades cercanas?

(TERCER PROBLEMA)

Al imaginar, concebir, ... etc., cierto hecho  $p$ , y con ello creer que  $\langle p \rangle$  es posible ¿qué otros hechos, y con cuánto detalle, necesitamos imaginar, concebir, ... etc., si es que nuestra creencia de que  $\langle p \rangle$  es posible ha de estar justificada?

### 4.3. contra Nozick

Un escepticismo modal mucho más simple que el de van Inwagen es el de Robert Nozick, expuesto en el capítulo 3 de su libro *Invariances* (Nozick 2001). Su argumento central es también mucho más claro—pero, creo, también mucho más fácil de refutar. Me concentraré, primero, en ese argumento central (p. 122):

Tales debates [sobre qué es necesario] serían evitados si poseyéramos una facultad de la razón que pudiera evaluar directamente la posibilidad de enunciados generales y sus negaciones. Esta facultad regresaría un juicio acerca de si  $O$  es posible o no, si no- $O$  es posible, y por ello si  $O$  es o no necesaria. De cualquier manera, no parece que tengamos tal facultad, y es implausible que los procesos evolutivos la podrían infundir en nosotros. Dado que nuestros ancestros evolucionaron en el mundo real, no había presiones selectivas para premiar la precisión acerca de todos los mundos posibles, y no había desventaja en estar en lo correcto sólo acerca del mundo real. Sin duda, el mecanismo más eficiente en llevar a nuestros ancestros hacia la verdad acerca del mundo real [*actual*] podría abarcar correctamente también otras posibilidades, pero es muy poco plausible pensar que abarcaría *todas* las posibilidades, no importa qué tan diferentes del mundo real. Una facultad más o menos precisa para generar o imaginar posibilidades podría haber servido bien a nuestros ancestros al decidir sobre acciones, pero no habría habido selección de una facultad suficientemente poderosa para generar *todas* las posibilidades lógicas. Las facultades más limitadas no tienen ninguna desventaja en el mundo real. Por lo tanto deberíamos ser cuidadosos al concluir que una oración  $O$  es necesariamente verdadera, simplemente porque nosotros y otros hayan sido incapaces de generar un contraejemplo a ella. Si nuestras capacidades imaginativas no generan todas las posibilidades que hay, o incluso una muestra aleatoria o representativa de ellas—si somos capaces de imaginar sólo posibilidades de cierto tipo—entonces la verdad de  $O$

podría no extenderse a posibilidades que no podemos imaginar.\*

De la manera en que estoy interpretando este párrafo, busca resolver las tareas ESCEPTICISMO1-3: decirnos qué hechos modales son los que no conocemos, qué principios epistémicos gobernarían tal conocimiento, y por qué no satisfacemos tales principios.<sup>86</sup> Vamos por turnos.

El argumento puede reconstruirse así (*cf.* Vaidya 2007, secc. 4):

(NOZICK1) No hay ventaja adaptativa en estar en lo correcto en todos los mundos posibles.

(NOZICK2) Si no hay ventaja adaptativa en estar en lo correcto en todos los mundos posibles, entonces no tenemos una facultad (o módulo) para detectar la verdad en todos los mundos posibles; y dado que esa es la definición de la necesidad metafísica, no tenemos un módulo para detectar necesidad.

(NOZICK3) Si no tenemos una facultad para detectar la necesidad, entonces ninguna de nuestras creencias acerca de la necesidad están justificadas.

(NOZICKC) Así, no estamos justificados en creer que algo es metafísicamente necesario.

Así, las respuestas a las tareas arriba señaladas serían las siguientes:

ESCEPTICISMO1: No conocemos hechos necesarios.

ESCEPTICISMO2: Podemos conocer hechos de una clase *C* sólo si tenemos una facultad evolutivamente seleccionada para detectar hechos de la clase *C*.

ESCEPTICISMO3: No poseemos una facultad para detectar hechos necesarios.

Consideraré ahora el argumento, y lo rechazaré. Rechazadas las tesis, no es necesaria una explicación de porqué creíamos poseer conocimiento modal que no poseemos—pues no nos habrá sido dada una buena razón para creer que no lo poseemos.

Uno de los peores defectos del argumento de Nozick es que es muy fácil de emular para probar cosas claramente absurdas. Por ejemplo, Alejandro Vázquez del Mercado alguna vez compartió<sup>87</sup> este ejercicio de sofistería, en tono de broma:

(TORMENTAS1) No hay ventaja adaptativa en resolver el cubo de Rubik.

---

\* "Such debates would be avoided if we possessed a faculty of reason that could directly assess the possibility of general statements and of their denials. This faculty would deliver a judgment about whether *S* was possible or not, whether not-*S* was possible, and hence whether or not *S* was necessary. However, we do not appear to have such a faculty, and it is implausible that evolutionary processes would instill that within us. Since our ancestors evolved in the actual world, there were no selective pressures to reward accuracy about all possible worlds, and there was no handicap to being right just about the actual world. To be sure, the most efficient mechanism to lead our ancestors to truth about the actual world might also correctly encompass other possibilities, but it is highly implausible to think that it would encompass *all* possibilities, no matter how different from actuality. A roughly accurate faculty for generating or imagining possibilities might have served our ancestors well in deciding upon actions, but there would not have been selection for a faculty powerful enough to generate *all* logical possibilities. More limited faculties bear no disadvantage in the actual world. We therefore should be wary of concluding that a statement *S* is necessarily true, simply because we and others have been unable to generate a counterexample to it. If our imaginative capacities do not generate all the possibilities there are, or even a random or representative sample of these—if we are able to imagine only possibilities of a certain sort—then *S*'s truth may not extend to possibilities that we cannot imagine."

86 Además de esto, Nozick ofrece varias tesis suplementarias para explicar por qué creíamos poseer conocimiento de hechos necesarios, con ello resolviendo la tarea ESCEPTICISMO4. De esto ya no hablo abajo.

87 Ver: <http://tormentas.tumblr.com/post/8391394820/evolutionary-skeptical-argument-against-rubik-cubes> (accesado en Julio de 2013)

(TORMENTAS2) Si no hay ventaja adaptativa en poder resolver el cubo de Rubik, entonces no hay módulo o facultad para resolver el cubo de Rubik.

(TORMENTAS3) Si no hay un módulo o facultad para resolver el cubo de Rubik, entonces no podemos resolver el cubo de Rubik.

(TORMENTASC) Así, no podemos resolver el cubo de Rubik.

Por supuesto, TORMENTASC es falso. Una de las cargas de la prueba, entonces, de un nozickiano está en mostrar cuál es la desanalogía entre el argumento de Nozick y el argumento sobre el cubo de Rubik. Asumiendo que no haya tal desanalogía, tenemos una muy clara indicación de que hay algo mal detrás del argumento de Nozick. Aquí intentaré mostrar qué es lo que anda mal: aunque la inferencia es, por supuesto, válida, creo que *las tres* premisas del argumento de Nozick son cuestionables, en el sentido en que o son altamente implausibles o su verdad debe ser decidida *a posteriori*, no con el razonamiento apriorístico de Nozick. Voy en orden.

Sobre NOZICK1. Concedamos la plausibilidad de afirmar que *no toda* cuestión necesaria tiene valor adaptativo. Aún así, es fácil argumentar que esto no puede ser general.

Supongamos, como es plausible (podríamos usar aquí cualquier otra tesis que con soporte empírico) que la capacidad de detectar si hay depredadores cerca de uno tiene un valor adaptativo. Entonces esta facultad sería “auto-localizadora”, en el sentido en que siempre se restringe a una región  $r$  (el lugar en el que uno está y sus alrededores) en un cierto momento  $m$  (el momento en que uno detecta si hay o no depredadores en  $r$ ). Es plausible suponer que las regiones y momentos tienen “identidad transmundo”: una y la misma región (momento) *podría* haber tenido diferentes propiedades—una y la misma región existe en diferentes mundos posibles. Ahora, supongamos que en un momento  $m_i$  y en una región  $r_i$ , en la cual estoy, detecto que hay un depredador en la cercanía. Sucede que en todo mundo donde existen  $m_i$ ,  $r_i$  y yo, y donde yo estoy en  $m_i$  y  $r_i$  preguntándome si hay un depredador, había tal depredador (quizá la historia inmediatamente anterior lo determinaba, quizá fue mera “suerte modal”—de cualquier manera, ese conjunto de mundos es relativamente pequeño). Entonces he descubierto un hecho necesario (en todo mundo  $w$  es el caso que: si en  $w$  existen  $m_i$ ,  $r_i$  y yo, y donde yo estoy en  $m_i$ ,  $r_i$  preguntándome si hay un depredador, hay tal depredador), y hacerlo tiene valor evolutivo. De nuevo, el ejemplo específico es insustancial; todo lo que se necesita es (1) fijar un hecho  $h$  cuyo conocimiento tenga valor evolutivo para cierta especie; (2) argumentar que las características específicas de  $h$  se dan en todos los mundos posibles donde existe ese hecho (quizá sólo existen en un mundo), y (3) concluir que: en todo mundo  $w$ , si existe  $h$ , entonces conozco  $h$  y eso tiene valor evolutivo. Esto nos da hechos necesarios cuyo conocimiento tiene valor evolutivo.

Por poner otro ejemplo, sea  $h$  un hecho cuyo conocimiento tiene valor adaptativo. Sin embargo,  $h$  y  $h_{@}$  (de hecho,  $h$ ) son lógicamente equivalentes. Así que parecería que, si sabemos  $h$  sabemos  $h_{@}$ . Desafortunadamente para Nozick,  $h_{@}$  es un hecho necesario: como sabemos por pura lógica modal (cuando se le suplementa el operador “*actually*”, “@”, “de hecho”), si  $p$ , entonces de hecho  $p$ , y si de hecho  $p$ , entonces necesariamente de hecho  $p$ .<sup>88</sup> Ergo, tenemos muchos conoci-

88 La inferencia es formalmente válida. Sea “@” el operador “de hecho”. Su semántica se da en marcos  $\langle W, R, a \rangle$  donde  $a$  es un elemento distinguido de  $W$  que representa al mundo real. La cláusula de verdad para fórmulas  $\lceil @p \rceil$  es:

mientos de hechos necesarios con valor adaptativo.

De nuevo, este argumento sólo muestra que NOZICK1, *así puesta*, tiene contraejemplos. Pero muy probablemente ése no es el “espíritu” de la premisa—más bien, quizá la idea pretendida es que no hay valor adaptativo en *detectar la necesidad* de un hecho, como opuesto a detectar hechos *que resulten ser* (incluso por pura lógica modal) hechos *necesarios*.

Eso tampoco es absolutamente claro: hay casos *imaginables y plausibles* en los que detectar la necesidad de un hecho tiene valor adaptativo (doy ejemplos abajo). Además, Nozick parece sostener NOZICK1 sobre bases igual de especulativas, como vimos en la cita. Simplemente no apela a evidencia empírica que apoye su discusión. Creo que esta evidencia es necesaria, pues, como mostraré abajo, *sí* podemos imaginar casos en que tales facultades tengan valor adaptativo—por ello, dado que el debate de si tales facultades tienen tal valor está en peligro de convertirse en un debate acerca de qué casos podemos imaginar o no, necesitamos evidencia empírica que decida la cuestión.

He aquí un caso imaginable y plausible donde la “detección de necesidades” tendría valor adaptativo: el caso de los condicionales *modales*. Hay casos en los que saber que cierta relación condicional entre dos proposiciones es necesaria, tiene valor adaptativo. Por ejemplo: “si veo huellas hacia mi cueva, oigo ruidos en los matorrales, y percibo el olor de un depredador, debe ser porque hay un depredador en las cercanías”. O: “si el bebé está llorando y tiene tales y cuales gestos en la cara, debe ser porque algo le molesta”.

Ahora, por supuesto, tales condicionales podrían no siempre resultar en una implicación necesaria, sino en mera correlación en un grado muy alto. Pero entonces podríamos imaginar que la noción de correlación, siendo una noción probabilística, es más fina que la noción de implicación necesaria—pues presupone conceptos cuantitativos más finos que los conceptos cualitativos de la noción modal. Esto podría llevarnos a pensar que creer en necesidades presupone menos sofisticación cognitiva que creer en correlaciones, y que por ello un animal que hiciese lo primero sin lo segundo podría ganar el valor adaptativo sin los costos que significa la obtención de un aparato conceptual más fino. Además, la creencia en necesidades como las mencionadas no tendría, *ceteris paribus*, un mayor costo que la creencia en correlaciones—dado que en estos casos un falso positivo es mucho menos costoso que un falso negativo: es mucho más costoso creer que no hay un depredador, cuando sí lo hay, que creer que sí hay un depredador, cuando no lo hay. Por lo tanto, la detección de conexiones de implicación necesaria podría tener valor adaptativo, incluso si hay casos en los que tales conexiones no se dan, sino que sólo se dan correlaciones.

De nuevo, no quiero comprometerme con que estos casos decidan la cuestión. La única fuerza dialéctica que deseo darles es esta: dado que Nozick puede imaginar que la detección de necesidades no tiene valor adaptativo, y dado que yo puedo imaginar que sí lo tiene, necesitamos un

---


$$w \models @p \text{ sii } a \models p, \text{ para cualquier } w$$

Ahora, supongamos que  $p$  es verdadera, entonces es verdadera en  $a$  (verdad es verdad en este mundo, y eso es sólo verdad real, pues este mundo es el real). Pero entonces todo mundo ( $R$ -accesible a  $a$ ) hará verdad que  $@p$ . Por definición, entonces,  $\Box @p$ —que era lo que queríamos.

Nótese que  $\langle @p \rangle$  es una proposición distinta de  $\langle p \rangle$ , por lo que este argumento no implica que *toda* verdad sea necesaria—implica que toda verdad necesariamente ocurre en el mundo real. (Esta inferencia falla cuando “@” se interpreta indexicalmente, “refiriéndose” al mundo de evaluación.)

juez externo—evidencia empírica—que nos ayude a decidir la cuestión. Mientras tanto, la cuestión sigue abierta, y NOZICK1 no sólo queda con su verdad puesta en duda, sino sin poder dialéctico alguno.

Sobre NOZICK2. El problema principal de esta premisa es que asume que los rasgos de los seres vivos son *únicamente* fijados por presiones adaptativas. Esta premisa es altamente controversial en la biología y su filosofía (Gould & Lewontin 1979; Sober 1996; Stich 1990: §3.3). Por ejemplo, Sober propone que, para las poblaciones sujetas a la selección natural, los rasgos con mayor *fitness* tienden a convertirse en los más comunes—*siempre que no haya fuerzas exteriores que lo impidan*:

Quando la selección es la única fuerza que guía la evolución de una población, el fenotipo con más *fitness* disponible evoluciona. De cualquier manera, cuando otras fuerzas se entrometen, otros resultados son posibles.

La selección natural “pura” tiene resultados predecibles, pero el mundo nunca es puro. Por ejemplo, las poblaciones nunca son infinitamente grandes, lo que significa que la deriva aleatoria siempre juega algún papel, tan pequeño como pueda ser. De todos modos, permanece la pregunta de qué tan cercanamente la naturaleza se aproxima al caso puro. Es una cuestión empírica el si la selección natural fue la única influencia importante sobre la evolución de un rasgo particular en una población particular, o si fuerzas no-selectivas también jugaron un papel importante.\*

Además, como nota Fodor (2000: 191), “el éxito reproductivo es producido por una división biológica del trabajo entre los órganos [de un organismo]”,\*\* por lo que no todo órgano, bajo la postura evolucionista, necesita dedicarse *específicamente* al éxito reproductivo—por ejemplo, el corazón es *para* bombear sangre, y esta función tiene un valor para el organismo con corazón que contribuye derivativamente al valor adaptativo. Si esto es así, la premisa de Nozick no considera la posibilidad abierta de que tengamos un módulo de detección de necesidades *no* porque tal módulo nos de una ventaja adaptativa *directamente*, sino porque, al interactuar con otras capacidades, *resulte* una ventaja adaptativa. De hecho, al inicio de este capítulo vimos que se han propuesto ideas de este estilo en la ciencia cognitiva.

Así, es una cuestión empírica la de si no tenemos un módulo o facultad evolutivamente diseñado para detectar posibilidades. Concluyo que NOZICK2 es una premisa que debe decidirse *a posteriori*; y que, de hecho, parece tener evidencia empírica en contra. NOZICK2 queda también sin poder dialéctico.

Sobre NOZICK3. Esta premisa parece requerir que sólo las creencias directamente basadas en la información procesada por *un* módulo o *una* facultad son fiables. Pero eso no es claro. Para

---

\* “When selection is the only force guiding a population's evolution, the fittest of the available phenotype evolves. However, when other forces intrude, other outcomes are possible.

'Pure' natural selection has predictable results, but the world is never pure. For example, populations are never infinitely large, which means that random drift always plays some role, however small. Still, the question remains of how closely nature approximates the pure case. It is an empirical matter whether natural selection was the only important influence on the evolution of a particular trait in a particular population, or if non-selective forces also played an important role.”

\*\* “[...R]eproductive success is produced by a biological division of labor between their organs”

mostrarlo, asumiré que las “facultades” en cuestión deben ser entendidas como módulos fodorianos (o algo suficientemente cercano a ello; *cf.* Fodor 1983, cap. 3)<sup>89</sup> y argumentaré que la(s) capacidad(es) claramente fiable(s) que los matemáticos tienen no se debe a un único módulo fodoriano. Si eso es el caso, la fiabilidad de una creencia no requiere de una capacidad dedicada a producir creencias de ese tipo, y NOZICK3 es falsa. (No quiero comprometerme con la idea de que la mente de hecho sea modular, en el sentido de Fodor o uno parecido. Creo que es ha de ser decidido empíricamente. Pero lo asumiré por mor de la discusión.)<sup>90</sup>

Tomaré como ejemplo el caso de la cognición matemática—un área de las ciencias cognitivas todavía en desarrollo, pero con importante evidencia que apunta hacia ciertas tesis. Por ejemplo, la evidencia disponible indica que infantes de muy temprana edad tienen la capacidad de discriminar numerosidades pequeñas (*i.e.*, de notar cambios en la cardinalidad de un conjunto con pocos miembros) y hacer operaciones aritméticas muy simples (sumas y restas de pocos elementos), y algo parecido sucede con algunos primates no-humanos: esto indica que al menos ciertas capacidades matemáticas básicas son innatas y con larga historia evolutiva (Butterworth 2005a; Dehaene 1997; Hauser *et al.* 1996; Izard *et al.* 2009; Wynn 1992). Además, parecen haber sistemas neurales funcionalmente especializados para tales capacidades (Dehaene 1997; Dehaene *et al.* 2003; Piazza *et al.*, 2003). También se muestran patrones característicos y específicos de daño (Butterworth 2005b). Todo eso parece evidenciar que tenemos una facultad—un módulo—para ciertas capacidades matemáticas *básicas*: el reconocimiento de números y operaciones aritméticas simples. Sin embargo ¿tenemos un módulo para capacidades matemáticas más complejas? ¿Tenemos un módulo para, digamos, la habilidad de hacer topología algebraica o teoría de categorías?

Esta cuestión es todavía más difícil, principalmente porque ha sido mucho menos estudiada. En primera, ya desde la cognición matemática básica, tenemos interacción con la cognición espacial (Dehaene *et al.* 1993; Hubbard *et al.* 2005), y los expertos (no necesariamente matemáticos) en

---

89 Según la tradición influida por el trabajo de Fodor, y siendo bastante esquemáticos, las características *típicas* de un módulo son (*cf.* Fodor 1983, cap. 3):

- \* *Especificidad de dominio*. Un módulo se especializa en cierto tipo de insumo.
- \* *Disparos obligatorios*. La operación de los módulos es obligatoria, en el sentido en que no está abierta a decisión voluntaria o influida por creencias.
- \* *Accesibilidad central limitada*. La información que es procesada por un módulo es accesible al sistema central sólo hasta cierto punto.
- \* *Rapidez*. La manera en que los módulos procesan información es más rápida que el procesamiento fuera de los módulos.
- \* *Encapsulación informacional*. Un módulo procesa información con independencia de la información procesada por otros módulos.
- \* *Superficialidad (“shallowness”) de las salidas*. Un módulo devuelve como salida típicamente una representación “a medio nivel” en una escala de abstracción.
- \* *Arquitectura neural fija*. Hay una arquitectura neural específica para cada módulo.
- \* *Patrones característicos y específicos de daño*. Relacionado con su arquitectura neural fija, los módulos tienen patrones de daño específicos a cada uno.
- \* *Ontogenia característica*. Los mecanismos neurales que subyacen a los módulos se desarrollan de acuerdo a patrones específicos que obedecen estímulos ambientales.

90 Agradezco los comentarios de Alejandro Vázquez aquí.



hacer cálculos hacen un uso no trivial de la memoria episódica (Butterworth 2001). Estas facultades cognitivas no son específicas de dominio, lo que haría pensar que aprender matemáticas más allá del nivel más elemental involucra a otras capacidades además del—si es que las hipótesis arriba mencionadas son ciertas—“módulo de la aritmética básica”. En segunda, algunos estudios han mostrado que sujetos con discalculia adquirida (la incapacidad de hacer cálculos aritméticos) todavía pueden hacer cálculos algebraicos (Dehaene 1997: 199; Hittmair-Delazer et al. 1994, 1995), lo que sugiere que “el módulo de la aritmética” puede dañarse, y aún así la capacidad para hacer álgebra seguir funcionando.<sup>91</sup> Esto lleva a pensar que el módulo de la aritmética, si hay tal cosa, simplemente no basta para dar cuenta de todo el conocimiento matemático disponible—una matemática tiene que usar otras facultades, diferentes de tal módulo, para obtener su conocimiento matemático.

¿Por qué es esto relevante para la epistemología modal? Bueno, NOZICK3 afirma que necesitamos de un módulo para detectar necesidades, si es que hemos de tener conocimiento de ellas. En los últimos dos párrafos he argumentado que, aunque probablemente tengamos un módulo para la aritmética básica, no parece que tengamos un módulo para las matemáticas más complejas. ¿Muestra esto que no tenemos justificación para nuestro conocimiento matemático? No lo creo. Si acaso muestra algo, es una *reductio* a una premisa análoga a NOZICK3 para el caso del conocimiento matemático. Por analogía, NOZICK3 queda puesta en duda también—pues no considera la posibilidad de que nuestro conocimiento de necesidades no se base en un sólo módulo, sino que provenga de información procesada por varias capacidades cognitivas. (De hecho, una hipótesis interesante, y con carga empírica, es que—muy en analogía con nuestra cognición matemática—tenemos una facultad para el pensamiento modal “básico”: para imaginar posibilidades a niveles de representación que emulan la percepción diaria. Esto, creo, recibe cierto soporte de la evidencia empírica repasada en la §4.1, arriba. Una segunda hipótesis es que las bases cognitivas de nuestro conocimiento modal más complejo—de nuevo en analogía con la cognición matemática más compleja—involucran diferentes capacidades, que ayudan a una filósofa a llegar a la experticia, justo como las otras capacidades ayudan a un matemático a llegar a ella.)

He argumentado que las tres premisas del argumento de Nozick son o implausibles, o presuponen información empírica de la que de hecho no disponemos. Concluyo que, gracias a esta característica, su argumento es suficientemente débil como para no tener una fuerza dialéctica considerable en contra del defensor del conocimiento de necesidades. Con ello también, las respuestas de Nozick a las tareas ESCEPTICISMO1-3 son puestas en duda.<sup>92</sup>

Quiero cerrar esta sección notando una curiosa forma de auto-“refutación”, por decirle de alguna forma, en el argumento de Nozick. Parece que una intuición detrás de él es que (por decirlo crudamente), para cazar, reproducirnos o detectar depredadores no necesitábamos conoci-

91 También notemos que, más especulativamente, las matemáticas más allá del nivel de la escuela primaria presuponen heurísticas que no son sólo aritméticas: algunas, por ejemplo, son espaciales (e.g., la geometría), otras simbólicas (e.g., la lógica matemática), y otras aún más abstractas que la aritmética (e.g., el álgebra, o el álgebra abstracta); otras aún son análogas (e.g., el análisis), y no discretas como la aritmética básica.

92 ¿Podríamos considerar teorías alternativas a la fodoriana, como el conexionismo? Sería interesante, aunque el nozickiano no ha mostrado que se pueda extender el razonamiento para tales modelos más sofisticados de la cognición. (Agradezco la necesidad de esta nota a Lourdes Valdivia)

miento de necesidades, y que por ello no hay razón alguna para pensar que nuestro aparato cognitivo sería sensible a lo que es necesario. El problema con esta intuición es que para cazar o reproducirnos o detectar depredadores tampoco necesitamos mecánica cuántica, criptografía o ingeniería de elevadores. Con este razonamiento, la intuición de Nozick tendría amplias e implausibles consecuencias escépticas. Y, sobre todo, al pensar que “sin ventaja evolutiva, no hay conocimiento real”, por así decir, Nozick se corta las propias manos: seguramente la biología evolucionista no sirve para cazar, reproducirnos o detectar depredadores—pero, si no es conocimiento real, Nozick no podría usarla para argumentar en contra del metafísico modal.

#### 44. conclusión, y una nota sobre el problema de Benacerraf

He examinado dos importantes retos a nuestro conocimiento modal y de esencias, y he argumentado que ninguno de ellos muestra que de hecho *no* poseamos tal conocimiento. Por supuesto, una cosa es eso—y muy otra es mostrar que *de hecho* lo poseemos. El capítulo siguiente pone, por así decir, los fundamentos para una epistemología que busca vindicar la asunción de posesión; el capítulo final—el sexto—propone una epistemología modal construida sobre los fundamentos fiabilistas del quinto capítulo, que, si exitosa, vindica tal asunción.

Me gustaría cerrar este capítulo comentando sobre otra posible fuente de escepticismo modal: el problema de Benacerraf para los hechos sobre objetos abstractos (Benacerraf 1973; *cf.* Field 1989, secc. 7.2). En breve, el problema de Benacerraf surge cuando notamos que en nuestras teorías matemáticas y científicas estándar ocurren términos que, presuntamente, refieren a objetos matemáticos (“2”, “el grupo de Lorentz del espacio de Minkowski”), y ocurren en posición referencial. Es decir, si tales teorías han de ser verdaderas, entonces deben existir objetos matemáticos. Pero una asunción compartida por la mayoría de los filósofos de las matemáticas es que los objetos matemáticos no son físicos—y por ello, no pueden entrar en interacciones causales con nosotros. Pero, entonces, ¿cómo se supone que podemos tener conocimiento acerca de tales cosas? Se supone que para que conozcamos algo—o, según las teorías causales de la referencia iniciadas por Kripke (1980), para que podamos referirnos a ese algo—esa cosa tiene que poder *afectarnos* de cierta manera. Pero sin influencias causales ¿cómo podría hacerlo?

Hay quienes creen que un problema análogo se extiende para el conocimiento modal (*cf.* Hawthorne 1996; Peacocke 1999: caps. 1, 4; Vaidya 2007: secc. 4). Dado que se asume que el fundamento de los hechos modales son hechos acerca de mundos posibles, y dado que se asume que tales cosas son objetos abstractos, surge de inmediato un problema benacerrafiano para nuestro conocimiento modal.

No consideré a detalle ese problema en este capítulo porque, de entrada, la metafísica modal que yo favorezco (desarrollada en la primera parte de esta tesis) *no* asume que el fundamento de los hechos modales sean hechos sobre mundos posibles—de hecho, invertí la flecha explicativa y argumenté que los mundos posibles están fundamentados en hechos sobre “constreñimientos modales”: cosas como leyes, principios generales, o incluso consecuencias de las esencias de los objetos (si las hay). En la imagen de los hechos a favor de la cual yo argumenté, tenemos prime-

ro hechos modales dictados por los constreñimientos; luego usamos a los mundos posibles—esas construcciones matemáticas—como un *medio representacional* de esos hechos modales.

La pregunta de cómo conocemos tales constreñimientos modales es una pregunta justa—pero *distinta* del problema benacerrafiano—y que consideraremos en lo que resta de la tesis (sobre todo en el sexto y último capítulo).

## 5. Fiabilismo externista

*For me, normative epistemology is a branch of engineering. It is the technology of truth-seeking [... ] It is a matter of efficacy for an ulterior end, truth or prediction. The normative here, as elsewhere in engineering, becomes descriptive when the terminal parameter is expressed.*

—W.v.O. Quine, “Reply to Morton White”

Este capítulo es un *excursus* en la epistemología general—la investigación sobre las condiciones del conocimiento, sea este modal o no. Aquí defenderé una versión particularmente tolerante del externismo, así como una versión particularmente general del fiabilismo. Antes de ello, haré una diferencia importante—y que muchas veces se pasa de largo—entre teorías sobre la *naturaleza* del conocimiento y teorías sobre la *atribución* del mismo; en el camino, haré algunos comentarios sobre meta-filosofía: el papel de las intuiciones en la epistemología y su relación con filosofías particulares acerca de los conceptos y el lenguaje. Entonces pasaré a argumentar que el externismo defendido (que llamo “mínimo”) es compatible con varias tesis internistas; por lo que en principio se abre la posibilidad de tener una epistemología de “dos niveles”—aunque aquí ya no me ocuparé de esto último. Entonces explicaré cómo es que entiendo a la tesis fiabilista general que explotaré en el último capítulo, y la defenderé de uno de sus mayores problemas: el problema de la generalidad. Argumentaré que ese problema se basa en una exigencia que parece ser casi irracionalmente demandante: la exigencia de satisfacer presupuestos que, si no inconsistentes, al menos están en una fuerte tensión entre sí; por lo que la carga de la prueba sobre el fiabilista se disminuye considerablemente.<sup>93</sup> Con una motivación y defensa del externismo y del fiabilismo en la mano, pasaré, en el siguiente y último capítulo, a proponer una epistemología modal que se ajuste a estos criterios.

Pero antes una nota importante. No me ocuparé aquí del debate acerca de si el conocimiento puede o no ser analizado. La tradición analítica del siglo XX dedicó una enorme parte de los esfuerzos en epistemología a intentar ofrecer condiciones necesarias y suficientes para el conocimiento—y todavía más después del revolucionario artículo de Gettier (1963). Sin embargo, se ha argumentado que los contraejemplos de Gettier surgirán *inevitablemente* si uno asume que la condición de justificación en el análisis propuesto es *fiabilista*: es decir, si se admite que uno puede tener justificación para una creencia falsa (cf. Zagzebski 1994 y Floridi 2004, quien ofrece una prueba en términos de la lógica epistémica). Si esto es correcto, para que un análisis del conocimiento sea inmune a problemas estilo Gettier, debe negar que se pueda estar justificado para creer una falsedad. Este tipo de infalibilismo es problemático por varias razones, de las que no me ocuparé aquí (cf. Hetherington 2012).

Todavía más, el libro de Williamson (2000) puso en duda la necesidad de analizar al conocimiento en factores supuestamente más básicos. Además, mostró la viabilidad del proyecto primitivista, de acuerdo al cual el conocimiento es un estado mental inanalizable, que más bien

<sup>93</sup> Una versión anterior incluía una discusión de otro supuesto problema para el externismo, el así llamado “problema de la circularidad epistémica”, a veces no claramente distinguido del “problema del conocimiento fácil”. Por razones de espacio he tenido que suprimir tal discusión.

puede utilizarse para definir otros conceptos epistémicos (como *evidencia*). Pero eso no quiere decir que no dijera nada sobre la *estructura* del conocimiento. Caracterizó su estructura lógica como la de un operador de estados mentales fácticos (2000: introducción), como el estado mental fáctico más general (cap. 1), compatible con la filosofía externista de la mente (cap. 2), no-decomponible (cap. 3) ni accesible directamente a la introspección (cap. 4), y también como un estado que se puede *caracterizar* (no *definir*) como el estado en que uno está cuando está en lo correcto y no podría haberse equivocado fácilmente (cap. 7; 2009).

Esto nos brinda evidencia de que un proyecto que no busque analizar al conocimiento todavía puede darnos una epistemología informativa, clara y plausible (además de fructífera, como Williamson mostró al aplicar sus ideas a otras ramas de la epistemología). Yo tampoco me comprometeré con un conjunto particular de condiciones supuestamente necesarias y suficientes del conocimiento. De hecho, no busco comprometerme con la idea de que el conocimiento sea analizable—ni con que no lo sea. Simplemente dejo la cuestión de lado, y las tesis externistas y fiabilistas que defenderé se ofrecen como *caracterizaciones* de su estructura y de sus relaciones con otros estados y propiedades filosóficamente importantes—que, a lo más, son condiciones necesarias. Por ponerlo metafóricamente: si el proyecto de analizar al conocimiento busca entenderlo “desde adentro” (desde sus condiciones constituyentes), proyectos como el de Williamson buscan entenderlo “desde su estructura externa”, por cómo ayuda a resolver y enfocar otros problemas, y cómo se comporta en diferentes contextos. Lo que pretendo en este y el siguiente capítulo es un proyecto que quizá caiga mejor en el segundo campo, pero no se limita a ello.

Una idea intuitiva y cruda que tenemos acerca del conocimiento bastará para poder avanzar por este capítulo, y es con lo único que me quiero comprometer de entrada: intuitivamente, tenemos conocimiento—sabemos algo—*cuando nuestra mente se apega con cierta seguridad al mundo*. Esto codifica al menos la idea de que el conocimiento es un estado mental representacional que es fáctico y fiable; pero no es la única intuición que se ha explotado acerca del conocimiento.<sup>94</sup> En lo que sigue, refinaré esta caracterización intuitiva del estado acerca del cual esta-

---

<sup>94</sup> Otras epistemologías importantes, de cuño relativamente reciente, creen que el conocimiento es un cierto tipo de *logro*, de manifestación de una habilidad cognitiva (Sosa 2007), que puede ser una facultad fiable (para Sosa) o un rasgo estable del carácter de la persona (para Zagzebski, 1996). Es posible que la idea intuitiva del conocimiento como logro pueda incluirse o fusionarse naturalmente con la intuición del conocimiento como representación segura del mundo: representarse de manera estable al mundo es cierto tipo de logro, para el cual se necesitan tanto equipamiento cognitivo confiable—nuestras facultades mentales—como rasgos estables de carácter que de manera confiable lleven a tales representaciones.

Todavía otras, muy probablemente venidas de la tradición de Descartes, ven al conocimiento como un estado que de manera sistemática necesita esfuerzo consciente del agente—para ajustar sus creencias a la evidencia, por ejemplo, o para notar que sus creencias forman una red coherente de soporte mutuo, o que pueden retro-traerse a creencias básicas. Es posible que también la idea fiabilista básica enunciada arriba sea compatible con estas ideas. Pues si los defensores de ellas logran mostrar que formar redes de soporte mutuo o estar basadas en creencias básicas brinda representación del mundo correcta y estable, sus ideas serán compatibles con la intuición fiabilista.

Todavía otra tradición, que bebe de Platón y se alimenta de otras ideas contemporáneas, ve al conocimiento como esencialmente construido *en sociedad*—un platónico dirá que depende de la actividad dialéctica; otros dirán que dependen de mecanismos de poder que implantan ciertos sistemas representacionales. De nuevo, estas ideas son compatibles con el fiabilismo siempre y cuando logren mostrar que los mecanismos sociales en cuestión nos llevan, con relativa seguridad, a representar correctamente al mundo.

No podré tocar más todas estas interesantes ideas en lo que sigue.

mos hablando.

### 5.1. preeliminares: ontología/atribución, y la función de las intuiciones

Lo primero de todo es hacer una distinción, central, entre dos tipos de teorías epistemológicas: las teorías sobre la *naturaleza* del conocimiento, y las teorías acerca de la *atribución* de este. Hay cierta interdependencia de una sobre la otra, en un sentido que intentaré aclarar abajo.<sup>95</sup>

Distingamos, pues, entre dos tipos de teorías epistemológicas: *ontológicas* y *atributivas*.

Una teoría *ontológica* acerca de una propiedad o estado epistemológicamente relevante  $x$  (como el conocimiento, la justificación o la racionalidad) es una teoría acerca de la *naturaleza* de  $x$ —acerca de qué es  $x$ . Así, una teoría ontológica especificará las condiciones de verdad para el esquema  $\lceil$ La creencia de  $S$  de  $\langle p \rangle$  en un momento  $t$  tiene la propiedad  $x$  $\rceil$  (tomé este esquema de Goldman 1979: p. 107).

Una teoría *atributiva* acerca de  $x$  es una teoría acerca de cómo los miembros de una comunidad lingüística se atribuyen, mediante el lenguaje, la posesión de  $x$ . Especificará las condiciones de verdad para el esquema:  $\lceil S_1$  atribuye a la creencia de  $S_2$  en  $\langle p \rangle$  en un momento  $t$  la propiedad  $x$  $\rceil$ .

Afirmé que hay cierta inter-dependencia entre los dos tipos de teorías. Esto lo creo porque al atribuir justificación (u otras propiedades epistémicas) a otros agentes—reales o incluso meramente imaginados—hacemos explícitas las *intuiciones* que tenemos acerca de lo que *es* el conocimiento. Aquí hay dos extremos en los tipos de acercamientos posibles a la función de estas intuiciones—a uno lo llamaré el acercamiento “intenso” y al otro el acercamiento “relajado”. Entre estos dos extremos se tiende una gama de acercamientos más, o menos, intensos o relajados. Veamos.

El acercamiento *intenso* a la función de las intuiciones está cercanamente relacionado con el descriptivismo en la filosofía del lenguaje y con cierto acercamiento a las representaciones mentales, y existen varias variantes de él.<sup>96</sup> Aquí tomaré un núcleo radical, que después suavizaré. Así, la versión más fuerte de tal acercamiento nos dice que las intuiciones (i) *fijan* la referencia de los términos o conceptos que usamos, (ii) son *necesarias* para que tales términos o conceptos refieran, y (iii) son *constitutivas* de la posesión del concepto o el uso del término.

95 Esta distinción está relacionada con, pero no es la misma que, la distinción de Goldman entre justificación *fuerte* y la *débil* (Goldman 1988). De acuerdo a Goldman en ese artículo, podemos distinguir entre la justificación *de facto* que una creencia tiene (que bajo la postura fiabilista de Goldman se deberá a su ser producto de procesos fiables), y la justificación que una creencia puede tener debido a haber sido adquirida de manera no culpable (“*blamelessly*”), es decir, adquirida con el mejor esfuerzo de parte del creyente, aún cuando el proceso que produjo la creencia en cuestión de hecho no sea fiable. (Por ejemplo, un agente en una cultura con una ciencia poco desarrollada, podría hacer su mejor esfuerzo al adquirir creencias sobre un tema dado, y esto es compatible con que su método de adquisición sea la consulta al oráculo—si es que en esa cultura los expertos coinciden en que consultar al oráculo es un método fiable; por más que *en realidad* no lo sea.) La primera justificación es la *fuerte*; la segunda es la *débil*. La distinción de Goldman es sólo para la justificación epistémica, mientras que la distinción que he trazado arriba pretende cubrir también otras propiedades o estados epistémicos.

96 Por ejemplo: Jackson (1995; 1998: cap. 2), Lewis (1966, 1970, 1994) y Peacocke (1992). Para críticas, veáanse los trabajos de Kripke, Putnam y Donnellan citados abajo; ver también (Williamson 2007: caps. 3 y 4).

Tomemos el término “conocimiento” en español. La idea (i) del acercamiento intenso nos dice que las *intuiciones*—aquellas creencias, que se asumen presentes antes de hacer epistemología, sobre cómo aplicar el término—de los hablantes competentes del español expresan las propiedades que el referente intencional tiene. El criterio (ii) nos dice entonces que, una vez sistematizadas tales intuiciones, las propiedades en cuestión *deben* satisfacerse si es que “conocimiento” ha de tener referencia. Si no se satisfacen, entonces “conocimiento” no refiere—no hay, estrictamente hablando, conocimiento. El criterio (iii) nos dice que nadie que esté hablando español puede contar como usando el término “conocimiento” si no tiene tales intuiciones.

Algo parecido pasa con los conceptos: supongamos que hay un concepto expresado por el término en español, “conocimiento” (es el concepto de, por supuesto, el *conocimiento*, y se asume que también es expresado por términos en otros lenguajes: ‘knowledge’, ‘connaissance’, ‘wissenschaft’, ‘conoscenza’, ‘episteme’, ‘scientia’, etc.). Entonces, (i) nos dice que los usos (u ocurrencias en estados mentales intencionales) de tales conceptos nos dan evidencia de las propiedades que caracterizan al *referente intencional* del concepto (aquello que pretendemos referir con tal representación mental). La idea (ii) nos dice que esas propiedades son necesarias para que el concepto refiera: si nada las ejemplifica, entonces el concepto no se aplica en el mundo. La idea (iii) nos dice que, para que alguien cuente como poseyendo el concepto de *conocimiento*, debe tener, al menos disposicionalmente, las intuiciones asociadas a él.

Por supuesto, la postura intensa aquí descrita es bastante radical, y tiene cada vez menos adherentes. (quizá porque parece estar comprometida con una línea clara entre lo analítico y sintético, algo que Quine (1952) puso, famosamente, en duda—ver Machery 2009: 40).

Pero existen versiones más suaves de ella (que no son todavía la postura “relajada”, que describo abajo). Por ejemplo, las condiciones (i)-(iii) pueden relajarse, afirmando que la aceptación de *todas* las intuiciones asociadas a un término o a un concepto *no* son condición necesaria para el uso o posesión de tal término/concepto. Se podría afirmar que sólo cierto subconjunto de ellas contiene las intuiciones “privilegiadas” o, quizá más plausiblemente, que ninguna es privilegiada, pero que sí es necesario que el hablante o pensante concuerde con al menos una cantidad notable de ellas (al estilo de la “*cluster theory*” de Searle (1958)). Otro tipo de debilitaciones son posibles, pero esto nos da una idea básica del tipo “intenso” de acercamiento a la función de las intuiciones.

El segundo acercamiento, el *relajado*, va más en consonancia con la teoría causal de la referencia (o “de la referencia directa”), motivada por el trabajo de Kripke (2005), Putnam (1975) y Donnellan (1966), y con el externismo semántico (la idea de que las propiedades semánticas de un término, concepto o pensamiento pueden ser fijadas por factores externos al sujeto que los tiene o usa).<sup>97</sup> Aquí encontramos dos ideas: (i) la referencia de un término o concepto es *fijada* por las intuiciones asociadas a él, pero (ii) *no* se necesita que tales intuiciones sean verdaderas para que tal concepto o término refiera, *ni* la aceptación de tales intuiciones es necesaria para la posesión del concepto o uso del término. Regresemos al ejemplo de arriba: el uso del término “conocimiento” puede estar asociado con intuiciones de variados tipos. Por ejemplo, supongamos que los hablantes del español tendieran a no aplicarlo cuando no estuvieran frente a infali-

97 Para la aplicación de un acercamiento “relajado” a las propiedades que le interesan a los epistemólogos, ver Kornblith 2002, capítulo 1. Otro ejemplo interesante es el de la simultaneidad en la teoría de la relatividad, en Field 2009.

bilidad: si alguien no es infalible acerca de un tema, los hablantes no le atribuirían conocimiento de tal tema, así que no usarían el término en cuestión. Un epistemólogo que aceptara algún tipo de acercamiento intenso podría verse obligado<sup>98</sup> a juzgar que el conocimiento necesita de la infalibilidad: como ella es una marca necesaria, no hay conocimiento sin ella. No así un epistemólogo que aceptara algún tipo de acercamiento *relajado*: él podría afirmar que tales intuiciones sólo ayudan a *fijar* el referente intencional, pero que tal referente intencional puede resultar no poseer las propiedades que se le atribuyen en tal caracterización.

Aquí no voy a asumir un acercamiento “muy intenso” a las intuiciones de los hablantes y pensantes, pero tampoco uno “muy relajado”. Mi hipótesis de trabajo es que las intuiciones que asociamos con un concepto o término (i) caracterizan el referente intencional, es decir, *fijan* la referencia del concepto o término pero (ii) tal referente intencional puede resultar ser muy distinto a como se lo caracteriza según ellas, con lo que varias, o todas, de las intuiciones puede resultar ser falsas, aunque (iii) debe haber algún vínculo entre nuestra concepción resultante y la primera, la intuitiva—en el sentido en que debe poder trazarse una evolución continua desde nuestra concepción pre-teórica y la resultante, afinada por la teoría.

Con la idea (i), pretendo ser consonante con el requisito que tanto el acercamiento intenso como el relajado aceptan, mientras que con (ii) muestro que mi acercamiento es más relajado que intenso, pero con (iii) hago notar que, para que podamos decir que nuestros términos o conceptos de hecho resultan referir (para que podamos decir que, por ejemplo, *hay* conocimiento, o mesas, o libre voluntad) por más que resulte que nuestra pre-concepción de ellos es errónea (por más que resulte que no tenemos infalibilidad, o que sólo hay partículas arregladas en forma de mesa, o que hay determinismo), tenemos que poder trazar una continuidad desde nuestra pre-concepción a la concepción refinada por la teoría, de tal manera que al responder la pregunta de si existe  $x$  con algo como “no existe precisamente lo que tú crees que es  $x$ , pero existe  $y$ ” no estamos simplemente cambiando de tema. Tales casos son bastante comunes al hacer teoría—y no deberíamos asumir de inicio que la mayoría de nuestras creencias pre-teóricas tienen que ser verdaderas, si es que hemos de lograr referirnos a fenómenos objetivos.

Por supuesto, qué tipo de “continuidad” deba trazarse es el problema clave. Por ejemplo, suponiendo que una física fundamental que postule partículas (y no campos extendidos por todo el espaciotiempo, por ejemplo) resultara ser la correcta, ¿deberíamos contar eso como el éxito de Demócrito y otros adherentes al atomismo de antaño? ¿Es la concepción atomista, revelada en las intuiciones que los atomistas de antaño expresarían, suficientemente continua con el atomismo de la física futura, de tal manera que los átomos postulados por esa física puedan contar como átomos “suficientemente democríteos”? No estoy seguro de que haya una respuesta general para este tipo de preguntas. Quizá la pregunta de si una pre-concepción es “suficientemente continua” con una concepción refinada por la teoría es, además de obviamente vaga (debido al término “suficientemente”, que introduce vaguedad), capaz de ser respondida en varias maneras, de tal suerte que no haya una regla que brinde sistematicidad a esta cuestión. Por el momento, al menos, no sé de ninguna pista que nos dé evidencia para pensar que hay tal sistematicidad. Así que, por el momento, mi creencia se inclina hacia pensar que los criterios que fijan si un referente cuenta o no como el referente de un término, o concepto, cuya pre-concepción aso-

---

<sup>98</sup> Dependiendo, claro, de qué tan intensa sea su aceptación del acercamiento intenso—es decir, de qué tanto haya debilitado tal acercamiento.



ciada resultó ser al menos mayoritariamente errónea, son altamente dependientes de contexto, y además poco sistemáticos.

Ése es, pues, el acercamiento que tomaré a la función de las intuiciones. Ahora podemos regresar a la dependencia entre las teorías ontológicas y atributivas de la que hablé arriba. La teoría ontológica depende de la atributiva porque es mediante la teoría atributiva que podemos sistematizar las intuiciones asociadas a la aplicación de términos y conceptos epistémicos: es gracias al uso de tales términos o conceptos, expresados en la atribución epistémica, que sabemos qué marcas se le asocian a estados como el conocimiento, o propiedades como la justificación o la racionalidad. Con estas intuiciones en mano, podemos pasar a hacer una teoría sobre qué cosa en el mundo—si es que alguna—las hace verdaderas; al menos, qué cosa es referida por aquellos conceptos o términos que se asocian con tales intuiciones. Así, en el *ordo cognoscendi*, vamos de la teoría atributiva a la teoría ontológica. (Esto involucra, por supuesto, una cierta idealización: como nadie puede afirmar, todavía, poseer la última teoría atributiva, no podemos afirmar que tengamos una base segura y completa sobre la cual fundar nuestra teoría ontológica—el proceso es más dinámico y holista.) Asumiendo el acercamiento medianamente relegado que mencioné arriba, varias de, o todas, nuestras intuiciones pueden resultar ser falsas—pero, si encontramos que las propiedades o estados epistémicos, para los cuales tenemos buenas razones para creer que existen, son rastreables a nuestras pre-concepciones, entonces sí que habrá *conocimiento*, *razonabilidad*, etcétera.

Por supuesto, en el *ordo essendi*, la corrección de nuestras atribuciones epistémicas depende de *qué sean* aquellas propiedades o estados que atribuimos. En este sentido, una teoría atributiva depende de una teoría ontológica: las condiciones de verdad de las atribuciones son fijadas por el comportamiento lingüístico de los atribuyentes, pero su satisfacción depende de aquello que tales agentes atribuyen—y qué sea aquello que atribuyen nos es dado por la teoría ontológica. (Claramente, aquí hay el mismo tipo de idealización que mencioné arriba: la relación real entre ambas teorías es mucho más dinámica y holista.)

Ahora podemos pasar a la defensa que propondré del fiabilismo y el externismo. Para ella, como veremos en §5.4, será crucial la diferencia entre teorías ontológicas y atributivas.

## 5.2. motivando un externismo mínimo

En esta sección examinaré a, y argumentaré a favor de, la tesis que llamo “externismo mínimo”, que es una tesis acerca de nuestro conocimiento. Tiene dos componentes: uno ontológico y uno atributivo, que expondré y comentaré por turnos.

(EXTERNISMO ONTOLÓGICO MÍNIMO)

Una teoría ontológica *T* es aceptable sólo si:

- (a) *T* acepta que estamos en posesión de facultades (procesos, habilidades, heurísticas, métodos, etcétera) que *conectan* a un agente epistémico con su ambiente—en el sentido en que cada una de esas facultades transporta información desde el ambiente hasta la mente del agente.<sup>99</sup> Llamo a tales facultades, etc., *fuentes de información externa* (FIEs);

<sup>99</sup> Pretendo que esta noción de “transportar información” sea compatible con que los mecanismos cognitivos tam-

- (b) *T* acepta que las FIEs pueden transportar tal información sin que nosotros seamos conscientes de ello o de cómo lo hacen;
- (c) *T* acepta que el que poseamos o no conocimiento de al menos muchas de las proposiciones que de hecho conocemos, está fijado esencialmente por las FIEs, en el sentido en que, para que conozcamos esas proposiciones, además de la creencia verdadera es *necesario y suficiente* que las FIEs funcionen *bien*.

Comentarios: he sido deliberadamente poco exacto al formular la idea de que las FIEs funcionen “bien”, pero abajo diré más acerca de este tema. También hay que notar que he hablado de *conocimiento*, y no de *justificación*. Tengo dos razones para esta decisión. Primera, quien asume que el conocimiento es analizable, regularmente asume que la justificación (o el *entitlement*, o el *warrant*, o alguna propiedad parecida) es un componente necesario, pero no suficiente, para el conocimiento. Pero lo que hace que el conocer una proposición dada sea un hecho describable por una teoría externista, y no una internista, es que la justificación en cuestión sea externista. Pero que la justificación sea o no externista termina dependiendo sobre factores del estilo de (a)-(c). Así que no perdemos generalidad si dejamos la justificación y nos pasamos al conocimiento directamente. Segunda, no todos creen que el conocimiento sea analizable (Williamson 2000), por lo que caracterizar el debate internismo/externismo como un debate, de entrada, sobre la justificación, resulta en pérdida de la generalidad. Pero como vimos con el punto antes mencionado, caracterizarlo como un debate sobre el conocimiento no resulta en tal pérdida. Otra cuestión es que (c) parece contradecir la idea de *no comprometerse con un análisis* del conocimiento, pues dice que para “al menos muchas de las proposiciones” (abajo diré más sobre cuáles) es equivalente el que las FIEs funcionen “bien” con que se tenga conocimiento. Pero esto no necesariamente es un análisis, pues puede resultar que el *buen* funcionamiento de las FIEs en cuestión sólo se pueda entender en términos del conocimiento mismo (*cf.* Williamson 2000, cap. 7; 2009).

Ahora podemos formular el segundo componente del externismo mínimo que defenderé:

(EXTERNISMO ATRIBUTIVO MÍNIMO)

Una teoría atributiva *T* es aceptable sólo si: *T* acepta que nuestras prácticas de atribución de conocimiento en muchos casos *únicamente* toman en cuenta las cualidades de las FIEs del agente cuya creencia es evaluada.

De nuevo, he hablado de “muchos casos”, lo cual es poco específico, y abajo diré más sobre esto.

Me referiré a la conjunción de ambas tesis como *EXTERNISMO MÍNIMO*. Quiero ahora mostrar por qué *EXTERNISMO MÍNIMO* es un *externismo*, luego diré por qué es un *externismo mínimo*.

En primera, la tesis es externista porque afirma que las características de las FIEs son relevantes para la evaluación epistémica. Así, según *EXTERNISMO MÍNIMO*, no necesitamos tener *acceso consciente* al funcionamiento de las FIEs, incluso si estas son una presuposición del conocimiento; además, la información que proveen no proviene de la mente del agente, de lo que resulta que el conocimiento presupone información externa. Nuestros estados de conocimiento (i) son fijados—al menos en *muchos* casos—por factores externos a nuestra mente, y que además (ii) no necesitan de nuestro acceso consciente para que ocurran. Estas dos características

---

bién la *filtren*, *procesen* o *transduzcan* de varias maneras. Además, la noción es deliberadamente amplia, pues especificar qué es transportar información del mundo a la mente es, creo, un trabajo que queda para la ciencia cognitiva.

son definitorias de la postura externista.

EXTERNISMO MÍNIMO no sólo enuncia una tesis que es externista, sino que es *mínima* en el sentido en que no implica que *todo* nuestro conocimiento sea como lo caracteriza el externista. Por todo lo que la tesis dice, alguna otra tesis internista podría ser verdadera—es sólo que el pre-requisito de nuestro conocimiento está bien descrito, según EXTERNISMO MÍNIMO, por una tesis externista.

Mi entendimiento del internismo aquí se enfoca en lo que se ha llamado “internismo de acceso”, que caracterizo así:

(INTERNISMO DE ACCESO)

Un agente epistémico  $x$  está justificado en creer  $\langle p \rangle$ , sólo si  $x$  tiene acceso consciente a los factores que justifican tal creencia.<sup>100</sup>

Ahora bien, esta condición es, argüiblemente, altamente restrictiva, y puede fácilmente resultar en escepticismo particularmente fuerte (cf. Pappas 2005, §§2-3; Van Cleve 2003). Así que asumiré que el internista debe restringir su postura. Pero, ¿cómo?

Una salida es restringir la postura internista para casos de conocimiento que, de una manera putativamente intuitiva, merezcan mayor “esfuerzo” y por ello sean más altamente valoradas. La idea, cuyo *pedigree* puede rastrearse al menos hasta Descartes o incluso hasta Platón, es que, como diría la frase coloquial: “todos somos iguales—pero unos somos más iguales que otros”. El conocimiento nacido de la percepción es de, por así decir, un grado más bajo que el conocimiento nacido de la reflexión y el razonamiento—y por ello, es valorable en mayor medida. Sosa y Klein formulan así la idea:

Lo que debería ser obvio es que el conocimiento del tipo de los detectores no es lo que es distintivo de los humanos adultos—o al menos de aquellos de nosotros que buscan ser agentes epistémicamente responsables que hemos examinado nuestras creencias y buscan sostener sólo aquellas que, después del examen, son dignas de creerse. Desde esta perspectiva, las creencias sin examinar todavía no son dignas de creerse. Buscamos tener buenas razones para nuestras creencias.\* (Klein 2007, p. 5)

Y Sosa:

Hay que admitir que hay un sentido en el que incluso una puerta del supermercado “sabe” cuándo alguien se acerca, y en el que un sistema de calefacción “sabe” cuándo la temperatura en una habitación sube más allá de cierto ajuste. Tal es el conocimiento “servo-mecánico”. Y hay una variedad inmensa del conocimiento animal, instintivo o aprendido, que facilita la sobrevivencia y la prosperidad en una diversidad sorprendentemente rica de modos y ambientes. El conocimiento humano está en un plano más alto de sofisticación, de cualquier manera, precisamente por su coherencia

<sup>100</sup>Véase Gibbons 2006. Esta no es la única manera de entender al internismo: algunos podrían entenderlo como la tesis de que los factores justificatorios dependan únicamente de eventos mentales (Connee & Feldman 2004), o como la tesis, análoga a una tesis discutida en filosofía de la mente, de que los factores justificatorios dependan únicamente de la naturaleza intrínseca del sujeto en cuestión (Nuccetelli 2003). Aquí trataré únicamente con INTERNISMO DE ACCESO.

\* “What should be obvious is that the detector type of knowledge is not what is distinctive of adult humans—or at least those of us who seek to be epistemically responsible agents who have examined our beliefs and aim at holding only those which after that examination are worthy of belief. From this perspective, unexamined beliefs are not yet worthy of believing. We seek to have good reasons for our beliefs.”

mejorada, su comprensividad, y su capacidad para satisfacer la curiosidad auto-reflexiva. El fiabilismo puro es cuestionable como una adecuada epistemología de tal conocimiento.\*\* (Sosa 1991, p. 95; cf. también Sosa 2007)

Esto ha llevado a filósofos con esta postura a formular epistemologías “bi-nivel”: para el conocimiento “animal” o “servo-mecánico” (que regularmente se identifica con el conocimiento obtenido a través de la mera percepción), se afirma que el fiabilismo externista basta para caracterizarlo; pero se afirma que una teoría internista se necesita para caracterizar al conocimiento “humano adulto” o “reflexivo” (en el caso de Sosa, se opta por el coherentismo; en el caso de Klein, por el infinitismo).<sup>101</sup> Así, estas teorías motivan una restricción de INTERNISMO DE ACCESO que resulta en:

(INTERNISMO DE ACCESO DÉBIL-RESTRINGIDO)

Para cierto tipo de contextos  $C$ : si  $x$  sabe  $\langle p \rangle$  en  $C$ , entonces  $x$  tiene acceso consciente a los factores *esenciales* que brindan tal conocimiento.<sup>102</sup>

Esta tesis es internista de acceso en el sentido obvio. Es *restringida* porque restringe el internismo a *ciertos* contextos. ¿A cuáles? Bueno, a los contextos en los que el conocimiento en cuestión es conocimiento “humano adulto” o “reflexivo”.

(Esto nos da, a su vez, una manera de entender la doble aparición de “muchos” en EXTERNISMO MÍNIMO: es para aquellas proposiciones que *no* necesiten ser caracterizadas como conocimiento “humano adulto”, para las cuales EXTERNISMO MÍNIMO afirma que vale la tesis externista (en su componente ontológico y atributivo). Si hay o no más proposiciones que estas, es una cuestión posterior.)

En fin: la tesis INTERNISMO DE ACCESO DÉBIL-RESTRINGIDO es *débil* porque no afirma que los factores que deben ser descritos por el internismo sean *todos* los factores que brindan tal conocimiento, sino que sólo a los esenciales. Hago esta acotación porque puede argumentarse que cuantificar sobre *todos* esos factores también resulta en una tesis altamente implausible (Pappas 2005).

Notemos que no es obvio qué tan defendible sea, ultimadamente, la clasificación del “bi-nivel” del conocimiento—quizá la separación entre conocimiento “animal” y “humano adulto” no termina siendo teóricamente interesante, recortando una juntura no-arbitraria sino natural en la realidad; quizá sí. Quizá el conocimiento “humano adulto” tampoco puede caracterizarse como

---

\*\* “Admittedly, there is a sense in which even a supermarket door “knows” when someone approaches, and in which a heating system “knows” when the temperature in a room rises above a certain setting. Such is “servo-mechanic” knowledge. And there is an immense variety of animal knowledge, instinctive or learned, which facilitates survival and flourishing in an astonishingly rich diversity of modes and environments. Human knowledge is on a higher plane of sophistication, however, precisely because of its enhanced coherence and comprehensiveness and its capacity to satisfy self-reflective curiosity. Pure reliabilism is questionable as an adequate epistemology for such knowledge.”

101 Un proyecto relacionado, pero distinto, es el de Foley (2002), quien cree que la teoría del conocimiento—como estado que conecta a la mente con el mundo—y la de la justificación—como propiedad que tienen las creencias de un agente que las ha revisado y cuidadosamente adquirido o modificado—deben separarse: para la primera, el externismo es la mejor idea; mientras que para la segunda lo es el internismo.

102 Véase Gibbons 2006. Esta no es la única manera de entender al internismo: algunos podrían entenderlo como la tesis de que los factores justificatorios dependan únicamente de eventos mentales (Connee & Feldman 2004), o como la tesis, análoga a una tesis discutida en filosofía de la mente, de que los factores justificatorios dependan únicamente de la naturaleza intrínseca del sujeto en cuestión (Nuccetelli 2003). Aquí trataré únicamente con INTERNISMO DE ACCESO.

conocimiento internista; quizá sí. Quizá, incluso, no tengamos tal conocimiento, si es que es como los internistas lo caracterizan, o tengamos mucho menos del intuitivamente esperado (cf. Kornblith 2002, cap. 4). Quizá no haya tales problemas. Aquí no me ocuparé de si tales teorías compatibilistas son, a final de cuentas, aceptables; sólo me interesa notar que EXTERNISMO MÍNIMO no las prohíbe—tratando con ello de eliminar una posible fuente de resistencia, desde el campo internista, contra tal tesis.

Ahora que hemos aislado una tesis internista plausible, la compatibilidad de INTERNISMO DE ACCESO DÉBIL-RESTRINGIDO y de EXTERNISMO MÍNIMO se muestra notando que las FIEs que son postuladas por la segunda tesis se toman como *prerrequisitos* de *todo* conocimiento, pero también como relevantes para la evaluación epistémica en *al menos algunos* casos (a saber, aquellos donde no se necesite describir al conocimiento como “reflexivo”). Esto es compatible con decir que *no en todos* los casos de evaluación epistémica evaluamos apelando a las características y funcionamiento de las FIEs: quizá, si algún bi-nivelismo resultara ser verdadero, hay ciertas dimensiones de evaluación para las que requerimos factores que van más allá de las FIEs. (Por eso el externismo en cuestión es *mínimo*.) Nuestra teoría de la justificación resultante aceptaría, entonces, que hay una dimensión *externista* en nuestra evaluación epistémica, pero que también existe una dimensión *internista*.

Con las restricciones y detalles relevantes puestos en su lugar, procedo a argumentar por EXTERNISMO MÍNIMO. Primero, a favor de la condición (a).

Necesitamos FIEs para tener creencias *sobre* el mundo. Pues supongamos que no tuviésemos ninguna FIE. Entonces toda la información que pudiera ser procesada por nuestro sistema cognitivo, si es que hubiera alguna, sería información *creada* por tal sistema—pues, por la suposición, nada de ella viene desde el ambiente (para ello, se necesitaría una FIE). Así, sin FIEs, podríamos tener información sólo sobre nuestra propia mente—siempre que tal información tampoco proviniera del ambiente, como los externistas semánticos argumentan que sucede en un rango importante de casos (Burge 1979; Putnam 1975; Williamson 2000: caps. 2, 3). Por lo tanto, necesitamos FIEs *si es que* queremos tener información acerca de otra cosa que nuestras propias mentes—o hasta un rango muy limitado de ellas.

Uno podría pensar que aún así seríamos capaces de tener una cantidad sustantiva de información—a saber, aquella que puede ser adquirida *a priori*. Pero esto también es muy dudoso. Para adquirir información *a priori*, regularmente se afirma, necesitamos primero poseer los conceptos necesarios. Y a menos que uno crea que *todos* ellos son innatos, para ello necesitaremos de la experiencia; por lo que también necesitaremos de FIEs. Por ejemplo, para tener información matemática—que típicamente se asume como *a priori*—requerimos poseer los conceptos necesarios. No podríamos hacer geometría analítica sin las nociones de *coordenada*, de *función trigonométrica*, o de *ecuación*; pero no es plausible suponer que tales nociones tan sofisticadas sean innatas (cf. §4.3, arriba). Por ello, aunque para *hacer matemáticas* no necesitemos, estrictamente hablando, *percibir* algún objeto o *realizar* algún experimento, sino que sólo necesitamos *razonar*, para razonar necesitamos conceptos desde los cuales parten tales razonamientos. Pero en buena parte de los casos, esos conceptos los adquirimos en la experiencia. Seguimos necesitando FIEs, incluso para adquirir información *a priori*.

Ahora argumentaré por la condición (c) en EXTERNISMO ONTOLÓGICO MÍNIMO, primero por la necesidad y luego la suficiencia.

El conocimiento requiere la obtención de información. Esto porque es universalmente aceptado que el conocimiento es un estado con *contenido* y el contenido en cuestión es información (que en filosofía se suele individuar como *proposiciones*). Por lo tanto, dado que para obtener información debemos tener FIEs, el conocimiento requiere FIEs: para representarnos al mundo, y representárnoslo con la estabilidad que el conocimiento requiere, necesitamos primero tener la capacidad de obtener información *acerca* del mundo. Por lo tanto, una teoría que no acepte FIEs será radicalmente escéptica. En esta tesis estoy asumiendo que el escepticismo es falso; dado esto, una epistemología aceptable debe aceptar nuestra posesión de FIEs.<sup>103</sup>

De manera converso, la obtención de información a través de las FIEs *debe* contar, al menos en un amplio rango de casos, como *suficiente* para el conocimiento. Pues supongamos que en ningún caso la obtención de información por las FIEs bastara para el conocimiento.<sup>104</sup> Si no aceptamos el escepticismo global, entonces algo además de las FIEs convierte esa información en conocimiento. Sea ese proceso o procesos, *G* (p. ej., razonamiento coherente, introspección, etc.).

Arriba argumenté que necesitamos información externa para representarnos al mundo—para tener creencias *sobre* el mundo, en particular. Pero, por asunción, *G* no es una FIE; así que o no dispone de información acerca del mundo, o lo hace y por ello depende de la información proporcionada por las FIEs. Si *G* no dispone de información acerca del mundo, ¿por qué su operación convierte información que sí lo es en conocimiento? ¿Cómo es que un procedimiento o facultad que no recibe información acerca del mundo puede ayudar a que la información acerca del mundo lo represente mejor? Del otro lado del dilema, suponer que *G* sí recibe información del mundo es suponer que recibe información de las FIEs. Pero entonces ¿por qué procesar información de las FIEs con más información de las FIEs produce conocimiento que no estaba ahí? Parece *ad hoc* postular que procesar información, mediante medios que sólo disponen de información que no se conoce, puede brindar conocimiento. Si no hay al menos una base de conocimiento desde la cual partir, es implausible pensar que *G* la pueda sacar de su propia operación.

Regresemos a la idea de conocimiento como representación estable. Si la información que usamos para representar al mundo nunca puede brindar, de por sí, representaciones estables, ¿cómo es que un mecanismo que, a final de cuentas, sólo tiene insumos de tal tipo de información, puede llegar a hacerla estable?

Por supuesto, mi argumento no convencerá a quienes no compartan la idea de que el conocimiento tiene que ver, esencialmente, con representar al mundo—por ejemplo, a quienes creen que el conocimiento tiene la responsabilidad del agente como condición necesaria (véase también la nota 94, arriba). Con este choque en las maneras iniciales al concebir el conocimiento, es poco probable que se llegue a muchos acuerdos posteriores. Lo único que podría decir es que quien piense que el conocimiento no se trata, fundamentalmente, de representar al mundo, me

103 Queda un pequeño detalle en el argumento: si algún tipo de idealismo radical es el caso, no necesitaríamos FIEs. Pues según el tipo de idealismo radical en cuestión, no conocemos otra cosa que los productos de nuestra mente—los objetos de conocimiento son constructos mentales. Así, este idealismo puede ser verdadero sin FIEs. Como en esta tesis también asumo la falsedad del idealismo radical, sigo sosteniendo la necesidad de las FIEs para nuestro conocimiento.

104 Uno podría sostener que la noción de *información* es fáctica, de tal manera que toda información es automáticamente verdadera. Esta es una cuestión que pasaré de largo en este texto. Por razones expositivas, sin embargo, hablo de la información como si fuera no fáctica. Las mismas afirmaciones podrían hacerse sin este compromiso, a costo de tener una redacción más rebuscada.

parece que o está cambiando de tema, o simplemente mezcla ideas de otros temas—la ética, por ejemplo—que pueden oscurecer el tema y prejuzgar cuestiones epistemológicas importantes. Me parece mejor idea tener una concepción relativamente estable de la representación que hacemos del mundo mediante el conocimiento, para después ver cómo interactúa con conceptos de otras áreas—la filosofía social o la ética, por ejemplo.

Si aceptamos los argumentos de arriba, EXTERNISMO MÍNIMO surge como una tesis sobre los *pre-requisitos* de nuestro conocimiento: para que podamos tener cualquier conocimiento, tiene que ser verdad cierto tipo de externismo ontológico.

He argumentado por (a) y (c) en EXTERNISMO ONTOLÓGICO MÍNIMO. Para (b), notemos cómo es que las FIEs operan sin que nosotros tengamos acceso a ellas. Supongamos, para propósitos del argumento, que nuestra capacidad de la visión es una FIE. Nadie cree que para obtener conocimiento visual, primero tengamos que saber cómo opera nuestro sistema nervioso, nuestro par de ojos, etcétera. Si así fuera ¡nunca tendríamos tal conocimiento! Pues mucho de lo que sabemos en la ciencia de la visión depende—de conocimiento obtenido a partir de la visión. Muchos experimentos y teorías sólo pueden verificarse o diseñarse utilizando la visión. Si se requiriera que, para conocerlas, supiéramos primero cómo funciona la visión, caeríamos en una imposibilidad de principio. Por lo tanto, al menos la FIE que es la capacidad de la visión nos brinda conocimiento sin que nosotros tengamos acceso consciente a su manera de operar.

Por supuesto, alguien con intuiciones internistas puede replicar que esto nunca fue lo que quiso decir. “No es que”, afirmaría, “necesitemos saber cómo funciona la visión, *antes* de poder obtener conocimiento sobre la base de su funcionamiento. La idea es sólo que necesitamos tener evidencia (o coherencia, o infinitas razones disponibles, o probabilidades subjetivas antecedentes altas—esto depende de qué teoría internista uno acepte) de que tal facultad *de hecho* nos brinda conocimiento, y debemos ser capaces de acceder conscientemente a tal evidencia.” Si este es el esquema de objeción que un internista tiene en contra de la cláusula (c) de EXTERNISMO ONTOLÓGICO MÍNIMO, creo que está equivocada—y por razones parecidas a la del párrafo anterior (de nuevo, véase también Pappas 2005, §§2-3; Van Cleve 2003). Que la evidencia (o coherencia, o infinitas razones disponibles, o probabilidades subjetivas antecedentes altas) deba poder ser accesible a la mente, presupone que tengamos información acerca del mundo—incluyendo en este, por supuesto, a nuestra misma mente (que, *pace* idealismos, es parte del mundo también). Pues sin ella, tal evidencia (coherencia...), si es que resultara que existe alguna, sería una mera invención de nuestra muy creativa mente. Pero ¿cómo es que tal evidencia (etc.) es un pre-requisito para el conocimiento, si no tiene nada que ver con lo que sucede en el mundo, mientras que el conocimiento es un estado que se da sólo si su contenido es verdad—*i.e.*, ocurre en el mundo? No podemos requerir evidencia (coherencia, etc.) si no aceptamos que hay FIEs—que puedan transformar la información obtenida del ambiente en contenidos en nuestra mente, a los que *después* podemos asignarle probabilidades subjetivas, podemos organizar coherentemente, podemos tomar como razones o como evidencia. Pero, como vimos arriba, tampoco podemos requerir que sepamos acerca del funcionamiento de las FIEs *antes* de poder obtener conocimiento basado en sus insumos.

Concluyo, como dice (b) de EXTERNISMO ONTOLÓGICO MÍNIMO, que debemos aceptar que las FIEs pueden transportar información del mundo a la mente sin que nosotros seamos conscientes de ello o de cómo lo hacen—y aún así brindarnos conocimiento con ello.

Habiendo argumentado por EXTERNISMO ONTOLÓGICO MÍNIMO, argumentaré ahora por EXTERNISMO ATRIBUTIVO MÍNIMO. Argumento desde ejemplos.

Solemos creer menos en el testimonio de nuestros sentidos cuando sabemos que no están funcionando adecuadamente. En la oscuridad de la noche, por ejemplo, tenemos menor confianza en nuestros juicios visuales; así como creemos menos en los juicios sobre lo que hemos escuchado decir a una persona cuando estamos en un ambiente ruidoso. También creemos menos en los juicios de una persona intoxicada, pues asumimos que sus facultades no están funcionando con toda su capacidad. Por otro lado, confiamos más en los juicios de quienes han entrenado sus capacidades: creemos más lo que nos diga un director de orquesta sobre cuál es la nota que acaba de ser tocada, por ejemplo, que lo que nos diga un escucha sin entrenamiento.

Estos ejemplos de la vida cotidiana muestran que nuestras consideraciones sobre las FIEs—tanto las propias como las de otras personas—juegan un papel importante al evaluar epistémicamente las creencias formadas acerca de la información que esas FIEs transmiten. Esto *no* es decir que el funcionamiento de las FIEs es *todo* lo que consideramos *siempre* que atribuimos conocimiento—es decir que es *uno* de los factores que consideramos, y uno importante entre los demás, *al menos* cuando no necesitamos atribuir conocimiento “humano adulto”. Además, con los ejemplos también se muestra que, al menos en casos con un parecido relevante a los ejemplos, no consideramos otra cosa que el funcionamiento de las FIEs. No nos preguntamos si los agentes en cuestión han reflexionado sobre su funcionamiento, o tienen un sistema coherente de creencias de acuerdo al cual las FIEs sí funcionan adecuadamente. Simplemente les atribuimos conocimiento asumiendo que sus FIEs funcionan bien.

Así mostramos EXTERNISMO ATRIBUTIVO MÍNIMO. Como expuse en la sección anterior, individualizamos nuestras prácticas atributivas desde las intuiciones generadas por ejemplos.

¿Qué criterio utilizamos para evaluar el funcionamiento de las FIEs? Los ejemplos dados muestran que, al menos en esos casos y los que relevantemente parecidos, evaluamos a las FIEs por su *fiabilidad*: en los casos mencionados, creemos más la información que nos es entregada por FIEs que o se encuentran en un *ambiente propicio* o *están entrenadas*. En general, entre más razones tenemos para creer que una FIE nos está entregando más información verdadera que información falsa sobre el ambiente, solemos asignarle mayor grado de creencia a tal información.

He defendido EXTERNISMO MÍNIMO, y en lo que resta de la tesis asumiré su verdad. En el siguiente capítulo, ofreceré una teoría que busca explicar cómo podemos tener conocimiento modal, bajo los constreñimientos que EXTERNISMO MÍNIMO pone.

Como noté arriba, no he dicho cómo es que concibo al conocimiento. En breve, lo concibo bajo un modelo fiabilista—bajo el cual una capacidad de poder representar nos brinda conocimiento cuando es *fiable*. El sentido exacto de la fiabilidad es el tema del capítulo sexto y último, pero la idea intuitiva es clara: una habilidad, facultad, método, etc., nos brinda conocimiento (al menos el conocimiento sujeto a EXTERNISMO MÍNIMO) cuando de manera regular y estable produce representaciones que son *fieles* a la realidad. La facultad de la visión nos brinda conocimiento porque produce, de manera estable, representaciones (creencias perceptuales) que son fieles a la realidad (que son verdaderas). Los métodos de la ciencia nos brindan conocimiento porque de manera estable producen representaciones (modelos y teorías) que son fieles a la realidad—que son, si no siempre verdaderas, al menos aproximadamente verdaderas. Que una representación sea fiel a la realidad no siempre es accesible desde nuestra mente, pero aún así se necesita y, en muchos ca-



sos, basta, para producir y atribuir conocimiento—en total consonancia con EXTERNISMO MÍNIMO.

En la siguiente sección me concentraré en un importante problema para el fiabilismo que acepta EXTERNISMO MÍNIMO y que asumiré en el siguiente capítulo: el *problema de la generalidad*. Una vez que hemos rechazado este problema, libramos al fiabilismo de una de las piedras más grandes que bloquean su aceptación. No lo libramos de todas, por supuesto. Pero hacemos un avance importante.

### 5.3. el fiabilismo

Como mencioné arriba, la motivación intuitiva para el fiabilismo es clara: *una habilidad, facultad, método, etc., nos brinda conocimiento* (al menos el del tipo sujeto a EXTERNISMO MÍNIMO) *cuando de manera regular y estable produce representaciones que son fieles a la realidad*.

Esta formulación del fiabilismo es algo más general que la de sus proponentes más famosos (e.g., Goldman 1979, 1986), quienes suelen restringirse a procesos productores de *creencias*, mientras que yo hablo de *representaciones* en general. No me parece claro que tal restricción sea necesaria. No es claro que la creencia sea el único tipo de estado mental representacional, y sí es claramente falso que sea la única manera de representarnos al mundo que poseemos. Representamos al mundo—que, por supuesto, nos incluye a nosotros y a nuestra mente—además de con creencias, con lenguajes, con imágenes, con sonidos, con banderas, con símbolos, con teorías, con modelos tridimensionales (como una maqueta), o con gestos. Aunque quizá no todas, muchas de estas maneras de representarnos pueden evaluarse en cuanto a su éxito representacional: así como una creencia es exitosa cuando es verdadera, una maqueta de un edificio es exitosa cuando lo representa de manera suficientemente detallada, o un nombre es exitoso cuando refiere a su referente intencional, o un gesto es exitoso cuando representa un estado mental.

Esto me lleva a dudar de que sea obvio que *sólo* las creencias, de entre las demás maneras de representarnos al mundo con que contamos, sean las que pueden brindarnos *conocimiento*—o estados perfectamente análogos a este. Por ejemplo, Moss (2013) argumenta que las *credences*—estados doxásticos graduales—analizadas por la tradición bayesiana, también pueden constituir conocimiento—y están sujetas a muchas de las tesis y discusiones de la epistemología tradicional. Así, por ejemplo, quizá una teoría científica, cuyo contenido se puede concebir como el conjunto de sus modelos (como en la tradición estructuralista) pueda constituir conocimiento.

El segundo aspecto en que mi formulación es más general, consiste en hablar de la “*fidelidad*” a la realidad que tiene una representación. La función de tal término es esta: así como el fiabilismo tradicional habla de la *verdad* de las creencias, si aceptamos que otros estados o representaciones además de las creencias pueden ser conocimiento, no es claro que siempre podamos hablar de *verdad* para tales estados o representaciones (por ejemplo, Moss argumenta que ciertas *credences* estrictamente *falsas* pueden constituir conocimiento). Esto es porque el criterio de corrección o adecuación de algunas representaciones no es la verdad—por ejemplo, un nombre es “correcto” cuando se refiere a su referente intencional, y un modelo probabilístico es “adecuado” cuando describe una situación real con cierto grado de idealización, de la misma manera en que una creencia es “correcta” o “adecuada” cuando es verdadera. En los tres casos, estamos ha-

blando de representaciones que son, en un sentido muy intuitivo, *fieles* a la realidad—que la describen bien, que la representan, que apegan su estructura a lo que la realidad es.

Pero incluso si sólo las creencias pueden constituir *conocimiento*, todavía es posible argumentar que otro tipo de representaciones—mapas, *credences*, nombres—tienen sus propios criterios de *adecuación*, y que pueden constituir estados *análogos* al conocimiento para las creencias: *una representación que de manera estable es adecuada* (o “fiel a la realidad”). Quizá, por ejemplo, la rigidez de un nombre es una propiedad *análoga* a la propiedad de las creencias de constituir conocimiento. Por ejemplo, Jenkins (2008a) argumenta que la propiedad análoga para los conceptos es una que ella llama *estar fundamentados* (ser “*grounded concepts*”). Creo que estos avances en generalidad desde el punto de vista fiabilista pueden ser útiles y teóricamente interesantes. De hecho, es dable conjeturar que, a largo plazo, pueden comenzar a brindarnos una explicación epistemológica más general de nuestra vida representacional en términos fiabilistas—que ya no se quede en el enfoque sobre las creencias o, en menor medida, nombres y conceptos. (Hablaré más de todo esto en el siguiente capítulo.)

He caracterizado intuitivamente—y en una versión general—al fiabilismo. Ahora haré un poco de “ingeniería inversa”, buscando los principios de los cuales puede extraerse el fiabilismo. Como también haré algo de historia intelectual, viendo cómo es que Goldman y otros llegaron al fiabilismo, mucha de la siguiente exposición es acerca del fiabilismo de proceso *acerca de la justificación doxástica*, pero creo que las notas que he argumentado arriba dan pistas para traducir tal discusión al nivel más general con el que comencé esta sección. Veamos.

En su primera examinación del fiabilismo procesual en 1979, Goldman se propuso dar una teoría *descriptiva* (como opuesta a una *prescriptiva*) de la justificación. Por una “teoría de la justificación”, Goldman entendía “un conjunto de principios que especifiquen condiciones de verdad para el esquema ‘La creencia de *S* de que *p* en un momento *t* está justificada’” (Goldman 1979: p. 107).<sup>105, 106</sup>

105 Las referencias a 1979 son hacia la reimpresión como el capítulo 6 de Goldman 1992. Véase también Goldman 1986, especialmente los capítulos 3 y 4.

106 En esta discusión temprana, Goldman asumió que tales condiciones de verdad deberían ser “condiciones para la satisfacción de este esquema en todos los casos posibles” [“conditions for the satisfaction of this schema *in all possible cases*”] (*ibid.*, mi énfasis), pero no es claro qué tipo de posibilidad tenía en mente. Parecería, por ejemplo, que la posibilidad metafísica impondría aquí un constreñimiento bastante fuerte sobre toda teoría de la justificación: se requeriría que el esquema en cuestión dictara lo que tendría que pasar en todo mundo donde haya justificación epistémica—independientemente de si esa justificación es poseída por humanos, robots, alienígenas, delfines ultra-desarrollados, o dioses. Si esto fuera así, tendríamos un caso claro donde los filósofos exigen algo al menos muy cercano a la *esencia* del fenómeno en cuestión: dado que en los años ‘70 y en los ‘80 todavía era común identificar lo esencial a *x* con lo que le es necesario, este requisito sobre las teorías de la justificación equivaldría a exigirles que identificaran la *esencia* de la justificación epistémica.

En trabajo posterior, Goldman hizo uso de la idea de los “mundos normales”, cuya extensión está determinada por nuestras creencias acerca de qué posibilidades son “realistas”; la idea es que nuestro concepto de justificación se aplica sólo en esos mundos (Goldman 1986, secc. 5.5). Así, uno podría pensar que la clase de mundos involucrados en la posibilidad de arriba serían sólo los mundos normales, no todos (y quizá no sólo) los mundos metafísicamente posibles. Sin embargo, después Goldman mismo criticó este acercamiento, y en otros trabajos distinguió entre una teoría de la *atribución* de justificación, y una teoría sobre la *naturaleza* de tal justificación (a su vez, esta última podría separarse en dos: una teoría sobre tal naturaleza *de acuerdo a nuestros conceptos folk*, y una teoría sobre tal naturaleza *de acuerdo a la mejor epistemología disponible*, que se permite refinar a nuestros conceptos *folk* con la ciencia cognitiva). Véase Goldman 1992 y 2008.

Aparte de la condición de descriptividad, él limitó el espacio de las teorías admisibles con otras dos asunciones. Las tres asunciones, en total, son las siguientes:

(DESCRIPTIVA)

Una teoría de la justificación adecuada debería tratar de *explicar* nuestros estándares ordinarios de justificación, a diferencia de *prescribir* estándares que mejoren a nuestros estándares ordinarios.<sup>107</sup>

(TÉRMINOS NO EPISTÉMICOS)

La teoría de la justificación adecuada debe ser dada sólo en términos que no sean epistémicos.

(EXPLICATIVIDAD)

La teoría de la justificación adecuada debe enunciar condiciones necesarias y suficientes para la justificación, y tales condiciones también deben ser apropiadamente *profundas* o *revelatorias*.

Es plausible que se pretendía correlacionar al contenido de DESCRIPTIVA con el contenido de los otros dos constreñimientos. Pues una teoría perfectamente descriptiva acerca de la justificación podría simplemente describir las circunstancias en, y los mecanismos lingüísticos por los cuales, usamos el término “justificación” (y otros términos del mismo campo semántico); pero eso no es lo que se quiere aquí. Más bien, como dijo Goldman, lo que se busca es “una teoría explicativa, una que explique en una manera general *por qué* ciertas creencias cuentan como justificadas y otras como injustificadas” (*ibid.*; mi énfasis). Así, EXPLICATIVIDAD se invoca para requerir que nuestra teoría descriptiva haga algo más que simplemente describir cómo solemos utilizar “justificación”: nuestra teoría también debería explicar *por qué* hacemos tal uso. Y se invoca a TÉRMINOS NO-EPISTÉMICOS, al parecer, porque una teoría puesta en términos epistémicos no explicaría tales términos—tal teoría, parece estarse asumiendo, sólo estaría presuponiendo lo que se necesita explicar.

Pero no es claro qué significan exactamente “profunda” o “revelatoria” aquí, ni por qué es que una teoría formulada usando términos epistémicos no sería lo suficientemente “profunda” o “revelatoria” como para merecer atención. Para ver esto último, consideremos una definición funcional de los términos epistémicos, al estilo de la definición funcional de los estados mentales que Lewis ofreció (1966; 1970; 1994), o, para un caso más cercano al nuestro, la definición funcional de los términos morales de Jackson y Pettit (1995). Uno definiría una oración de Ramsey,  $(\exists x_1, \dots, x_n)(\Phi[x_1, \dots, x_n])$ , donde  $\Phi$  especifica cómo cada propiedad epistémica se relaciona con otras propiedades epistémicas y quizá otros estados no-epistémicos (como el creer con verdad

---

La hipótesis de trabajo en este capítulo es que los epistemólogos sí buscan una teoría que describa la *esencia* de la justificación (o, en su caso, del *conocimiento*). Quizá, como mencioné al inicio del capítulo, no nos acerquemos al conocimiento tratando de analizarlo—pero todavía podemos, al caracterizarlo, mostrar cómo es que su esencia se relaciona con otras propiedades y estados epistemológicamente relevantes.

107 Esta condición es suplementada en Goldman 1992, en donde Goldman cree que, además de una teoría completamente descriptiva, debe existir una teoría prescriptiva, guiada por la ciencia cognitiva. Pero el problema de la generalidad surge desde la primera tarea, la descriptiva. Véase la siguiente sección de esta tesis.

\* “an explanatory theory, one that explains in a general way *why* certain beliefs are counted as justified and others as unjustified”

que  $p$ , o el tener la intención de que  $q$ ). Tal definición funcional todavía invoca lenguaje epistémico, pero no es claro cómo es que falla en dar una teoría suficientemente explicativa de él.

Sea como sea, así es como Goldman fraseó sus requisitos, y motivó al fiabilismo de proceso. En trabajo posterior, el fiabilismo también fue motivado por lo que Goldman después desarrolló en una postura sobre el valor epistémico: el *veritismo*.<sup>108</sup> De acuerdo a esta idea, un—quizá *el*—valor epistémico *fundamental* (en el sentido de no deber lo valorable que tenga a otros valores epistémicos) es la adquisición de creencias verdaderas o, al menos, de creencias verdaderas sobre los temas que nos interesan. Dada esta tesis, argumentó Goldman, necesitamos una teoría *consecuencialista* de la justificación epistémica, de acuerdo a la cual los procesos, metodologías o heurísticas de producción de creencias, tienen su valor justificatorio de manera derivada por el cuánto ayudan a promover la obtención de ese valor epistémico. Esto motiva la siguiente definición (asumiendo que las creencias vienen en grados):

(VALOR VERITISTA)

Una creencia a un grado  $g$  ( $g \in [0, 1]$ ) en una proposición verdadera  $p$  tiene valor veritista  $g$ . Una creencia a un grado  $g$  en una proposición falsa tiene valor veritista  $1-g$ .

Y entonces, con una formulación simple de la fiabilidad,

(FIABILIDAD SIMPLE)

Un proceso de producción de creencias  $P$  es fiable sii  $P$  tiende a producir *más* creencias verdaderas que falsas,

podríamos generalizar desde creencias verdaderas a valor veritista, si asumimos:

(FIABILIDAD VERITISTA SIMPLE)

Un proceso de producción de creencias  $P$  es fiable sii  $P$  tiende a producir *suficiente* valor veritista.

Esto, por supuesto, es todavía un bosquejo. Ahora le daremos un poco de más detalle.

Empecemos dejando de lado la vaguedad de “más” y de “suficiente”: plausiblemente, *hay* vaguedad en qué tan “exitosa” puede ser cualquier cosa para ser fiable en una dimensión apropiada—específicamente, parece ser vago por *cuánto más* la producción de creencias verdaderas por un proceso tiene que exceder la producción de creencias falsas por ese proceso, si es que ha de contar como fiable. Uno podría afirmar que tal vaguedad ha de ser *eliminada* contextualmente: los contextos de uso de la palabra “fiable” fijan el estándar que un proceso (método, etc.) tiene que satisfacer si es que ha de contar como fiable. O, de manera más realista: el espacio determinado por los límites (*borderlines*) ha de ser *reducido* contextualmente: los contextos de uso de la palabra “fiable” fijan una clase relativamente pequeña de estándares admisibles, al menos uno de los cuales el proceso tiene que satisfacer para contar como fiable (*cf.* Goldman 1979: 114; Heller 1995).

Lo importante aquí es que, por más que “fiable” esté sujeto a vaguedad, es claro que *no* es un adjetivo *fáctico*, en el sentido en que cualquier cosa que sea fiable es por lo tanto exitosa. (En nuestro caso, para cualquier creencia  $c$ , que  $c$  sea producto de un proceso fiable no implica que  $c$  sea verdadera.) Como se indicó arriba con las definiciones de fiabilidad simple, la fiabilidad es una *tendencia*, y por ello un proceso fiable todavía podría, en algunos casos, no regresarnos una creencia verdadera (o, más generalmente, suficiente valor veritista). Pero esto, plausiblemente,

<sup>108</sup>Goldman 1986; 2001; 1999: cap. . 5; Goldman & Olsson 2009.

es una consecuencia bienvenida—porque entonces nuestra teoría resultante satisfará:

(FALIBILISMO)

Hay creencias justificadas con contenido falso.

¿Por qué tomar a FALIBILISMO como un principio aceptable? Bueno, porque parece que su negación es implausible: es la idea de que *sólo* las creencias verdaderas están justificadas. Pues parece que podemos evaluar una creencia como justificada por más que (en la metáfora común) “el mundo no coopere”. Un científico, haciendo el mejor de sus esfuerzos, tomando en cuenta toda la evidencia disponible que le es accesible, y siendo intelectualmente honesto, puede aún así llegar a una creencia falsa. Pero aún así, el científico estaba *justificado* en creerla—hizo los mejores razonamientos que pudo, consideró la evidencia, su aparato cognitivo y la tecnología que usaba estaban funcionando bien, etc.. Simplemente el mundo no siempre se apega a lo que creemos de él. Pero aún así podemos evaluar positivamente esas creencias que no lo describen; y para ello usamos el término “justificación”. Por estas razones se concluye FALIBILISMO.

Todavía tenemos que dar cuenta de qué es que un proceso *tienda* a producir alguna cosa. Goldman asumió que lo que se pretende es que el proceso tenga *objetivamente* la tendencia relevante, no como cuestión de las probabilidades subjetivas de ningún agente.

(OBJETIVISMO)

La noción de fiabilidad en FIABILIDAD (VERITISTA) SIMPLE es una noción de probabilidad *objetiva*.

(Hay que notar que esta asunción es mucho más clara si se asume para una teoría *ontológica*—y no una atributiva—de la justificación.)

Ahora bien, dado OBJETIVISMO, Goldman considera brevemente dos propuestas para dar cuenta de la noción de fiabilidad en cuestión (1979, p. 114.):

(PROPENSIÓN)

La noción de fiabilidad de un proceso se puede entender como la *propensión* del proceso de producir más creencias verdaderas que falsas.

(FRECUENCIA REAL A LARGO PLAZO)

La noción de fiabilidad de un proceso se puede entender como la *frecuencia* real a largo plazo de producir más creencias verdaderas que falsas.

Por supuesto, estas dos propuestas están modeladas bajo dos diferentes y populares interpretaciones de la teoría matemática de la probabilidad: el frecuentismo y la teoría de propensiones (*cf.* Háyek 2007). De acuerdo a Goldman y plausiblemente, nuestra concepción común de la justificación no es suficientemente determinada como para decidir entre estas opciones; una decisión entre ellas será altamente teórica.

Con todo esto, la primera propuesta de Goldman resultaba en:

(FIABILISMO DE PROCESO-VERSIÓN SIMPLE)

Una creencia *b* está justificada sólo si *b* es producida por un proceso *P* que es fiable.<sup>109</sup>

<sup>109</sup>Goldman mismo notó contraejemplos inmediatos a esta tesis, y propuso bloquearlos agregando condiciones extra (como condiciones que eliminaban derrotadores, “*defeaters*”). En esta tesis no voy a considerar tales consideraciones extra, sino que me quedaré con el “fiabilismo crudo” que se enfoca sólo en la fiabilidad de los procesos de producción de creencias. La razón de ello es que las cuestiones son ya de por sí complicadas sin consideraciones extra, y además que asumo que tales consideraciones son una tarea para todo fiabilista, que yo puedo adi-

Ahora ya podemos ver cómo es que FIABILISMO DE PROCESO-VERSIÓN SIMPLE satisface los requisitos de Goldman para una teoría de la justificación: satisface DESCRIPTIVA porque correlaciona nuestros juicios epistémicos con la fiabilidad de los procesos, mientras que se abstiene de brindar un juicio normativo acerca de si así es, o no, como debería hacerse; satisface TÉRMINOS NO-EPISTÉMICOS porque la teoría de la fiabilidad de proceso no necesita ponerse en un lenguaje que contenga ideología epistémica; satisface EXPLICATIVIDAD porque las condiciones propuestas (plausiblemente) son suficientemente explicativas; y satisface FALIBILISMO porque se asume que la fiabilidad implica la justificación, pero la fiabilidad de un proceso es compatible con que produzca también algunas creencias falsas.

Todo estaría bien hasta ahora, si no fuera porque FIABILIDAD (VERITISTA) SIMPLE no nos dice qué es un “proceso de producción de creencias”. Si no lo hace, nuestro entendimiento de FIABILIDAD (VERITISTA) SIMPLE no puede ser total—a uno de sus términos centrales todavía no se le ha asignado contenido definido. Y aunque, trivialmente, sabemos que un proceso de producción de creencias es un *proceso* que termina en una *creencia*, y aunque quizá sabemos que un proceso es una secuencia de eventos con una conexión causal apropiada, esto no es suficiente. Lo que FIABILISMO DE PROCESO-VERSIÓN SIMPLE necesita es una teoría de los procesos tal que, al menos, las tendencias de esos procesos *no sean triviales*: no deben ser siempre 0 o 1. Si todo proceso tuviera fiabilidad 0, entonces ninguna creencia estaría justificada. Pero intuitivamente algunas creencias forman conocimiento, así que DESCRIPTIVA nos obliga a rechazar la idea de que todo proceso tiene fiabilidad 0. Y si todo proceso tiene fiabilidad 1, entonces toda creencia está justificada, pero eso de nuevo va contra la intuición y DESCRIPTIVA termina igualmente bloqueando la idea. Así, el requisito DESCRIPTIVA nos obliga a postular fiabilidades no triviales.

Pero supongamos que quisiéramos rechazar DESCRIPTIVA. Entonces no necesitaríamos buscar una teoría que se limitara por lo que nuestras intuiciones nos dicen, y los perfiles de fiabilidad de los procesos productores de creencias no nos dirían mucho acerca de lo que intuitivamente llamamos “justificación”. La única manera de correlacionar tales perfiles de fiabilidad con la justificación intuitivamente caracterizada sería hacerlo con fuerza normativa (“la justificación es gradual, pero no la fiabilidad; entonces ¡debemos *corregir* nuestra idea *folk* de la justificación!”). Pero será mejor que resulte que tal normatividad no sea epistémica en ningún sentido razonable—porque si lo es, TÉRMINOS NO-EPISTÉMICOS también tendría que irse (la opción que permanece parece ser el tipo de descripción funcional considerada antes, pero no es claro que Goldman alguna vez haya considerado favorablemente tal proyecto). Y si no es epistémica, no es claro a qué tipo de normatividad se podría apelar para correlacionar justificación con fiabilidad.

Pero supongamos que la justificación no fuera explicada en términos de la fiabilidad, conservando DESCRIPTIVA: entonces dejaríamos ir a EXPLICATIVIDAD, de tal manera que la teoría resultante se apegaría a la descripción intuitiva que hacemos de la justificación, pero sin buscar *explicarla* en el sentido relevante.

Dado lo inestable que se vuelve la postura si no se asume que la fiabilidad en cuestión debe poder estar entre 0 y 1, debemos aparejarla con una teoría de los procesos de producción de creencias que les permita tener perfiles de fiabilidad no-triviales:

(FIABILIDAD NO-TRIVIAL)

Los procesos en el dominio de  $P$  en FIABILISMO DE PROCESO-VERSIÓN SIMPLE deben ser capaces

---

cionar a la teoría resultante de lo que viene abajo.

de tener un perfil de fiabilidad que esté estrictamente entre 0 y 1.

Ahora bien, FIABILIDAD NO-TRIVIAL *no* es satisfacible por un *token* de un proceso del mundo real. Un proceso deja de existir cuando alcanza su resultado; por lo tanto no hay manera en la que un proceso de producción de creencias podría tener un número de fiabilidad diferente al 1 (si la creencia que produjo tenía contenido verdadero) o 0 (si la creencia que produjo tenía contenido falso). Lo que necesitamos son procesos “extendidos” o “durables”—que puedan tener más de una creencia como resultado, y con ello que puedan tener números de fiabilidad no triviales (bajo FIABILIDAD (VERITISTA) SIMPLE).

Hay dos maneras de encontrar tales procesos con fiabilidades no triviales que inmediatamente vienen a la mente. Cada una corresponde, a su vez, con PROPENSIÓN O FRECUENCIA REAL A LARGO PLAZO, consideradas arriba:

(MODAL)

Cada proceso  $P$  existe en más de un mundo posible. El perfil de fiabilidad de  $P$  se define en términos de su resultado en cada uno de los mundos posibles en el cual existe.

(TIPOS)

Cada proceso  $P$  cae al menos bajo un *tipo* de proceso. El perfil de fiabilidad de  $P$  se define en términos de los resultados de los procesos que caen bajo el tipo de  $P$ .

Veamos MODAL. Si los procesos se extienden a través de mundos posibles como lo dice esta teoría, entonces, si es que alguna vez hemos de saber qué perfil de fiabilidad tiene *este* proceso, tenemos que saber cómo “mantener nuestro dedo apuntado” a él en otros mundos posibles. Este es, por supuesto, el famoso “problema de la individuación trans-mundo”: asumir que una y la misma cosa existe en más de un mundo posible, con diferentes propiedades en cada uno de ellos, nos trae la carga explicativa de cómo explicar que una y la misma cosa pueda existir con dos conjuntos distintos (e incluso *incompatibles*) de propiedades—dado que, por supuesto, una y la misma cosa no puede tener propiedades distintas o incompatibles con las que ella tiene.

Una solución es invocar al *esencialismo*—según el cual la individuación trans-mundo se basa en las esencias:  $x$  (existiendo en  $w$ ) e  $y$  (en  $v$ ) son la misma cosa si la esencia de  $x$  = la esencia de  $y$ . Así, una y la misma cosa puede tener diferentes propiedades porque las tiene *relativas* a mundos, de igual manera en que una y la misma cosa puede tener diferentes propiedades porque las tiene en momentos distintos. Y es la *misma* cosa la que tiene diferentes propiedades (relativo a mundos distintos), porque es la misma *esencia* la que está presente.

Supongamos que los procesos tuvieran esencias individuales cualitativas: para cada proceso  $P$ , hay cualidades  $Q_1, \dots, Q_n$ , necesarias y suficientes para la existencia de  $P$ —para cada mundo  $w$ , lo que sea que ejemplifique  $Q_1, \dots, Q_n$  en  $w$  es, por ello,  $P$ . Entonces, para seguir a  $P$  a través de los mundos, sólo necesitamos seguir a  $Q_1, \dots, Q_n$ —y aquellas cosas que ejemplifiquen tales cualidades serán, según MODAL, las que determinen la fiabilidad de  $P$  (Cf. Heller 1995, nota 7, p. 515.).

Hay algunos problemas aquí. En primera, ha resultado difícil mostrar que los objetos *tienen* esencias cualitativas. En segunda, ha resultado difícil, incluso asumiendo que los objetos tienen esencias (cualitativas o no), mostrar cómo podemos llegar a conocer tales esencias. En tercera, incluso asumiendo, en particular, que los procesos tienen esencias cualitativas, el problema es identificar tales esencias de una manera *principada* y que tenga como resultado que los procesos

tienen un perfil de fiabilidad no-trivial.

Estos problemas pueden resultar no ser fatales. Por ejemplo, no es obvio que no podamos tener conocimiento de las esencias cualitativas de los procesos. Pero, por el momento, todavía necesitamos más para saber cómo podríamos aplicar MODAL (bajo esta interpretación esencialista) para entender la postura fiabilista.

Una segunda solución para el problema de identificación es invocar a la teoría de contrapartes de Lewis (Lewis 1986, cap. 4). (Re-)interpretaríamos MODAL como afirmando que el perfil de fiabilidad de un proceso  $P$  se define en términos de los resultados de las *contrapartes* de  $P$ —por dar una teoría simple, por ejemplo, el perfil de fiabilidad de  $P$  sería el resultado de promediar el número de creencias verdaderas producidas por las contrapartes de  $P$ , dado el número total de creencias resultado de las contrapartes de  $P$ . (Por supuesto, quizá tales contrapartes podrían ser infinitas, pero por simplicidad estoy asumiendo que tendríamos a la mano alguna herramienta estadística para lidiar con esos casos.)

Ahora, como la noción de contraparte es dependiente de contexto (Lewis, *ibid.*), esto haría que la fiabilidad sea dependiente de contexto: un proceso sería fiable sólo de manera relativa a un contexto  $C$ : si, en  $C$ , las contrapartes de ese proceso resultan en más casos en creencias verdaderas que falsas, el proceso es fiable; pero en otro contexto  $C^*$  el *mismo* proceso puede tener diferentes contrapartes, por lo que puede dejar de ser fiable. Así, si la fiabilidad es al menos condición necesaria para la justificación, entonces dos creencias producto del *mismo* proceso, pueden no estar justificadas—dependiendo del contexto en el que sean evaluadas.

Aquí no evaluaré qué tan plausible sea este tipo de contextualismo fiabilista resultante (*cf.* Heller 1995). Pero notemos que, de todos modos, una variante del problema parece seguir aquí. Pues qué relación de contraparte sea la que esté en juego en un contexto depende, por supuesto, de las presuposiciones de los hablantes en ese contexto. Así, cuáles sean las propiedades modales de un proceso a las que nos referimos en un contexto, dependerá de la información que en el contexto en cuestión está en juego. Pero, ¿cómo es que esa información nos ayuda a individuar procesos? Esta es la misma pregunta que venimos presionando desde arriba. Tanto la interpretación esencialista como la de la teoría de las contrapartes sólo nos ayudan a rephrasearla, no a responderla.

Tomemos entonces TIPOS. La manera en que se frasea—“El perfil de fiabilidad de  $P$  es dado con referencia a los resultados de los procesos que caen bajo el tipo de  $P$ ”—requiere que  $P$  caiga bajo un *único* tipo, o al menos uno *distinguido*. Es decir, necesitamos un tipo tal que *ése* tipo único (o distinguido) determine la fiabilidad de  $P$  de la manera requerida. Pues supongamos que no hubiese tal tipo saliente:  $P$  cae bajo los tipos  $\underline{P}_1, \dots, \underline{P}_n$ , y ninguno de ellos,  $\underline{P}_i$ , es distinguido en el sentido de que la fiabilidad de  $\underline{P}_i$  determine la fiabilidad de  $P$ . Entonces, no hay garantía de que lo siguiente nunca suceda: para  $j \neq k$ , como un *token* de  $\underline{P}_j$ , la fiabilidad de  $P$  es un cierto número; como un *token* de  $\underline{P}_k$ , su fiabilidad es otro número. Además de todo, no hay garantía de que los números de fiabilidad resultantes no diverjan ampliamente, por lo que tomar alguna estadística (como la *moda* o el *promedio*) sea poco informativo acerca de la fiabilidad del proceso en cuestión.<sup>110</sup> Veamos un ejemplo.

110 Esta formulación hace particularmente claro que lo que abajo consideraremos como el problema de la generalidad, es un caso del problema de la clase de referencia (Háyek 2007b). En esa medida, una solución al problema de la generalidad que no exploraré consiste en simplemente apelar a las soluciones que se han propuesto al pro-



El proceso,  $P$ , que resultó en mi creencia de que tengo enfrente una computadora, cae bajo diferentes tipos de proceso.  $P$  es un caso, por ejemplo, del tipo de procesos que tienen como insumo (presumiblemente) haces de luz, que estimulan mi retina, pasan por mi sistema visual y gracias a mis capacidades cognitivas pueden transformarse en la creencia de que tengo enfrente una computadora. Este tipo de proceso, plausiblemente, es fiable: plausiblemente, mis capacidades para formar creencias con base en la información visual son fiables. Pero  $P$  también es un caso del tipo de proceso que se inicia con información recibida por alguna FIE los viernes por la noche, y termina en una creencia. Este tipo de proceso, plausiblemente, es mucho menos fiable que el primer tipo. Sin embargo,  $P$  es un caso de ambos tipos de proceso. Entonces, necesitamos un criterio para distinguir de entre ellos (y de todos los demás tipos de procesos de los cuales  $P$  es un caso) un tipo *distinguido* de proceso—en el sentido en que sea el tipo de proceso que determina la fiabilidad de  $P$ .

Por supuesto, uno puede extraer, quizá, el promedio—u otra medida estadística, como la media, moda, etc.—de la fiabilidad de *todos* los tipos de proceso de los cuales  $P$  es un caso; pero entonces necesitamos criterios no-arbitrarios para elegir entre extraer promedios, medianas, modas, etc., para entonces obtener la fiabilidad de  $P$  como el número resultante.<sup>111</sup>

#### 54. el problema de la generalidad

Llegamos, entonces, a lo que Connee y Feldman (1998) popularizaron como “el problema de la generalidad para el fiabilismo de proceso”. Este problema se puede formular así:

(PROBLEMA DE LA GENERALIDAD)

Digamos que un conjunto de procesos  $\underline{P}$  *determina la fiabilidad* de un proceso concreto (un *token*)  $P$ , sii existe un estadístico muestral  $\sigma$  tal que  $\sigma(\underline{P}) =$  la fiabilidad de  $P$ :

¿Cuál es el tipo de proceso *distinguido* bajo el cual cae  $P$ , donde un tipo de proceso  $\underline{P}$  es *distinguido para*  $P$  sii el conjunto de procesos que cae bajo  $\underline{P}$  determina la fiabilidad de  $P$ ? (para TIPOS)

¿Cuál es la *esencia cualitativa* de  $P$ ,  $Q(P)$ , tal que el conjunto de procesos que ejemplifican  $Q(P)$  en otros mundos posibles determina la fiabilidad de  $P$ ? (para MODAL, interpretado bajo el esencialismo)

¿Cuál es la clase de *relaciones de contraparte* para  $P$ ,  $C(P)$ , tales que el conjunto de las  $C$ -contrapartes de  $P$  en un contexto dado, determinan la fiabilidad de  $P$  en ése contexto? (para MODAL, interpretado bajo la teoría de contrapartes)

En esencia, pues, el problema de la generalidad existe porque no hay una manera obvia de que el fiabilismo encuentre el proceso cuya fiabilidad deba medirse para determinar la justificación de

---

blema de la clase de referencia, por ejemplo invocando a los *criterios de selección de modelos estadísticos* (Cheng 2009). No exploraré demasiado esta solución, porque creo que se acerca, y es menos específica para el caso del fiabilismo, a la solución de Beebe, que consideraré en §5.5. Véase también Connee & Feldman, 2004a.

<sup>111</sup>Wunderlich (2003) considera una opción mucho más compleja, que también se basa en extraer *todos* los tipos de proceso de los cuales  $P$  es un caso, junto con sus números de fiabilidad. Sin embargo, el desarrollo de tal idea sigue sin determinar un número de fiabilidad único para cada proceso, a menos que se postule algún criterio para distinguir entre los diferentes números resultantes. Véase §2.6. del artículo de Wunderlich.

una creencia.

Para simplificar la exposición, me concentraré en el problema formulado para TIPOS; además, claramente, una respuesta a ése problema nos brindaría una respuesta a los otros dos problemas (para las dos formulaciones de MODAL).<sup>112</sup>

Debemos notar que el espacio de soluciones posibles al PROBLEMA DE LA GENERALIDAD está acotado por dos respuestas que parecen claramente inapropiadas, que siguiendo a Conee y Feldman llamaré “el caso único” y “el caso de no distinción”.

El *caso único* consiste en resolver el PROBLEMA DE LA GENERALIDAD identificando el tipo distinguido de  $P$  como el tipo bajo el cual *sólo* cae  $P$ . Esta es una mala respuesta, pues, como vimos al postular FIABILIDAD NO-TRIVIAL, la asunción DESCRIPTIVA nos lleva a postular que  $P$  pueda tener una fiabilidad estrictamente entre 1 y 0—pero un tipo de proceso bajo el cual sólo cae  $P$ , tendrá fiabilidad 0 (si  $P$  resulta en una creencia falsa), o 1 (si  $P$  resulta en una creencia verdadera), yendo así en contra de FIABILIDAD NO-TRIVIAL.

Otra respuesta en el otro extremo es el *caso de no distinción*. Si el tipo distinguido para  $P$  fuera demasiado amplio, entonces determinará una fiabilidad que puede no ser la de  $P$ . Tomemos por ejemplo el proceso que produjo mi creencia de que tengo una computadora enfrente. Si su tipo distinguido incluye a procesos que producen creencias mediante el mero *wishful thinking*, dado que éstos procesos regularmente terminan en creencias falsas, el tipo distinguido del proceso que produjo mi creencia sobre la computadora no será fiable, y con ello mi creencia no estará justificada. Pero parece que tal creencia, con una base meramente perceptual, está bastante más justificada que mis creencias producidas por *wishful thinking*. Necesitamos, entonces, un criterio fino para individuar tipos distinguidos, que se corresponda con la fiabilidad que, *intuitivamente*, tienen mis procesos perceptuales—y así, con la justificación de mis creencias perceptuales.

Con esto, una solución al PROBLEMA DE LA GENERALIDAD tiene que manejarse entre dos extremos: entre el extremo del proceso que o no es nada fiable (fiabilidad 0) o lo es absolutamente (fiabilidad 1), porque su tipo distinguido es tan específico que sólo lo incluye a él, y entre el extremo del proceso al que se le asigna una fiabilidad muy alejada de la que intuitivamente tiene, debido a que su tipo distinguido es tan poco específico que incluye a procesos muy distintos, con fiabilidades muy dispares. Al manejarse entre esos dos extremos, la respuesta en cuestión tiene que darnos un criterio para individuar los tipos distinguidos de proceso, que intuitivamente se correspondan con la fiabilidad de los casos de esos tipos de procesos.

Sin una solución al problema de la generalidad, no tenemos, en principio, una manera de entender lo que FIABILISMO DE PROCESO-VERSIÓN SIMPLE dice. Además de los dos extremos a evitar que hemos visto, Conee y Feldman (1998: secc. I.B) pusieron los siguientes requisitos sobre las soluciones admisibles, que yo renombraré:

(SISTEMATICIDAD) Debe ser *principada*, en el sentido en que no es equivalente a una selección caso-a-caso de tipos que coincida con nuestras intuiciones.

<sup>112</sup>De la siguiente manera: si  $\underline{P}$  es el tipo distinguido de  $P$ , entonces, abstrayendo las propiedades que  $P$  tiene en común con todos los otros miembros de  $\underline{P}$ , habremos encontrado tanto  $Q(P)$  como  $C(P)$ :  $Q(P)$ , la esencia cualitativa de  $P$  que determina los  $P$ s en otro mundo que determinan la fiabilidad de  $P$ , consiste en aquellas propiedades que  $P$  tiene en común con los demás miembros de  $\underline{P}$ ; mientras que  $C(P)$  tiene como miembros a todas y sólo aquellas relaciones de contraparte  $c$  tales que:  $c(P, P^*)$  sii  $P^*$  tiene las propiedades que  $P$  tiene en común con los otros miembros de  $\underline{P}$ .

La idea detrás de esta restricción es que una selección caso-a-caso no imparte ningún contenido definitivo sobre la noción de “tipo distinguido”, y por ello, para tener una teoría fiabilista que sea informativa, primero tenemos que saber qué determina, en lo general, el “tipo relevante”.

(INTUITIVIDAD) La regla que determine los tipos relevantes debe implicar una *clasificación razonable*, en el sentido en que las creencias cuyo tipo relevante es  $\underline{T}$  deben tener un grado de justificación correlacionado con la fiabilidad de  $\underline{T}$ —en breve, la regla debe implicar una correlación *intuitivamente plausible* entre fiabilidad y justificación.

Notemos que, al poner este criterio a trabajar para eliminar varias soluciones propuestas al problema de la generalidad, Connee y Feldman asumen que la justificación de nuestras creencias está, al menos, muy fuertemente correlacionada *con nuestros juicios acerca de la justificación*—en el sentido en que, si tomamos a una creencia  $B$  como justificada, mejor que sea el caso que la solución en cuestión regrese el resultado de que el proceso *token* relevante que produjo  $B$  es suficientemente fiable.<sup>113</sup> Goldman parece haber asumido tal idea también, al postular

DESCRIPTIVIDAD.

Finalmente, tenemos:

(ESPÍRITU FIABILISTA) La solución debe tener un *espíritu fiabilista*: la regla que nos diga qué tipos son relevantes para qué creencias debe implementar la tesis fiabilista central de que la fiabilidad de los procesos de producción de creencias fija la fiabilidad de las creencias producidas por tales procesos. Una solución no puede dar una regla de relevancia que apele a una regla de evaluación epistémica no-fiabilista.

La idea aquí es clara: una solución que apele a teorías no fiabilistas, no resolverá el problema *para el fiabilismo*—mostrará solamente que *otra* teoría lo resuelve.

Veamos ahora qué podemos hacer con estas limitantes.

## 5.5. contra el problema de la generalidad

Entendamos *anti-realismo epistemológico* como una postura que toma a las propiedades y estados que son el objeto de interés de la epistemología como siendo dependientes de la mente.<sup>114</sup> Por otro lado, una *epistemología realista* es cualquier teoría epistemológica de acuerdo a la cual tales objetos son existentes e independientes de la mente—o, en breve, *objetivas*. Estas distinciones están al nivel de las *teorías ontológicas*, en el sentido que revisamos arriba (§5.1.)

En esta sección argumentaré que dos de los tres requisitos de Connee y Feldman (SISTEMATICIDAD e INTUITIVIDAD) están en tensión: mientras que uno tira hacia una epistemología realista, otro tira hacia una epistemología anti-realista. (ESPÍRITU FIABILISTA es neutral acerca de las dos opciones epistemológicas, y lo asumiré aquí sin más.) Si esto es así, los requisitos que Connee y Feldman ponen para resolver el PROBLEMA DE LA GENERALIDAD son excesivamente demandantes, y por ello el fiabilista no está obligado a satisfacerlos todos. Ahora veremos cada uno de estos puntos con mayor detalle.

<sup>113</sup>Al rechazar (en las páginas 7-9) varios intentos de resolver el problema apelando al sentido común, Connee y Feldman apelan categóricamente a lo que está o no justificado.

<sup>114</sup>Por supuesto, bajo otra manera de entender al anti-realismo, una teoría anti-realista de  $X$  es una bajo la cual  $X$  simplemente no existe. No consideraré tales tipos de anti-realismo.

Consideremos primero INTUITIVIDAD. Noté arriba que esto supone que la solución al PROBLEMA DE LA GENERALIDAD debe implicar que las fiabilidades resultantes están estrechamente correlacionadas con los juicios intuitivos acerca de las propiedades epistémicas para las cuales se ofrece una teoría fiabilista. Así, por ejemplo, una solución que satisfaga INTUITIVIDAD implicará que una creencia dada es producida por un proceso altamente fiable cada vez que la creencia en cuestión esté, *intuitivamente*, altamente justificada.

¿Por qué se postularía una limitante como INTUITIVIDAD? Bueno, como vimos en §5.1 arriba, los filósofos que tengan un acercamiento relativamente “intenso” a la función de nuestras intuiciones, asumirán que una teoría que no se apegue a nuestras intuiciones sobre *X*, simplemente no habla *de X*—sino, a lo más, de alguna otra cosa. Puesto como eslogan, el acercamiento intenso dice: o apego a intuiciones o cambio de tema. Una postura relativamente intensa se presupone al postular INTUITIVIDAD: la idea de que debe rechazarse una teoría fiabilista que no concuerde en al menos un alto grado con nuestros juicios intuitivos, presupone que una teoría acerca de *X* debe acordar con nuestras intuiciones acerca de *X*.<sup>115</sup>

Sin embargo, ¿por qué debería ocurrir lo último? Por poner un caso para contrastar, nadie asume que lo que resulte verdad acerca de, digamos, la composición última de la materia, debe seguir a nuestras intuiciones. ¿Por qué? Esta es una pregunta muy compleja, pero creo que plausiblemente pueden individuarse algunas razones, no todas independientes entre sí (*cf.* Omnès, 1999; Wallace 2010):

(1) *La materia es algo independiente de nuestra mente.* Se asume entonces que no tenemos un acceso epistémico privilegiado a las propiedades intrínsecas de aquello que es independiente de nuestra mente.

(2) *Nuestras intuiciones provienen de fuentes que no funcionan igual de bien para toda área.* Es decir, las intuiciones que formamos en la vida diaria las formamos con los recursos limitados de nuestra memoria y otras facultades mentales, incluso quizá con el adoctrinamiento cultural o en función de nuestro interés en ciertos temas. Tales fuentes de creencias, se asume, son menos fiables que otras fuentes (como los métodos de las ciencias), al menos para dominios como la constitución de la materia. (Quizá uno podría buscar un argumento evolutivo a favor de esta idea: plausiblemente, la función primaria de nuestras facultades mentales era devolvernos las características de nuestro ambiente inmediato, así que es sólo con entrenamiento y regimentación metodológica que tales facultades pueden aspirar a conocer otros dominios. Véase las secciones 4.2 y 4.3, arriba.)

(3) *Se cree que la teoría correcta acerca de la naturaleza última de la materia puede, quizá con ayuda de otras teorías, explicar nuestras intuiciones.* Por ejemplo, algunos teóricos de la mecánica cuántica asumen que mecanismos como la decoherencia pueden explicar por qué no hay efectos cuánticos, y sí comportamiento clásico, a nivel macro. Otros creen que el comportamiento cuántico sí está presente a nivel macro. En cualquier caso, se asume que aunque en el estado presente de la ciencia no podamos explicar por qué nuestras intuiciones

---

<sup>115</sup>Véase, por ejemplo, la nota 6 de Conee & Feldman 1998, donde asumen que Alston no está interesado en *el problema de la generalidad* (aunque le llame por ese nombre) al no asumir que una solución a este debe casar con nuestras intuiciones. Según ellos, el problema de la generalidad es “[...] aquel de formular una teoría y de *estar en lo correcto*” (“[...] that of getting the theory stated and *getting it right*”), con lo que asumen que una “solución” que no case con nuestras intuiciones simplemente no puede estar en lo correcto (o “*get it right*”).

sobre el comportamiento de la materia no casan con lo que la teoría sobre su constitución nos dice, en algún momento tendremos disponibles teorías que nos ayuden a resolver tal problema. Así, una teoría sobre la composición última de la materia no está forzada a respetar todas nuestras intuiciones—sólo está forzada a reconocer que es un problema que puede ser resuelto por teorías puente que muestren la compatibilidad de la “imagen manifiesta” con la imagen que nos da la teoría.

La pregunta ahora es: *¿por qué no podría pasar lo mismo con una teoría sobre la justificación epistémica?* Si aceptamos una epistemología realista, entonces por definición aceptamos que las propiedades y estados epistémicamente relevantes son objetos en el mundo, que no dependen de la mente: *tener conocimiento* sería un estado tan natural, objetivo e independiente de la mente como *tener carga eléctrica*, o *ser una bacteria*. ¿Por qué pensar que una teoría sobre propiedades objetivas como estas tienen que casar *perfectamente* o incluso *altamente* con nuestras intuiciones? Quizá ellas sólo nos han guiado para determinar el referente intencional de ciertos términos—“conocer”, “saber que”, etcétera—pero, una vez que las buscamos en el mundo, no hay una garantía *a priori* de que tales intuiciones vayan a resultar, todas o incluso su mayoría, verdaderas. No es claro que cambiemos de tema—pues podemos trazar una historia suficientemente continua desde nuestros acercamientos intuitivos hasta el resultado teórico contra-intuitivo, que nos muestre qué tipo de problemas y principios generales nos llevaron a la otra orilla contra-intuitiva.<sup>116</sup>

Esto es presionar el campo del “acercamiento intenso” a las intuiciones desde el campo del acercamiento “relajado”, por supuesto—pero son exactamente el tipo de consideraciones que son motivadas por la filosofía de la referencia influida por Kripke y Putnam.

En resumen: si asumimos una epistemología realista, entonces las propiedades y estados epistémicos son objetivos—tan independientes de la mente como los quarks o las galaxias. Cuando una entidad es objetiva, hay poca presión en asumir de entrada que una teoría sobre esa entidad es aceptable sólo si implica que todas, o muchas, de nuestras intuiciones acerca de tal entidad resultan verdaderas. (Se evade la objeción del cambio de tema apelando a un acercamiento a las intuiciones motivado por la filosofía externista y causal de la referencia.) Así, parece que INTUITIVIDAD está en fuerte tensión con una epistemología realista.

¿Qué pasa cuando consideramos INTUITIVIDAD a la luz de una epistemología anti-realista? Por definición, esto es comprometerse con la idea de que las propiedades y estados epistémicos son productos de la mente: estados como *tener conocimiento* o propiedades como *ser racional* son tan dependientes de la mente como (putativamente) que algo tenga un sabor *feo*, que cierto *equipo* cuente como habiendo *anotado gol*, o que ciertas personas cuenten como *casadas*. Es decir: sin mentes suficientemente sofisticadas en el universo, nada podría haber contado como teniendo conocimiento—así como nada hubiera sido un *gol* (incluso aunque por azares del destino objetos redondos con manchas negras y blancas hubieran cruzado una línea blanca en el paso, trazada bajo un rectángulo) o nadie hubiera estado casada (o divorciada, o viuda).

Una idea que naturalmente se apareja con la tesis anti-realista es que las propiedades epistémi-

---

116 Véase también Weatherson (2003) para una defensa de la idea de que aceptar la objetividad del conocimiento—la idea de que el conocimiento es “natural”, en el sentido de Lewis (1984)—nos justifica en rechazar algunas intuiciones que parecen centrales en el concepto *folk*. Weatherson arguye sobre esta base que podemos aceptar que los sujetos en los casos Gettier *saben*, por más que la intuición *folk* parezca decir lo contrario.

cas son *evaluativas*. Hay goles porque *nosotros* evaluamos de tal y cual manera que la pelota cruce la portería (Searle 1997); así, según esta postura, hay conocimiento, justificación o racionalidad porque *nosotros* evaluamos de tal y cual manera ciertos estados mentales (Field 2009; Williams 1991: cap. 3).

(Aquí hay que eliminar de entrada cierta posible confusión. Es un hecho trivial que algo se llama “quark”, o cae bajo el concepto de *quark*, porque *nosotros* tenemos tal palabra o usamos un concepto. Pero eso, en y por sí mismo, *no* hace que los quarks dependan de nosotros. Los quarks son cosas “ahí afuera”, determinadas por la naturaleza, no por nuestro lenguaje. La diferencia con objetos dependientes de la mente es que, *para que ellos existan*, debe haber gente que piense en ellos o que use el lenguaje de cierta manera. Aún si no hubiera lenguaje, habría quarks— aunque nada se llamaría “quark”, porque no habría lenguaje. Pero si no hubiera lenguaje, simplemente *no habría* matrimonios.)

Ahora, INTUITIVIDAD parece una mucho mejor idea bajo una epistemología anti-realista. Pues entonces *nuestras* intuiciones acerca de los objetos dependientes de la mente son las que fijan sus características y existencia: ser dependiente de la mente es estar esencialmente determinado por la mente. ¿Qué otra cosa, si no nuestras creencias acerca de ellos, podrían fijar a los objetos epistémicamente relevantes, si estos dependen de la mente?

Bien, aunque sean creencias, quizá no *nuestras* creencias. Quizá alguien quiera afirmar que no todo objeto dependiente de la mente dependa de *una sola* mente, o de un grupo de mentes específico. Tomemos de nuevo a la institución del matrimonio: que dependa de la mente no es decir que *cualquiera* está autorizado en crear un matrimonio. (Si hoy yo digo “¡Sea yo, ahora, esposo de J.K. Rowling!”, no crearé un matrimonio.) Plausiblemente, Buffy Summers (la protagonista de la serie *Buffy The Vampire Slayer*) es un objeto creado por y dependiente de la mente. Pero no de toda mente: Joss Whedon tiene más autoridad sobre las propiedades de Buffy que, digamos, yo. Es decir: que un objeto dependa de la mente no es decir que *cualquier* mente puede crearlo, o fijar sus propiedades. Hay ciertas *autoridades* en ese asunto.<sup>117</sup>

Esto muestra que INTUITIVIDAD debe restringirse. ¿Cómo? No creo que haya una respuesta clara y definitiva a esta pregunta. Las teorías que buscan capturar las intuiciones que los hablantes de cierto lenguaje *L* asocian a cierto término *t* regularmente buscan los usos de *t* sin errores de ejecución de los hablantes competentes de *L*, pero eso todavía no individua un conjunto bien deli-

<sup>117</sup>Entonces alguien podría preguntar: ¿qué ventaja tiene, entonces, asumir un anti-realismo para salvar INTUITIVIDAD, si ni siquiera el anti-realismo se ve compelido a rescatar *toda* intuición? Sigue siendo verdad una idea con tintes realistas: *que hay posturas privilegiadas al determinar las propiedades epistemológicas*—si en el realismo era el “punto de vista” de la realidad, en el anti-realismo es el punto de vista de los creadores privilegiados de objetos dependientes de la mente. INTUITIVIDAD entonces parece no casar bien con el anti-realismo tampoco.

La respuesta es que es plausible y justificado interpretar INTUITIVIDAD como *también* teniendo tal compromiso. Cuando INTUITIVIDAD requiere que una solución al problema de la generalidad sea intuitiva, no necesariamente tiene que afirmar que las intuiciones en cuestión sean las intuiciones de *todo* usuario de la terminología epistémica. De hecho, no debería—porque eso sería inconsistente. Supongamos que una propiedad o estado epistémico *X* es caracterizado como *F* bajo cierta teoría, y como no-*F* bajo otra teoría. Entonces quienes aceptan la primera teoría, si alinean sus intuiciones con ella, contarán como usuarios del término “*X*” que asocian con estas intuiciones; y algo análogo para los defensores de la teoría no-*F*. Entonces, bajo la interpretación de INTUITIVIDAD que estoy rechazando, una teoría defendible de *X* tendría que incorporar las dos intuiciones—pero eso la haría inconsistente, pues las dos intuiciones son incompatibles entre sí.

mitado de usos: ¿quién es un *hablante competente*, por ejemplo? En el caso de la epistemología que nos ocupa, no pueden ser sólo los hablantes expertos, por supuesto—pues privilegiar sus usos de los términos epistemológicos prejuzgará varias cuestiones teóricas que deberían dejarse abiertas para investigación posterior, no resolverse por un mero recurso al significado. Pero tampoco pueden ser sólo las personas no educadas en la epistemología analítica, porque es claro que también los hablantes expertos son competentes, aunque tengan teorías divergentes.

Hay cuestiones delicadas aquí, pero, en resumen, INTUITIVIDAD en un marco anti-realista todavía *no* nos dice de *quién* son las intuiciones a las que la teoría buscada debe obedecer. Le falta contenido en este respecto. No pretendo ofrecer esto como un argumento definitivo en contra de INTUITIVIDAD—después de todo, una vasta mayoría de las posturas en casi cualquier rama de la filosofía apelan a las intuiciones del “folk”, al cual no se le especifica con mayor detalle, y no estoy dispuesto a rechazarlas por esta razón. Esto es un problema abierto en la filosofía de la filosofía (cf. Machery & Mallon & Nichols & Stich 2004; Dunaway & Edmonds & Manley 2012), y no busco resolverlo aquí.

Ahora revisemos brevemente otro posible problema. La idea es la siguiente. Si afirmamos—como dice el anti-realista—que las propiedades epistémicas son dependientes de la mente, uno podría preguntarse por qué entonces necesitamos hacer epistemología ¿si nosotros creamos su objeto de estudio! La asunción es lo *dependiente* de la mente es por ello *transparente* a ella. Pero esta asunción se ve contradicha por el hecho de que necesitamos hacer bastante filosofía, a través de varios años, para conocer la estructura de los objetos de estudio de la epistemología. ¿Cómo hemos de explicar esto?

Hay dos salidas, compatibles entre sí. En primera, la idea de que nuestros estados mentales—sean intuiciones o no—son transparentes es bastante problemática (Nuccetelli 2003; Williamson 2000, cap. 3). Así, podríamos tener intuiciones acerca de un objeto y aún así no tener un acceso transparente a ellas—quizá las tenemos, pero no sabemos que las tenemos, o no accedamos conscientemente a su contenido. La tarea de la epistemología podría verse como la de hacer explícitas estas intuiciones mediante casos concebibles (Jackson 1998, cap. 2) y luego sistematizarlas, quizá mediante el equilibrio reflexivo (Lewis 1973: 88). En segunda, arriba argumenté que la dependencia de la mente *no* implica dependencia de *cualquier* mente. Ahora, esto puede ser otra fuente de la no-transparencia que caracteriza a los objetos de la epistemología. Necesitamos hacer epistemología, bajo esta hipótesis, al menos en parte porque *no cualquiera* fija la naturaleza de los objetos de la epistemología: al menos varios debates en epistemología pueden verse como debates acerca de quién tiene el derecho de fijar la naturaleza de los objetos de la epistemología. De todos modos, como nos dice INTUITIVIDAD junto con una epistemología anti-realista, debemos esperar que la solución al PROBLEMA DE LA GENERALIDAD correlacione fuertemente la fiabilidad con juicios intuitivos sobre justificación—pues son *justamente* esos juicios los que, a final de cuentas, fijan la naturaleza de la justificación.

Hasta ahora he argumentado que INTUITIVIDAD es bastante más plausible cuando se asume una teoría anti-realista que cuando se asume una realista. Esto nos da evidencia para suponer que las intuiciones que soportan a INTUITIVIDAD son intuiciones *anti-realistas*. Sin embargo, ahora procederé a argumentar que las intuiciones que soportan al otro requisito—SISTEMATICIDAD—son, plausiblemente, intuiciones realistas.

Recordemos que, de acuerdo a SISTEMATICIDAD, una solución al PROBLEMA DE LA GENERALIDAD debe

ser *principada*, en el sentido en que no es equivalente a una selección caso-a-caso de tipos que coincida con nuestras intuiciones. Esto implica que una solución aceptable debe postular *reglas* para individuar el tipo relevante de cada proceso al que se le evalúa su fiabilidad.

¿Qué tan motivada está una limitante como SISTEMATICIDAD, dada una epistemología anti-realista? Para ver ello, asumamos por un momento el anti-realismo. Eso quiere decir que la existencia y naturaleza de las propiedades y estados objeto de la epistemología estarían fijados por la actividad mental de los seres humanos. La pregunta entonces es: ¿por qué los objetos epistémicos que creamos tendrían que estar gobernados por principios o reglas, como asume SISTEMATICIDAD?

Tomemos dos casos paradigmáticos de entidades creadas: los personajes de una ficción y los estados-nación. Para los primeros, no es claro qué tipo de reglas *tienen* que ser obedecidas al crearlos—parece que las únicas limitantes son los límites de la imaginación y de nuestras demás facultades. Las historias de ficción pueden ser metafísicamente imposibles, lógicamente incompletas (por ejemplo, puede ser falso tanto que Sherlock Holmes tiene un lunar en el tobillo derecho como que no lo tiene), o hasta inconsistentes. No es para nada obvio que la creación de entidades de ficción sean gobernadas por reglas—más allá de las limitantes de nuestra imaginación y nuestras capacidades.

Lo mismo sucede con los estados nación, donde parece que las únicas limitantes son cuestiones prácticas (o políticas)—como la utilidad de adoptar tal o cual ley. Pero su creación no parece necesitar de gran sistematicidad: incluso hay casos en los que las leyes que los gobiernan pueden resultar inconsistentes. Además, la basta variedad de modelos de estado—incluyendo los aspectos legal y económico—a través de la historia parece mostrar que, si hay algún tipo de regla objetiva gobernando la creación de estados, esta es muy abstracta. Aún así, yo estoy dispuesto a aceptar que los estados-nación existen. Por lo tanto, parece que podemos crear entidades dependientes de la mente con muy poca sistematicidad—o, al menos, gobernados bajo casi ninguna regla además de las cuestiones prácticas; o quizá incluso reglas muy, pero muy, abstractas que coincidentemente todavía no descubren las ciencias sociales.

Esto no quiere decir que todo objeto dependiente de la mente no tenga ninguna estructura sistemática en absoluto. Por ejemplo, filósofos del lenguaje y lingüistas buscan descubrir las leyes semánticas, sintácticas, etc., que gobiernan a nuestro lenguaje (aunque su existencia también es controversial, cf. Rayo 2008). Pero incluso en el caso del lenguaje—quizá el caso paradigmático de algo dependiente de la mente y con una estructura sistemática, incluso quizá matematizable—tenemos términos para los cuales es difícil establecer reglas completamente determinadas y sistemáticas. Por ejemplo, ¿cuál es la regla que gobierna la aplicación del término “alto” en español? Parece que no simplemente decimos que alguien es alto cuando tiene cierta estatura determinada. Una comunidad de pigmeos que hablara español diría que yo soy alto; un equipo de la NBA que hablara español no lo diría. Se puede decir que hay ciertos criterios generales para su aplicación en diferentes contextos, de todos modos. Quizá, por ejemplo, la regla para “alto” sea algo como: “alto” aplica a  $x$  en un contexto  $C$  sólo si  $x$  tiene una estatura mayor que un promedio de personas salientes en  $C$ . Pero ¿qué tan *mayor* tiene que ser la estatura? Y ¿qué determina cuál es el grupo de personas *saliente* en  $C$ , tal que relativo a *ese* grupo, se haga un promedio intuitivo para medir altura? No es claro que haya una manera sistemática de responder a estas preguntas. Quizá, en los términos de Heller (1995), “alto” es *ricamente sensible al con-*



*texto*: a diferencia de términos como “yo”, cuyo contenido es fácil de determinar relativo a un contexto—“yo” simplemente se refiere al hablante que pronuncia la palabra en el contexto de emisión—la referencia de “alto” parece ser afectada por una variedad mucho más amplia de factores contextuales que las que es posible determinar sistemáticamente.

Esto muestra que aunque los objetos dependientes de la mente pueden estar gobernados bajo ciertos principios, hay ciertas limitantes en el alcance de tales principios. Lo cual tampoco parece completamente sorprendente. Cuando creamos objetos—instituciones, ficciones, normas legales—regularmente no tenemos *a priori* una idea de cómo se comportarán en todo momento, y no siempre podemos estipular de manera exhaustiva y completa su naturaleza.<sup>118</sup> Eso parece ser distinto con objetos independientes de la mente, para los cuales parece ser verdad que están completos, que es determinado si poseen o no cualquier propiedad (real) dada—seguro esta es una de las razones por las que aceptar que hay contradicciones o fallas de la ley de tercero excluido *en el mundo* parece tan implausible.

Concluyo que, dada la postura anti-realista, es difícil justificar SISTEMATICIDAD—porque no tenemos una garantía *a priori* de que un objeto creado por la mente vaya a obedecer principios sistemáticos. Lo contrario sucede si asumimos una postura realista. Veamos.

Suponiendo el realismo, es más plausible que las entidades epistemológicas estén gobernadas bajo principios sistemáticos. Parece que la Naturaleza (por usar la manida metáfora) crea objetos siempre gobernados por leyes (aunque sean probabilísticas). Esta es la asunción principal de las ciencias naturales. De hecho, es justo porque la idea de leyes que gobiernan de manera absoluta al reino social es mucho menos evidente, que a muchos les ha parecido que no puede haber *ciencias* sociales, estrictamente hablando. Esto nos da otra disparidad, *prima facie* al menos, entre objetos creados por la mente—sociales—y objetos naturalmente existentes.

Si propiedades como la justificación o estados como el conocimiento son objetivamente existentes, tan creados por la Naturaleza como las galaxias o las neuronas, entonces es más plausible suponer que están gobernados por principios—a saber, las leyes naturales (de la misma manera en que las galaxias y las neuronas lo están.)

Bajo la postura realista, es mucho más plausible suponer que los objetos de la epistemología están gobernados por principios sistemáticos—no otros que las leyes de la naturaleza. Esto muestra que SISTEMATICIDAD está motivada por intuiciones realistas.

Por supuesto, no he mostrado que SISTEMATICIDAD *implique* el realismo, ni que INTUITIVIDAD *implique* el anti-realismo. Sólo he argumentado que cada constreñimiento parece mucho más plausible bajo la correspondiente posición ontológica que bajo su contraria, por lo que parece que las motivaciones intuitivas para cada constreñimiento vienen de diferentes campos. Pero de hecho, bajo asunciones más fuertes puede mostrarse que sí hay tales implicaciones.

Consideremos el siguiente principio:

Si una teoría acerca de una clase *C* de cosas no es correcta a menos que rescate (quizá salvo inconsistencia) *todas* las intuiciones que tenemos acerca de *C*, entonces las cosas en *C* son productos de la mente.

---

118 Puede que en algunos casos esto sí suceda. Por ejemplo, suponiendo que podemos crear objetos mediante recetas recursivas. En ese caso, la cláusula de clausura—la que dice “y nada más que lo anteriormente dicho cae bajo mi definición”—parece ayudarnos a definir completa e inambiguamente la naturaleza del objeto definido. (Esto sucede, por ejemplo, si es que los objetos matemáticos son dependientes de la mente.)

Este principio “puente” claramente hace que INTUITIVIDAD implique al anti-realismo, pero puede parecer demasiado fuerte. No quisiera atribuírselo a Conee y Feldman. Sólo he argumentado que una versión debilitada—no tan general, y requiriendo algo menos que implicación—de este parece estarse asumiendo al postular INTUITIVIDAD.

De manera análoga, el siguiente principio “puente” hace que SISTEMATICIDAD implique al realismo:

Si una teoría acerca de una clase *C* de cosas no es correcta a menos que implique que las cosas en *C* están siempre gobernadas por principios sistemáticos, entonces las cosas en *C* no son productos de la mente.

De nuevo, no le atribuyo algo tan fuerte a Conee y Feldman. Pero una versión debilitada de esta idea—quizá algo como que *la existencia de principios sistemáticos que gobiernen a los C es buena evidencia para pensar que los C no son dependientes de la mente*—parece estarse asumiendo al postular SISTEMATICIDAD.

Para lo que sigue, asumiré que mi argumento es correcto: INTUITIVIDAD y SISTEMATICIDAD *están motivadas por intuiciones inconsistentes*—una viene del lado del anti-realismo, otra desde el lado realista. Esto, por sí mismo, *no* hace que INTUITIVIDAD y SISTEMATICIDAD formen un conjunto inconsistente de requisitos sobre una solución aceptable (a menos que se acepten los dos principios “puente” de arriba). Si eso pasara, Conee y Feldman estarían exigiéndole lo imposible al fiabilista. Pero si pasa lo que he defendido, aunque no le exigen lo imposible, *le exigen algo casi igual de difícil*: le exigen una teoría que haga justicia a intuiciones realistas y a intuiciones anti-realistas. ¿Por qué el fiabilista tendría que cumplir tal requisito? ¿Por qué tendría que diseñar una teoría que *haga justicia a intuiciones inconsistentes*? Cuando hay tales conflictos, regularmente se acepta que una de las dos intuiciones en tensión debe rechazarse—por más que ello pueda tener costos intuitivos.

He argumentado que dos de los requisitos que Conee y Feldman ponen sobre cualquier solución aceptable para el PROBLEMA DE LA GENERALIDAD están en un fuerte conflicto, pues son motivados por intuiciones venidas de posturas inconsistentes entre sí. Además, noté que cuando eso pasa, generalmente no pedimos una solución que haga justicia a ambas intuiciones, sino que aceptamos que una de las dos tienen que rechazarse. Así, de aquí podríamos inferir que el fiabilista puede rechazar el realismo y quedarse sólo con INTUITIVIDAD y ESPÍRITU FIABILISTA; o puede rechazar el anti-realismo y quedarse sólo con SISTEMATICIDAD y ESPÍRITU FIABILISTA.<sup>119</sup> Veamos cómo se verían estos resultados.

Bajo la primera opción, se podría dar una teoría fiabilista que se ajustara en cada caso, o en la mayor cantidad de casos razonablemente posibles, a nuestras intuiciones: cuando la teoría en cuestión implicara que una creencia *c* tiene tal grado *g* de fiabilidad, podríamos revisar nuestras intuiciones y notar que en efecto, intuitivamente *c* tiene el grado (al menos aproximadamente)

---

<sup>119</sup>La tercera opción es, por supuesto, rechazar ESPÍRITU FIABILISTA, pero eso haría que la teoría resultante no fuera fiabilista—y sólo mostraría que *otra* teoría no se enfrenta con el problema de la generalidad. Sin embargo, se puede debilitar el requisito, de manera que la teoría resultante sea reconociblemente fiabilista, aunque incorpore elementos de otra teoría. Por ejemplo, Comesaña (2006) argumenta que una teoría fiabilista que combine elementos evidencialistas puede resolver el problema. La idea básica es que el proceso-tipo en cada caso es *basar la creencia en el cuerpo de evidencia disponible*, y que incluso los evidencialistas como Conee y Feldman están comprometidos con que se pueda individuar tal tipo.

g de justificación. Pero a la teoría *no* se le exigiría brindar principios completamente sistemáticos para individuar procesos. Podría afirmarse que la fiabilidad en cuestión es altamente dependiente de contexto, o se evalúa caso-a-caso de una manera que está incluso menos gobernada por reglas que la que está el lenguaje dependiente de contexto.

De hecho, ya se ha ofrecido al menos una solución sobre estas líneas (Heller 1995). Según Heller, “fiable” es un término altamente dependiente de contexto (a diferencia de términos como “yo”, cuyo contenido depende del contexto de emisión, pero se da por una regla constante), pero que aún así logramos aplicar fuera del contexto de la epistemología, como cuando decimos cosas como “mi auto es fiable”. El contexto ayuda, según Heller, no sólo a determinar qué tan fiable debe ser un tipo de proceso para que una creencia resultante de este esté justificada, sino que ayuda a individuar los tipos de procesos cuya fiabilidad se evalúa.

Conee y Feldman rechazaron la solución de Heller por dos razones: (1) que el contexto pocas veces brinda los elementos necesarios para individuar a un único proceso de producción de creencias, y (2) que el atribuyente del conocimiento (que puede ser el mismo sujeto al que se le atribuye) puede simplemente estar equivocado acerca de qué proceso llevó a la creencia siendo evaluada, pero si eso pasa entonces su contexto de atribución no fija la justificación de la creencia relevante (1998, pp. 20-24).

Creo que estas dos críticas muestran, de manera particularmente clara, cómo es que los requisitos de Conee y Feldman chocan entre sí: cuando alguien ofrece una postura anti-realista (motivada por INTUITIVIDAD y ESPÍRITU FIABILISTA), le exigen algo que sólo sería razonable exigir a un realista, como he venido señalando. Veamos.

A la primera crítica se puede responder que en esos casos, es simplemente indeterminado cuál es la justificación de la creencia relevante. Lo mismo sucede con otros objetos dependientes de la mente: cuando lo que fija su naturaleza no es suficiente, simplemente están indeterminados. Por ejemplo, puede ser indeterminado si cierto movimiento está o no permitido en un juego de mesa. Esta es una muestra más de cómo asumir la postura anti-realista se lleva bien con el rechazo de sistematicidad, y cómo es que Conee y Feldman parecen estar exigiendo soluciones realistas a las teorías anti-realistas.

A la segunda crítica se puede responder que, si el atribuyente de conocimiento puede estar equivocado, es porque se asume que *hay* un hecho *objetivo* acerca de qué proceso resultó en la creencia en cuestión. Así, Conee y Feldman le están requiriendo aquí al fiabilista que sea consistente con el realismo, cuando primero se le exigió que cumpliera con INTUITIVIDAD—criterio más claramente anti-realista.

El problema es que si un fiabilista quisiera dar cuenta de estas ideas realistas, tendría que aceptar que el evaluador está equivocado en tales ocasiones—lo cual haría que las intuiciones del evaluador simplemente chocaran con el resultado de la teoría fiabilista, yendo en contra de INTUITIVIDAD.

Esto es generalizable. *Qua* epistemólogos, nosotros somos evaluadores también. Nada nos impide equivocarnos. Eso implica que habrá casos en que, como el del evaluador que consideran Conee y Feldman en su objeción, tengamos creencias equivocadas acerca de qué proceso resulta en la creencia a evaluar. Pero podemos estar equivocados *porque* hay un hecho no constituido por la mentalidad acerca de qué proceso es relevante para evaluar la justificación de una creencia, y al estar equivocados nuestras intuiciones chocarán con los hechos. Esto es una muestra

concreta de cómo el realismo puede terminar chocando con el requisito INTUITIVIDAD.

Esto nos sugiere una solución conciliadora.<sup>120</sup> Una teoría fiabilista que se guíe bajo INTUITIVIDAD nos servirá para recuperar conceptos epistémicos *folk*: apegarnos lo más posible a la intuición nos dará lo que se entiende por “intuición” o por “conocimiento”. Armados con el concepto *folk*, podríamos buscar en el mundo, ayudados por la ciencia (por ejemplo, la ciencia cognitiva y la estadística), aquellos procesos que mejor balanceen dos requisitos: el ajuste al concepto *folk* y el ajuste a los criterios meta-teóricos de las ciencias. El primer requisito nos sirve para buscar en el mundo aquello que responde a nuestras intuiciones, el segundo nos guía en esa búsqueda bajo cualesquiera principios que guíen a las ciencias relevantes. Pero aquí debemos notar que la fuerza de INTUITIVIDAD es mucho menos importante. Cuando hacemos ciencia, buscamos descubrir *qué hay en el mundo*, y no tenemos por qué imponer ninguna garantía *a priori* de que aquello que halla debe ajustarse, completa o incluso sólo fuertemente, a nuestras creencias previas. Para la segunda teoría, el peso está sobre SISTEMATICIDAD: guiados por los principios que brindan sistematicidad a las ciencias, buscamos en el mundo los procesos cognitivos que dan lugar a nuestras creencias justificadas y nuestro conocimiento, incluso si el resultado no casa perfectamente con lo que creíamos de antemano. (Esto, por supuesto, está motivado por un acercamiento relativamente “relajado” a la función de las intuiciones, como argumenté al inicio del capítulo.) Las ciencias revisan una y otra vez nuestras intuiciones previas, que pasan, después de cierto tiempo, a ser consideradas simplemente *prejuicios*. No hay razón alguna, consistente con la manera en que la ciencia ha revisado nuestras creencias a través de la historia, para imponer un conservadurismo que le dé un lugar central a un requisito como INTUITIVIDAD.

Cerraré esta sección mostrando que de hecho ya *hay* soluciones que cumplen SISTEMATICIDAD, y por ello están mejor motivadas por una epistemología realista—aunque no se lleven bien con INTUITIVIDAD.

Una de ellas es la de Alston (1995), que fue criticada por Conee y Feldman (1998: pp. 10-16). Otra es la propuesta de Beebe (2004), quien propone una *condición de tres niveles* para individuar procesos: el tipo relevante de un proceso token  $t$  es un tipo  $T$  tal que todos los otros procesos que caen bajo  $T$  tales que:

- a) resuelven el mismo tipo de problema de procesamiento de información  $i$  resuelto por  $t$ ;
- b) usan el mismo tipo de procedimiento o algoritmo de procesamiento de información  $t$  usado al resolver  $i$ ; y
- c) comparten la misma arquitectura cognitiva que  $t$ .

Lo siguiente es postular un *criterio de relevancia estadística*: la idea es primero obtener un proceso-tipo  $A$  que, relativo a un token  $t$ , satisfaga la condición de tres niveles. Entendiendo a los procesos-tipo como conjuntos, particionamos a la clase  $A$  en celdas determinadas por factores que afecten la probabilidad del proceso-token  $t$ , relativo a la clase  $A$ . El proceso-tipo relevante para  $t$  corresponde entonces a la subclase más amplia de  $A$  tal que no pueda ser subdividida de manera estadísticamente relevante para la ocurrencia de  $t$ . Esta clase es la clase de procesos tal que añadirle o quitarle algún proceso haría que la ocurrencia de  $t$  fuera más, o menos, probable.

Hasta donde sé, Conee y Feldman no han objetado la propuesta de Beebe; pero el mismo Goldman tiene la siguiente duda: “permanece la pregunta de si siempre hay un conjunto de con-

<sup>120</sup>La siguiente idea está inspirada, aunque no es exactamente la misma, por las ideas de Goldman en su 1992; véase también su 2008.

diciones que satisfacen los estándares de Beebe, *i.e.*, que generan una partición apropiada” (2008: §4).<sup>\*</sup> (También se podría objetar que los resultados no son suficientemente intuitivos, pero arriba ya he respondido sobre esto.) No veo razón por la cual esto deba ser devastador, y cerraré esta sección argumentándolo.

Las condiciones de Beebe aseguran que los resultados del procedimiento de generar particiones estadísticamente relevantes, serán particiones estadísticamente equivalentes. Cuando dos procesos-tipo intuitivamente distintos sean estadísticamente equivalentes, relativo a sus procesos-*token*, el teórico fiabilista puede simplemente identificarlos. Identificar objetos equivalentes bajo las propiedades que se estudian es una práctica recurrente en la ciencia. Que sean distintos puede ser cierto, y que lo sepamos intuitivamente también. Pero hemos dejado de lado las intuiciones para buscar sistematicidad, y en busca de esta pasamos a identificar objetos equivalentes, para propósitos teóricos.

## 5.6. conclusión: a favor del fiabilismo externista

He argumentado a favor del externismo—básicamente, la idea de que el conocimiento depende de, y en una buena parte de casos está totalmente determinado por, las facultades cognitivas que nos brindan información del mundo. Ese conocimiento puede ser de “menor grado”, debido a su automatismo, que el conocimiento obtenido reflexionando o infiriendo a partir de otras creencias o cosas que sabemos—pero es conocimiento, de todos modos. Nuestra vida epistémica depende de él, y el externismo mínimo que defendí da cuenta de ello.

También defendí al fiabilismo, un determinado tipo de externismo mínimo, de uno de sus problemas más importantes: el problema de la generalidad. Argumenté que las limitantes que se ponen para las soluciones a tal problema están motivadas por intuiciones en una tensión tan fuerte que llega a la inconsistencia. Por un lado, las limitantes fácilmente llevan a una postura realista; por el otro, fácilmente llevan a un anti-realismo.

Podemos notar cómo es que los dos lados hacia los que tira pueden dar cuenta de diferentes cosas. Una solución del lado anti-realista puede ajustarse a una teoría de la *atribución* del conocimiento; mientras que una solución realista puede ajustarse a una teoría acerca de la *naturaleza* del conocimiento.

Con este fin, argumenté a favor de un acercamiento a la función de las intuiciones según el cual estas nos ayudan a entender de qué estamos hablando, pero no tenemos una garantía *a priori* de que todas ellas deban resultar verdaderas de su referente intencional. Este fue uno de los pilares en la defensa de la idea de que exigirle a una teoría que respete *todas* las intuiciones que tenemos acerca de su objeto, se acerca demasiado peligrosamente a exigirle que dé como resultado que el objeto en cuestión *de hecho está determinado por nuestras intuiciones*.

Como ya había notado, no he quitado *toda* piedra del camino para la aceptación del externismo y el fiabilismo. Pero al motivarlos y defenderlos, espero haber quitado *alguna importante*.

---

\* “[...T]here remains the lingering question of whether there is always a set of conditions that meet Beebe's standards, *i.e.*, that generate an appropriate partition.”

## 6. Una epistemología de la modalidad y la esencialidad: *representacionalismo fiabilista*

[... S]cience is possible because there is order in Reality. *The laws that structure the representation we form of Reality are an image of its own order.* The whole of science suggests such an answer, but science alone cannot establish or even formulate it, for this assertion is beyond science's own representation. Science is restricted to the region of Reality already explored; it cannot get out of it or assess it. To go beyond what is known amounts to proposing a hypothesis about the unknown, to leaving science and *entering metaphysics.*

—Roland Omnés, *Quantum Philosophy*

**H**an habido varios acercamientos al problema de cómo se supone que podemos conocer que algún hecho dado es posible, necesario, etcétera—en resumen, hay varios acercamientos a la *epistemología de la modalidad*.<sup>121</sup> En un importante número de casos, los epistemólogos de la modalidad han propuesto teorías sobre el conocimiento modal con una formidable cantidad de compromisos teóricos. El caso más claro de esto es la teoría de Chalmers (1996, 2002, 2004), quien para mostrar cómo y por qué (según él) la concebibilidad implica la posibilidad, ha diseñado un sofisticado aparato semántico—el *bidimensionalismo epistémico*. El acercamiento racionalista de Chalmers ha generado varias respuestas (e.g. Martínez 2013; Sturgeon 2010), pero parece que la mayor cantidad de objeciones y problemas van dirigidas a su aparato semántico (Alvarado Marambio 2007; Byrne & Pryor 2006; Nimtz 2004; Schroeter 2005; Speaks 2010; Stalnaker 2006; Yablo 2002). Por poner otro ejemplo, el acercamiento de Peacocke (1999) parece tener problemas debido a la lógica modal que motiva (Williamson 2002), y además tiene fuertes compromisos teóricos con cierta postura acerca de los conceptos (Peacocke 1992) que parece ser o empíricamente falsa o al menos problemática, además de metodológicamente objetable (Machery 2009: §2.3). Otro ejemplo más es el acercamiento de Williamson (2007, cap. 5), que se compromete con ciertas tesis acerca de la equivalencia *epistémica* entre ideología modal y contrafáctica (Jenkins 2008b, Vaidya 2007: §7), con la fiabilidad de nuestra capacidad de imaginar escenarios alternativos (Williamson 2010, 2013b), y con una teoría acerca de los juicios contrafácticos que es empíricamente falsa, o al menos problemática (Machery 2009: 43-46). Estos casos sugieren la necesidad de buscar una epistemología modal que no tenga que estar comprometida con un programa filosófico demasiado *específico*, sino que explote asunciones comunes y muestre cómo ellas dan una *base* para nuestro conocimiento modal—y, quizá, de esencias.

121 Mientras tanto, los acercamientos a la epistemología de la *esencialidad*, específicamente, son menos (Bird 2010, Hale (2013: cap. 11), Lowe 2013, Sveinsdóttir 2013), quizá debido a que la asunción metafísica de que la esencialidad se reduce a la necesidad (cf. §1.4, arriba) podría motivar la idea de que una epistemología de la necesidad es automáticamente una para la esencialidad. Cuando uno cree lo inverso—que la necesidad se reduce a la necesidad—es entonces dable suponer que la epistemología de las esencias será una epistemología de la necesidad, como Hale y Lowe en los textos citados.

Por supuesto, es altamente ingenuo suponer que se puede partir de cero. Pero en este último capítulo busco mostrar que se puede tener una epistemología modal (y esencial) que caiga bajo la epistemología general, con el menor número posible de asunciones especiales al caso modal (esencial). Así, aquí sólo asumiré: (1) cierto externismo mínimo—el que fue defendido en el capítulo anterior; (2) un acercamiento fiabilista al conocimiento, y (3) que tenemos la capacidad de representarnos al mundo, incluyendo en esta una capacidad de *simulación* (hice una rápida revisión de algo de evidencia empírica acerca de esta capacidad en §4.1, arriba).

Con lo que este último capítulo busca dar una epistemología de la modalidad que cumpla los siguientes requisitos:

- (1) sea *compatible* con el estructuralismo absolutamente primitivista, defendido en el capítulo 3 (y por ello, que *no necesite* un reduccionismo de la modalidad),
- (2) dé al menos un inicio de respuesta a los tres problemas individuados en el capítulo 4,
- (3) sin por ello caer en el escepticismo global (*i.e.*, que acepte  $\neg$ MODAL GLOBAL),
- (4) satisfaga EXTERNISMO MÍNIMO, defendido en el capítulo 5,
- (5) involucre una postura fiabilista, defendida en el capítulo 5, y
- (6) evada la mayor cantidad de los problemas que aquejan a las principales epistemologías de la modalidad disponibles.

Lo primero sobre la mesa es hablar sobre la noción de *representación* que asumo aquí. Argumentaré que postular su existencia es un compromiso altamente inocente—los filósofos de básicamente cualquier convicción deben aceptar que podemos representar. Habiendo aceptado ello, mostraré cómo es que la capacidad de representarnos—en particular, de *simular* con esas representaciones—puede brindarnos conocimiento modal. Que las creencias adquiridas por tal capacidad cuenten como conocimiento nos es asegurado por la tesis fiabilista-externista asumida aquí y defendida en el capítulo anterior. Si todo esto va por buen camino, para finales del capítulo habré construido una teoría que satisfaga (1)-(6).

## 6.1. la metáfora de las etiquetas que se pegan al mundo

Es conveniente pensar a una representación como una etiqueta. La idea es que una etiqueta se le pega a un objeto. A donde va el objeto, va la etiqueta. Pero no todas las etiquetas tienen un buen pegamento. Algunas se caen. Una etiqueta con un pegamento insuficiente se puede despegar de su objeto y pasar a pegarse a, digamos, el suelo. Entre mejor sea el pegamento, mejor servirá la etiqueta su propósito: mejor se pegará a su objeto.

Cuando queremos saber si un objeto se ha movido en el espacio con una cierta trayectoria, durante cierto intervalo de tiempo, un pre-requisito de ello es que *sigamos* al objeto—o, como diré, que lo *rastreemos*.

Pensemos en las etiquetas de nuevo. Supongamos que queremos saber si un objeto se ha movido en la oscuridad, y para ello lo rastreamos con una etiqueta fluorescente (por supuesto, a su vez tenemos que rastrear a la etiqueta.) Cuando la etiqueta tiene pegamento suficiente, permanecerá pegada a su objeto, y así, la etiqueta fluorescente es una excelente indicación del objeto. A donde vaya el objeto, irá la etiqueta. Si rastreamos la trayectoria de la etiqueta, habremos ras-

treado la trayectoria del objeto—de nuevo, cuando el pegamento es bueno.

Pero si el pegamento es malo, seguir la trayectoria de la etiqueta no será una indicación excelente de la trayectoria del objeto. Quizá durante cierto tiempo, la trayectoria de la etiqueta sea una buena indicación de la trayectoria del objeto—hasta que la etiqueta se caiga. Cuando se caiga, la trayectoria del objeto y de su etiqueta se separarán. Entonces la trayectoria de la etiqueta no indicará la trayectoria del objeto.

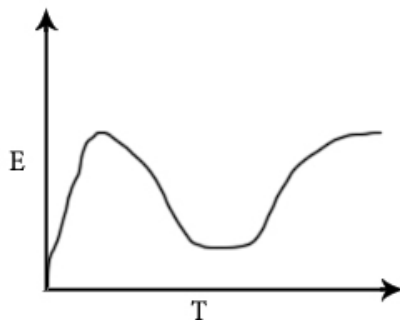
Así, si queremos saber hacia dónde se ha movido el objeto, es mejor que tengamos etiquetas con muy buen pegamento.

Pero el pegamento excelente no es siempre posible. A veces sólo podemos tener etiquetas que sigan al objeto unos momentos, para entonces caerse y dejar de ser buenos indicadores de la trayectoria de este. Si extendemos un poco la analogía, también tendremos etiquetas que después de caerse (como por magia) vuelven a pegarse a su objeto. Esas etiquetas serán mejores que las etiquetas que se caen para no volver a pegarse, pero peores que las etiquetas que jamás se caen.

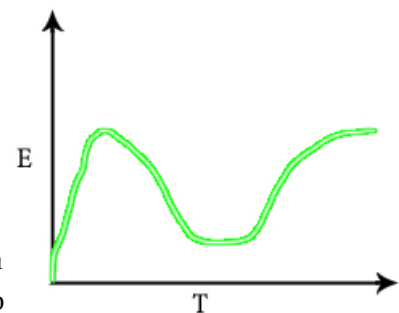
Podemos medir esto en grados. Una etiqueta que se le pega de manera definitiva a un objeto, será un indicador perfecto de su trayectoria: el 100% de las ocasiones, seguir la trayectoria de la etiqueta servirá para seguir la trayectoria del objeto. Una etiqueta que se pega al objeto para inmediatamente después caerse, seguirá al objeto un 0% de las ocasiones. Una etiqueta que se le pega al objeto durante la mitad de su trayectoria, pero se le despega después, lo seguirá a un 50%. Lo mismo para una etiqueta que se pegue al objeto el 25% de su trayectoria, se caiga otro 25%, regrese a pegarse otro 25%, y se caiga para el último 25% de la trayectoria. Y así: habrá etiquetas que sigan a su objeto un 62%, o un 37%, o..., etcétera, de las ocasiones.

Cuanto más alto es el porcentaje de la trayectoria del objeto en que su etiqueta lo sigue, más nos servirá seguir la trayectoria de la etiqueta para seguir la trayectoria del objeto. Una etiqueta que sigue a su objeto el 100% de las ocasiones es una etiqueta que, para propósitos representacionales, sustituye perfectamente al objeto: seguir al objeto o seguir a la etiqueta es, para el propósito de conocer la trayectoria del objeto, lo mismo.

He hablado del cambio en una dimensión: la espacial (bueno, en 3, suponiendo que el espacio es tridimensional, pero simplifiquemos). Tal cambio lo medimos respecto a otra dimensión: el tiempo. Así, tenemos una trayectoria del objeto en dos dimensiones: el espacio y el tiempo. Como en el diagrama:



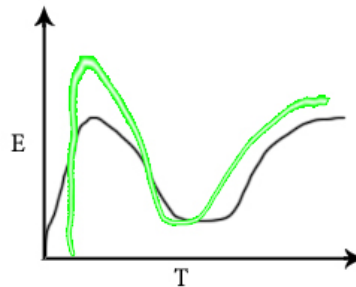
Luego tenemos la etiqueta, que represento con una línea verde. Cuando la etiqueta sigue perfectamente la trayectoria del objeto, las líneas que representan sus trayectorias se traslapan exactamente, como en este diagrama:



Cuando la etiqueta no lo sigue perfectamente, pero sí de manera relativamente estable, la trayectoria de la etiqueta y la del objeto se separan en ciertos puntos, pero se vuelven a unir en otros.



Habr  etiquetas que casi no sigan a su objeto, como esta:



Y quiz  haya etiquetas que no sigan en absoluto al objeto.

Vimos que un pre-requisito para saber cu l era la trayectoria de un objeto en determinado intervalo, era *rastrearlo*. He estado usando el ejemplo de las etiquetas para rastrear a objetos. Pero lo mismo podemos hacer con otras representaciones.

Pensemos en un objeto (digamos, un p ndulo) que se mueve en el espacio. Para conocer su trayectoria durante un intervalo  $t_1$ - $t_3$ , debo, antes que todo, rastrearlo. Supongamos que lo hago mediante la vista (lo rastreo con representaciones visuales.<sup>122</sup>) No me es muy  til ver primero al p ndulo en  $t_1$ , luego al suelo en  $t_2$ , y luego otra vez al p ndulo en  $t_3$ : m s me hubiera servido ver tambi n al p ndulo en  $t_2$ . Si mi representaci n visual *me obligara* a ver al suelo en  $t_2$ , no me ser a muy  til. Ser a mejor dejar esa representaci n visual por otra—por una, dicho metaf ricamente, que se pegara al p ndulo en los tres momentos.

Supongamos ahora que no s lo queremos rastrear la trayectoria de un objeto. Tambi n queremos rastrear el c mo ese objeto va cambiando en esa trayectoria.

De nuevo pensemos en objetos que se mueven en el espacio con relaci n al tiempo. Supongamos que queremos rastrear c mo una persona, Juan, cambia una playera por otra, caminando en el espacio de un punto  $e_1$  a uno  $e_3$ , durante un intervalo  $t_1$ - $t_3$ . En  $t_1$ , Juan est  en  $e_1$ . En  $t_2$ , Juan est  en  $e_2$ , y en  $t_3$ , Juan est  en  $e_3$ . En  $t_1$ , estando en  $e_1$ , Juan tiene una playera roja. En  $t_2$ , cuando ha caminado a  $e_2$ , Juan se ha quitado la playera y no tiene ninguna. (Suponemos que dej  la playera roja en  $e_1$ .) En  $t_3$ , ya estando en  $e_3$ , Juan se ha puesto una playera azul (que suponemos que la recog  en  $e_3$ ).

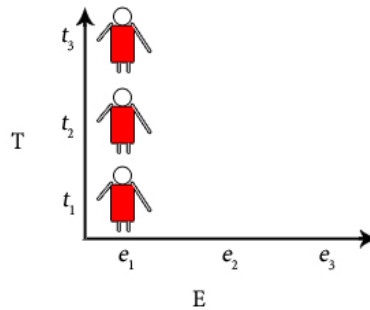
Para rastrear los cambios que Juan hace de su playera y en su posici n relativo al tiempo, de nuevo es un pre-requisito rastrear a Juan. No nos servir  de nada si en  $t_1$  vemos a Juan estando en  $e_1$  con una playera roja, y luego en  $t_2$  volteamos hacia otro lugar y dejamos la vista ah  hasta  $t_3$ . As  s lo sabremos c mo Juan era en  $t_1$ : estaba en  $e_1$  y usaba una playera roja. Pero su trayectoria en el espacio, as  como sus cambios de playera, han sido ignorados.

Hay otra moraleja: necesitamos elegir bien a qu  le ponemos nuestras etiquetas.

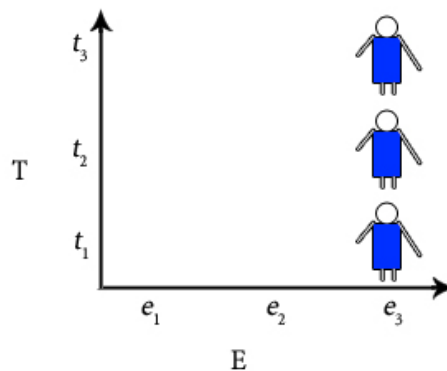
Supongamos que ponemos una etiqueta a la playera roja (que llamar  Roja) en  $t_1$ , y estipulamos que la usaremos para rastrear el estado de Juan: la etiqueta "nos dir " en d nde est  Juan y con qu  playera en cada uno de los tres momentos. Esto nos servir  muy bien para rastrear el estado de Juan en  $t_1$ : como tanto Juan como Roja est n en  $e_1$ , la etiqueta "nos dice" lo correcto en  $t_1$ : que Juan tiene una playera roja y que ambos est n en  $e_1$ . Pero para  $t_2$ , la etiqueta es in til

122 Estoy siendo poco cuidadoso aqu , salt ndome, por simplicidad, importantes debates en la filosof a de la percepci n.

para rastrear el estado de Juan: la etiqueta “nos dice” que Juan sigue en  $e_1$ , y con la playera roja. Lo mismo para  $t_3$ . Es decir, si pegamos la etiqueta a Roja, asumimos que nos dirá el estado de Juan, y dado que Juan deja a Roja en  $e_1$ , lo que la etiqueta “nos dirá” sobre los estados de Juan es esto:



Algo igual de problemático pasará si le pegamos la etiqueta a su playera azul (Azul) y asumimos que nos informará del estado de Juan:



De acuerdo con esa etiqueta, Juan siempre estuvo en  $e_3$ —la etiqueta “cree” eso porque Azul siempre estuvo en  $e_3$ .

Así que para representar bien los estados de Juan—tanto su posición en la línea del espacio ( $E$ ) como si lleva playera y de qué color—a través del tiempo ( $T$ ), queremos un sistema de etiquetado efectivo: uno que rastree de manera adecuada a Juan y sus estados a través del tiempo. Necesitamos, entonces, una “etiqueta” que no sólo rastree a Juan (para conocer su posición en el espacio), sino a si lleva una playera y de qué color. (Uso comillas porque estoy extendiendo la analogía. Para seguir con ella, asumamos que las etiquetas que rastrean las playeras de Juan y colores de estas también están pegadas a Juan.) Necesitamos, en principio, tres etiquetas: una pegada a Juan, tal que su posición en el espacio a través del tiempo nos informe sobre la posición de Juan en el espacio a través del tiempo (es decir, que nos informe sobre la trayectoria de Juan), una también pegada a Juan que diga “Sí” cuando Juan lleve una playera y “No” cuando no lleve ninguna playera (como en  $t_2$ ) y finalmente, una tercera etiqueta que diga “roja” cuando Juan lleve a Roja y que diga “azul” cuando Juan lleve a Azul.

(Quizá, por supuesto, podemos reducir el número de etiquetas a dos: cuando la tercera etiqueta diga “roja” o “azul”, inferimos que Juan lleva una playera. Si la tercera etiqueta tiene un tercer indicador para cuando Juan no lleva ningún color, inferiremos que Juan no lleva ninguna playe-

ra. Así, podemos usar sólo la primera etiqueta, que rastrea la trayectoria de Juan, y la tercera, que rastrea tanto si Juan lleva una playera como de qué color.)

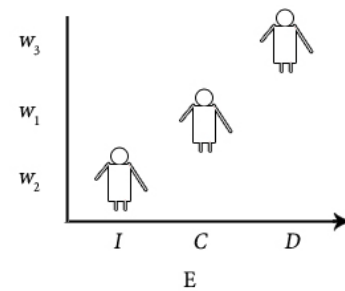
Por supuesto, todo esto asume que las etiquetas tienen muy buen pegamento: si la etiqueta que rastrea la trayectoria de Juan se le cae en  $t_2$ , nos será poco útil para saber en dónde está Juan tanto en  $t_2$  como en  $t_3$ . Y si la etiqueta que nos informa sobre qué color lleva Juan también se despegara, para, digamos, salir volando por el aire y llegar a pegarse a Roja en  $t_2$ , entonces nos informaría erróneamente que Juan lleva una playera roja tanto en  $t_2$  como en  $t_3$ . La situación del pegamento complica las cosas, por supuesto, pero no las hace imposibles: a veces logramos tener etiquetas con buen pegamento.

Ahora pensemos que en vez de tiempos, queremos rastrear los cambios en la posición espacial de Juan, relativos ahora a mundos posibles (dejamos de preocuparnos por el tiempo, por el momento.) De nuevo, es un pre-requisito tener una etiqueta que se le pegue a Juan de manera estable. Lo mejor sería tener una que se le mantenga pegada en todos los mundos posibles donde Juan exista, pero a veces sólo podemos aspirar a menos.

Supongamos que tenemos una etiqueta *perfecta*, que se pega a Juan en el 100% de los mundos donde él existe. Entonces podemos seguir las posibles trayectorias de Juan, siguiendo a su etiqueta.

Supongamos que Juan está en un punto “centro”,  $C$ , en el mundo  $w_1$ . Supongamos que en un mundo  $w_2$  Juan se mueve hacia la izquierda de  $C$ ,  $I$ . En un mundo  $w_3$ , Juan se mueve hacia la derecha de  $C$ ,  $D$ .<sup>123</sup> Así, tendríamos:

Si “seguimos” a la etiqueta de Juan a través de los mundos posibles, podremos rastrear a los posibles cambios de posición de Juan. Así, “siguiendo” a la etiqueta, podremos decir que *Juan podría haber estado en  $I$* : para que eso sea verdad, simplemente se requiere que (i) la etiqueta esté en  $I$  en  $w_1$ , y (ii) que la etiqueta se haya quedado pegada a Juan en  $w_1$ . Eso implica que de hecho Juan está en  $I$  en  $w_1$ .<sup>124</sup> Por razones parecidas, la etiqueta también “nos dice” que *Juan podría haber estado en  $D$* .



Por supuesto, esto es completamente general. En vez de seguir a la etiqueta de Juan a través de los mundos para rastrear su posición en el espacio físico, en principio podemos tener etiquetas que rastreen la posición de Juan en todo el espacio de cualidades, que llamaré  $Q$ .

(Podemos pensar a  $Q$  como el espacio infinito-dimensional de las cualidades—quizá incluso magnitudes—posibles. Un punto en  $Q$  determina un perfil cualitativo.)

<sup>123</sup>Los índices de los mundos son completamente insustanciales. Podemos definir, claramente, funciones de indización de mundos de manera arbitraria (a menos que el conjunto de mundos tenga más estructura; *e.g.*, que de hecho conforme un espacio métrico, como se suele especular en el tratamiento de las condiciones de verdad de los contrafácticos.) La representación gráfica resultante tendrá obviamente una geometría invariante bajo el cambio de índices.

<sup>124</sup>Para *saber que sabemos* que Juan podría estar en  $I$ , basta saber que (i) la etiqueta está en  $I$  en  $w_1$ , y (ii) que la etiqueta se ha quedado pegada a Juan en  $w_1$ . Por supuesto, dada la no-reducción de la modalidad a mundos, (i) y (ii) ya presuponen saber que la etiqueta podría estar en  $I$  y que se le podría haber quedado pegada a Juan en esa misma posibilidad—como lo muestra este mismo ejemplo, cuantificar sobre mundos es un recurso heurístico muy útil.

La etiqueta que rastree los estados posibles de Juan—sus estados en cada mundo—nos darán, bajo esta idea, infinita información: para cada cualidad, la etiqueta nos diría “Sí” o “No”, dependiendo de si Juan tiene tal cualidad; para cada cantidad, la etiqueta nos diría la magnitud en que Juan la tiene.

Por supuesto, no poseemos tales etiquetas de hecho. Pero este ejercicio nos muestra la generalidad en principio de las ideas que hemos venido manejando.

Quitemos las comillas. ¿Qué es “seguir” la etiqueta?

En el caso de Juan moviéndose (o cambiando de estado, en general) a través del tiempo, la idea era suficientemente clara: seguimos la etiqueta como seguimos cualquier otra cosa en el tiempo. Por ejemplo, si la etiqueta fuera fluorescente y estuviésemos a oscuras, para rastrear la trayectoria de Juan simplemente seguiríamos con la vista la etiqueta de Juan. O, si cerramos los ojos pero la etiqueta tiene un olor muy fuerte y distintivo (y tenemos *muy* buen olfato), podríamos seguir la etiqueta con el olfato. Mecanismos más complicados son posibles en principio: si estuviésemos muy lejos pero la etiqueta estuviera conectada por satélite a un mapa digital, podríamos establecer coordenadas en ese mapa y seguir la trayectoria de Juan en ese mapa (como en un GPS, digamos). Cuando se trata del cambio en el tiempo, entonces, no parece haber grandes problemas de principio para seguir a la etiqueta que nos representa a Juan. (De nuevo, estoy asumiendo todavía que tenemos etiquetas que se pegan *perfectamente*—que siguen a Juan en el 100% de su trayectoria.)

La cosa parece cambiar radicalmente cuando intentamos seguir la etiqueta de Juan a través de los mundos posibles. Es un hecho incontrovertido que no vemos, olemos, o tenemos GPS para los mundos posibles. Cuando Juan en  $w_1$  *podría haber estado* en  $I$ —lo cual lo representamos como Juan *estando* en  $I$  en algún mundo, que llamamos “ $w_2$ ”—no podemos corroborar eso de la misma manera en que podemos corroborar que Juan *estará* en  $e_3$ , cuando está en  $t_1$ .

Así que necesitamos un cambio de perspectiva. No podemos ver a Juan en sus estados posibles, y tampoco a su etiqueta en ellos. Ahora simplemente tenemos etiquetas que le pegamos a Juan, y las movemos en el espacio de cualidades, a través de los mundos posibles. La noción misma de “mundo posible” es representacional—cada mundo posibles es una enorme etiqueta, representando posibilidades. (Para eliminar, por el momento, esta segunda capa representacional, hablaré de posibilidades.)

No tenemos un acceso directo a las “trayectorias” de Juan en el espacio de cualidades, relativas a las posibilidades. Sólo podemos asumir que tenemos una etiqueta, y luego estipular que esa etiqueta “se moverá” por el espacio cualitativo, informándonos de los varios estados posibles de Juan. Si tenemos suerte, la trayectoria de esa etiqueta—de esa representación—por el espacio cualitativo, a través de las posibilidades, se corresponderá con la posible trayectoria de Juan en el espacio cualitativo a través de las posibilidades.

Pero no queremos suerte, porque no nos daría conocimiento. Además, es poco informativo y poco práctico, además de probablemente imposible en la práctica, estipular todas las trayectorias posibles de Juan. El mecanismo básico (usamos una representación de Juan para “moverlo” a través de las posibilidades) es interesante por sí mismo, pero sin confirmación independiente, sólo nos queda la suerte. Trataré con este problema abajo.

Un segundo y obvio problema es la noción de “mover” la representación de Juan a través de las posibilidades.

Una idea es simplemente pegar etiquetas a los objetos y esperar a que ellas “nos digan” cosas sobre las posibilidades de Juan. Pero ¿cómo revisamos lo que las etiquetas “nos dicen”? Para que las etiquetas representen a Juan, primero tuvimos que etiquetar a Juan, y luego asumir que lo que a ellas le pase, va a representar lo que le pasa a Juan. Así, en el caso de las playeras, asumimos que los estados de las etiquetas nos informan, “nos dicen”, sobre los estados de Juan. Cuando una etiqueta está en un lugar *e*, estipulamos que eso “nos dirá”, que eso *representará*, a Juan estando en *e*. Cuando una etiqueta está en el estado “rojo”, estipulamos que eso representará a Juan teniendo una playera, que además es roja.

Pero en el caso de los cambios de estado de Juan, de nuevo, pudimos saber cómo las etiquetas representaban a Juan, como *rastreaban* sus estados, al tener un acceso a las etiquetas mismas. Si teníamos medios independientes para contrastar el estado de las etiquetas con el estado de Juan, teníamos medios para calcular lo bien que las etiquetas se pegaban a él, y con ello lo bien que los estados de ellas representaban los estados de Juan.

Pero muchas veces no tenemos esos medios. Como en el caso del objeto en la oscuridad con una etiqueta fluorescente, muchas veces sólo nos queda simplemente fiarnos de lo que nos dicen las etiquetas. Quizá en otras ocasiones, pasadas o futuras, podremos contrastar a las etiquetas con Juan. Quizá no.

En todo caso, la pregunta permanece de cómo tuvimos acceso a las etiquetas de Juan en los mundos posibles. Simplemente no pudimos ver u oler u oír los estados de las etiquetas en tales mundos posibles, y tomar a esos estados como representando los estados de Juan (con la confianza que pudiéramos tener en esa representación).

Es natural, o al menos tentador, asumir que el acceso a esos estados posibles de las etiquetas no es empírico y que, por ello, salvo escepticismo, tendremos que confiarnos en un acceso racional. Nos hemos puesto la carga explicativa de mostrar cómo es que, sin apelar directamente a los sentidos, podemos saber algo *sobre Juan*—que podría estar en tales o cuales estados—a través de las etiquetas que le pusimos: a través de las representaciones que tenemos de él.

Surge la idea de tratar de resolverlo apelando a la naturaleza de las etiquetas. Decimos, por ejemplo, que aunque hay etiquetas muy malas que se despegan fácilmente, hay etiquetas—o eso esperamos—que permanecen siempre pegadas a su objeto. Tomamos esas etiquetas. Pero no podemos rastrear sus estados de manera empírica. Entonces usamos alguna *simulación*—que puede ser algún tipo de imaginación, concebibilidad, intuición, etc.—de los estados de la etiqueta. Si la simulación es “buena”, y si la etiqueta es del tipo que permanece pegada a Juan, podemos pasar a tener la seguridad de que la etiqueta representará los estados posibles de Juan.

Pero ¿qué es que la simulación sea “buena”? De nada nos servirá simular—imaginar, concebir, intuir—que la etiqueta podría estar en un estado, si *de hecho* no podría. Esto incluso aunque la etiqueta sea del tipo que siempre permanece pegada. Aún con tales etiquetas de pegamento perfecto, si simulamos que la etiqueta podría estar en *I* pero resulta que Juan en ningún mundo posible está en *I*—por poner un caso simple—entonces será incorrecto concluir que Juan podría estar en *I*. La pregunta surge de nuevo: ¿cómo vamos a *asegurarnos* que la simulación es “buena”?

Parece que estamos en un caso muy parecido al objeto que está a oscuras, pero que tiene una etiqueta fosforescente. En ese caso, dado que no podíamos comprobar directamente en dónde estaba el objeto, sólo nos quedaba confiar que la etiqueta le había permanecido pegada, de tal ma-

nera que seguir la trayectoria fosforescente, asumíamos, nos servía para seguir la trayectoria del objeto. Ahora también estamos en otro tipo de oscuridad. No podemos ver directamente al objeto en sus mundos posibles. Sólo nos queda seguir la trayectoria de la etiqueta en las simulaciones, y confiar que se le ha quedado pegada al objeto, y que la simulación de hecho representa estados posibles de la etiqueta.

No siempre tenemos acceso directo a los objetos, y muchas veces sólo tenemos acceso indirecto a ellos cuando las etiquetas que les pusimos permanecen apropiadamente pegadas a ellos. Esto lo muestra el caso de la oscuridad, y el caso de los mundos posibles. Pero si no somos escépticos acerca del conocimiento de objetos en la oscuridad, *esto*, todavía, no nos da razón para ser escépticos acerca del conocimiento de objetos en sus posibilidades. Quizá no tenemos acceso directo, cierto. Pero todavía podríamos tener acceso indirecto.

Debemos preguntarnos si el acceso indirecto a las posibilidades de los objetos es mediante algún tipo de simulación—nos *imaginamos* a los objetos, o los *concebimos*, o los *intuimos*, o algún otro proceso mental, en algún estado. Puede resultar, se piensa, que tales simulaciones de hecho sean etiquetas bien pegadas a los objetos, y con ello representarán sus estados posibles. Si no, no hay razón todavía para el escepticismo: tendremos un caso análogo al objeto en la oscuridad cuya etiqueta fluorescente simplemente se le despegó, por lo que la trayectoria de esta no representaba la trayectoria del objeto.

Una vez que hemos aceptado que no tenemos un acceso empírico a las posibilidades de los objetos, pero que de hecho tenemos algún tipo de acceso, la salida que nos queda es aceptar que no hay ningún canal de información entre esas posibilidades y nuestra mente, pero que aún así podemos conocerlas. Y podemos conocerlas *simulando* sus propiedades—siempre que esas simulaciones rastreen adecuadamente a las posibilidades.

No creo que sea tarea propiamente filosófica especular sobre la naturaleza particular de tal simulación. Claramente poseemos un tipo de proceso mental así, por lo que el escepticismo está fuera de lugar; pero aclarar cómo funciona de manera más fina es tarea experimental. Sin embargo, podemos hacer algunos comentarios epistemológicos sobre la base de la estructura de tal proceso (ver lo apartados siguientes).

Regresemos al objeto en la oscuridad con una etiqueta fluorescente. Como no podemos verlo, sólo podemos confiar en que la etiqueta se le haya quedado pegada—de tal manera que la trayectoria de la etiqueta sea un buen indicador de la trayectoria del objeto. La simulación de posibilidades funciona como una etiqueta así. Esperamos que la manera en que representamos los estados del objeto—en que simulamos que es—de hecho esté “pegada” a los posibles estados del objeto. Pero no tenemos manera de comprobarlo directamente—o, en la metáfora, de prender la luz y ver si la etiqueta ha permanecido pegada en el objeto.

Queda todavía la pregunta de si la simulación es, en un sentido sustantivo, *puramente racional*. La pregunta es si podemos tener conocimiento de las posibilidades de los objetos sin depender de manera alguna del conocimiento empírico. La respuesta que defenderé es que no: no hay buenas razones para pensar que el acceso a las posibilidades de los objetos puede, o deba, ser puramente racional—si es que tal acceso nos brinda conocimiento.

Es hora de volver a resquebrajar lo que Chalmers llama el “Triángulo Dorado”, que el creyó restaurado, y que conecta a la modalidad con el significado y la razón. Para ello, es necesario comenzar a dejar atrás las metáforas sobre etiquetas y a hablar literalmente.

## 6.2. tipos de representación

Representations are not in general intended to say how it is. They can be portrayals or delights. After our recent obsession with words it is well to reflect on pictures and carvings. Philosophers of language seldom resist the urge to say that the first use of language must be to tell the truth. *There should be no such compulsion with pictures.*

—Ian Hacking, *Representing and Intervening*

Una característica de los seres con mentalidad, es que podemos *representarnos* al mundo.<sup>125</sup> Los seres humanos nos representamos al mundo mediante conceptos, ideas y demás *representaciones mentales*, también mediante creencias, conocimiento, y demás *estados intencionales*. (Muy probablemente, también algunos animales no-humanos tienen representaciones mentales: Strevens 2004). También usamos banderas, palabras, signos escritos, sistemas de ecuaciones, gestos, señales de humo, etcétera, para representarnos al mundo. Se ha hecho (Greenberg 2011: §0.1) una distinción muy general entre dos maneras de representar: la *básica* y la *derivada*. La idea es que las banderas, señales de humo, gestos, etc., cuentan como representaciones de manera *derivada*, porque, en el orden explicativo, primero nos representamos mentalmente y *después* nos representamos con palabras u otros signos: si nouviésemos las capacidades y medios mentales para representarnos, no podríamos utilizar otros medios físicos—palabras, gestos—para representar otros estados. Así, *la manera básica* de representarnos es la mental, y todas las otras son *manera derivadas* de representarnos.<sup>126</sup>

Se dice (Greenberg, *loc. cit.*) que una representación mental es *privada*, pues ocurre en la privacidad de nuestra mente, y que una representación derivada es *pública*, pues es, al menos en principio, accesible directamente por otros sujetos. Así, una hipótesis es que la representación básica es siempre privada, y la derivada es siempre pública.<sup>127</sup>

Es muy plausible suponer que es *suficiente*, para que algo sea una representación pública, que

125 Es al menos epistémicamente posible que cierta variedad de *pansiquismo* sea verdadera, de acuerdo al cual todo individuo en el mundo tiene cierto tipo de mentalidad, aunque no todo individuo tenga mentalidad representacional. La tesis es que todo individuo tiene *qualia*, o mentalidad cualitativa, lo que se ha llamado "*what-is-it-likeness*" (*cf.* Llamas *manuscrito*). En este caso, la tesis de arriba se frasearía así: "Una característica de los seres con mentalidad, *al menos aquellos cuya mentalidad no es sólo cualitativa*, es que podemos *representarnos* al mundo". Esto hace que la afirmación sea neutral acerca de si el *what-is-it-likeness* tiene una función representacional.

126 Como nota Greenberg, la distinción no es absolutamente obvia. Regularmente, se asume que las proposiciones *representan* estados de cosas, pero también se suele asumir que no lo hacen en virtud de los estados mentales de nadie. Estas nos darían un caso de representación básica—no basada en la mente de nadie—pero que no es mental. La pregunta de si eso es, ultimadamente, coherente, es una pregunta que me parece interesante, pero que cae fuera de este trabajo.

127 La hipótesis no es completamente obvia, de todos modos. Si uno acepta un externismo al estilo de Burge (1979), Putnam (1975) o Heidegger (1927), muchas representaciones mentales (al menos muchos conceptos) resultan estar determinadas por aspectos sociales—que son típicamente públicos (bajo esta caracterización). Aquí no trataré este interesante tema.

sea usado con la intención de representar. Es así que podemos ver cómo el estatus de representación pública es fijado por la representación mental—por las *intenciones* (cf. Grice 1957). También, esto muestra que tal estatus es *funcional*: que algo sea una representación pública es que cumpla una función—ser usado con intenciones representativas—y *no* que tenga tales o cuales cualidades materiales. Así, diferentes medios pueden ser usados como representaciones: una palabra escrita o hablada pueden servir como representación de un mismo tipo de objeto, por ejemplo, aunque cada una tenga una base material distinta. Además, muestra que tal estatus es esencialmente *teleológico*: una cosa es una representación si se la usa *para* representar, aunque no alcance ese fin (cf. Priest 2006: §4.5, para quien las aseveraciones tienen como *telos* el decir la verdad). Por ejemplo, la palabra “fantasma” es una representación, aunque no logre representar a los fantasmas—pues no hay tales cosas—debido a que se la usa como dirigida hacia cierto fin. (Según la hipótesis de Dretske (2003), también la representación mental es esencialmente teleológica. Aquí no podré decir nada al respecto de esta cuestión.)

Pero no sólo nos representamos lo que *es*—lo que sucede—sino también lo que *no es*—lo que *podría ser* e, incluso, lo que *no podría ser*. La oración “Carlos, un buen día, se levantó a las cuatro de la mañana sintiendo sed, pero ese mismo día, exactamente una hora después de las tres de la mañana, se levantó sin sentir sed” representa un estado de cosas imposible—el estado imposible en el que yo tanto tengo como no tengo sed al despertarme a las cuatro de la mañana. Se dice también (Bueno & French 2011: §3), por ejemplo, que las ecuaciones de Maxwell en la teoría del electromagnetismo representan estados contradictorios del mundo. Y la descripción definida “El hombre gordo y delgado en la puerta de Quine” representa una entidad imposible. Digamos que una representación es *fáctica* cuando representa al mundo como es; que es *posibilista* cuando representa al mundo como podría ser, y que es *imposibilista* cuando representa al mundo como no podría ser.

Notemos además que la distinción entre la representación y su representado es conceptual, pues en algunos casos una cosa puede fungir como representación de sí misma: en la película *Traffic* (Stephen Soderbergh, 2000), por ejemplo, Julia Roberts hace el papel de... Julia Roberts. Así, Julia Roberts se representa a sí misma en la película: es un caso donde representación y representado son la misma cosa. En las paradojas por autoreferencia, por poner otro caso, una oración habla de sí misma (Kripke 1975): una oración puede fungir como representación de sí misma, que se representa teniendo cierta cualidad (la de ser falsa, por ejemplo).<sup>128</sup>

Esto nos lleva a hacer la primera distinción estructural. Ya sea con representaciones mentales o públicas, podemos decir con bastante generalidad que una representación representa a su representado *como siendo de cierta manera*. Una fotografía representa un evento, una “escena” en el mundo; Julia Roberts en *Traffic* se representa a sí misma como teniendo cierta historia; una aseveración (una “*assertion*”) representa cierto estado de cosas o hecho, que puede o no darse en el mundo.

Un caso que parece resistirse a esta tesis es el caso de los nombres propios: según la tradición milliana revivida por Kripke (2005), el único valor semántico de los nombres es su referente (según otros millianos como Salmon (1986), los valores semánticos tradicionalmente—o en la tra-

128 Creo que una adecuada teoría de la representación puede darnos *insights* interesantes en el problema de las paradojas por autorreferencialidad. Esta conjetura, desafortunadamente, queda por ser explorada en otro lugar (véase mi “The Paradoxes of Semantics and the Metaphysics of Representation”, en elaboración).



dición fregeana—atribuidos a los nombres son en realidad características *pragmáticas*). Suponiendo la tesis milliana, parecería que los nombres—a nivel semántico—no representan a los objetos como siendo de cierta manera, sino que simplemente representan objetos: “Carlos” no me representa a mí como siendo de cierta manera (a nivel de la semántica), sino que simplemente se refiere a mí.

Pero podemos acomodar la tesis milliana dentro de nuestra tesis general, si asumimos que *referirse* es una manera de representar. Cuando la única función de una representación es referirse a su *representado intencional* (aquel objeto que se busca representar), en lugar de distinguirlos como teniendo tales y cuales propiedades cualitativas, podemos decir que en esos casos se representa al referente *como siendo idéntico a sí mismo*. Por ejemplo, asumiendo el millianismo, mi nombre “Carlos” no me representa semánticamente como teniendo cierta cualidad u otra (a lo más, esto pasaría a nivel pragmático). Pero el nombre “apunta” a su referente—lo selecciona, por así decir. Su función semántica, la de *referir*, puede verse como *recogiendo* a su referente de entre los demás objetos (siguiendo en la asunción milliana, puede que el usuario del nombre no sepa qué cualidades tenga tal referente, pues ellas no forman parte del significado del nombre). Pero este “recoger” también puede verse como individuación que no se hace mediante cualidades—o al menos, no directamente. Se recoge al referente como aquel objeto que sea idéntico al referente intencional, tenga este las cualidades que tenga. Así, mi nombre, “Carlos”, me representa a mí *como siendo yo mismo* (y *no* como llamándome “Carlos”, sino como siendo yo: Carlos).

Entonces, en general, una representación representa a su representado como siendo de cierta manera (incluso de maneras no cualitativas). Esto no quiere decir que para *usar* una representación, *r*, tengamos que saber de antemano cómo es que *r* representa a su representado intencional. Por ejemplo, la oración “El número de páginas de la tesis de Carlos tiende peligrosamente a infinito”, representa al mundo como siendo de cierta manera—pero para comprenderla, y para evaluar si es verdadera o no—si representa o no correctamente al mundo—no necesitamos saber *completamente*, de entrada, la manera en que el mundo sería si fuera representado por ella. El argumento para esto consiste en apelar al bien conocido problema de Frege en “Sentido y Referencia” (1892/1997). Puede que los hechos macroscópicos, como el de mi tesis y sus páginas, sean idénticos a ciertas composiciones de hechos microscópicos. Pero si eso resultara ser el caso, parece todavía incorrecto decir que entender la oración en cuestión *es* saber cómo serían ciertos hechos microscópicos, y cómo estarían combinados. Es difícil ser más claro aquí sin usar terminología parcial y comprometedor, pero me lo permitiré: parece que al usar o comprender la oración, sólo necesitamos conocer al mundo bajo un “modo de presentación”—aún si eso que se nos presenta (los hechos microscópicos) no nos es completamente conocido.

Ya vamos caracterizando los primeros componentes para la relación que nos interesará en lo que resta del capítulo: la relación de *representar-como-bajo cierta convención*, o *RepAs*. Esta relación tiene cuatro lugares: uno para la representación, uno para el representado intencional o *contenido*, uno para *el aspecto* bajo el cual la representación representa al representado, y uno para la *convención* representacional (de la que hablaré abajo). Así, la relación objeto de lo que llamo “teoría general de la representación”, es:

$$\text{RepAs}(r, o, A(o), C)$$

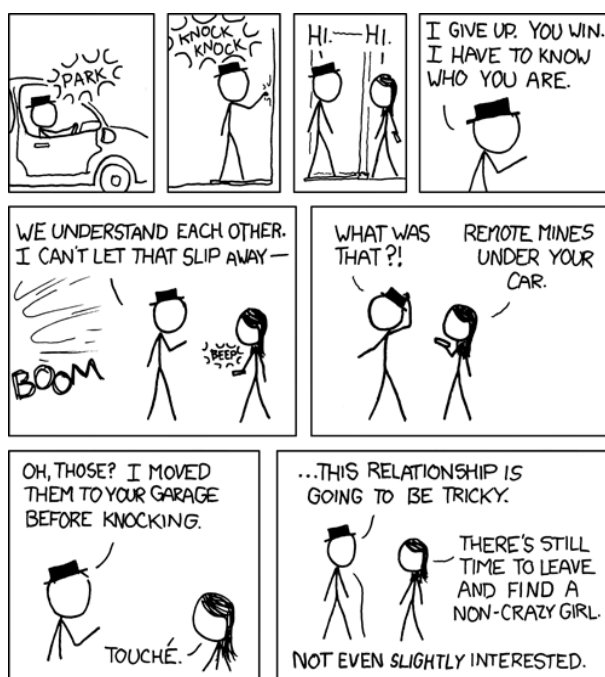
Antes de hablar de las convenciones representacionales, es importante tratar un poco con los

*objetivos, los fines, el telos, de las representaciones, para ayudar a caracterizar la relación RepAs.*

Una idea muy intuitiva es que, entre las *varias* funciones o *telos* de las representaciones, *una muy importante* es representarnos para *informar(nos)* acerca de cómo es el mundo, para *transmitir información* acerca de él. Los modelos de los científicos buscan representar al mundo, aunque intencionalmente contengan idealizaciones, y por ello no lo describan a la perfección. Las oraciones del lenguaje natural buscan describir al mundo, aunque puedan ser inexactas. Un diagrama o dibujo puede también buscar representar al mundo, aunque no sea completamente exacto al retratar.

Una tradición en filosofía de la ciencia enfatiza la función representacional, informacional, de las teorías y modelos científicos, mientras que otra enfatiza sus funciones prácticas (Bueno & French 2011; Chakravartty 2010). De acuerdo a la segunda tradición, la función esencial de las representaciones es brindarnos la capacidad de hacer inferencias acerca del mundo, lo cual presupone cierto dominio de esas representaciones. Algunos han creído que las dos funciones—las de describir al mundo y la de permitirnos manejarlo, a partir de las inferencias que hacemos sobre él—son opuestas, pero, de hecho, pueden convivir muy bien (Bolinska 2013; Chakravartty 2010). Ambas funciones pueden verse como complementarias. Si las representaciones pueden usarse para un fin práctico, es porque logran representar de manera suficientemente fiable al mundo; y es porque tienen un fin práctico que buscamos ajustarlas al mundo.

En adelante me enfocaré en la función informacional de las representaciones, y asumiré (lo cual a mí me parece obvio) que toda representación se usa *para* representar—aunque, plausiblemente, ésa no sea su *única* función. (Y aunque no toda cosa que cuenta como representación es *siempre*, mucho menos *necesariamente*, usada con la intención de representar. Podemos usar



las palabras para hacer metáforas—y tal función del lenguaje, como muchos creen, no siempre tiene como fin el describir al mundo (a menos que en esos casos, las palabras representen otras cosas—por ejemplo, las impresiones del poeta). Podemos usar una bandera para limpiarnos la nariz; o una maqueta para hacer fogatas. En esos casos, el objeto no necesita contar como representación—debido a la naturaleza teleológica de esta.)

Ahora diré algo sobre la noción de *convención representacional*, el cuarto lugar de la relación *RepAs*. Tomemos como ejemplo una tira del webcomic *xkcd* (Randall Munroe, <<http://www.xkcd.com/433/>>). Es claro que representan esos círculos y líneas: representan seres humanos. Lo que quiero notar ahora es que eso es, en parte, gracias a una *convención representacional*.

Una convención representacional simplemente es una *convención* que establece cómo vamos a *interpretar* ciertos objetos *que se toman como representaciones*. Aquí simplemente voy a asumir que tenemos un entendimiento suficientemente claro, aún si no *total*, de lo que es una con-

vención (*cf.* Lewis 1979). De una manera vaga pero suficientemente clara, una convención es un acuerdo, implícito o explícito, de realizar ciertas acciones dadas ciertas condiciones. Las convenciones que nos interesan aquí son las convenciones *representacionales*—que tomo, de nuevo de manera no completamente exacta, como convenciones para *usar* ciertos objetos como representaciones. Tales convenciones pueden estar funcionando incluso aunque haya *determinaciones* para representar o no, o incluso para usar una clase limitada de objetos como representaciones. El caso más claro es el del lenguaje natural: aunque claramente tenemos *determinaciones biológicas* para usar ciertos objetos—sonidos emitidos por la boca, por ejemplo—como representaciones, aún así es teóricamente útil contar al lenguaje natural como teniendo un importante componente convencional (Lewis 1979, Greenberg 2011: 6-7; para una interesante manera en que el lenguaje puede no ser convencional, véase Ramachandran & Hubbard 2001: 18-23).

Las convenciones son esenciales para la representación derivada (pública). Sin ellas, un círculo en una pizarra es *sólo* un círculo en una pizarra; con ellas, el círculo puede representar al Sol—o al cero. Así, una primera función de las convenciones son (1) establecer que un objeto *ha de usarse* para representar. Las convenciones traen representaciones al mundo—excepto cuando tratamos con la representación mental (o proposicional). Pues esta se suele tomar como totalmente determinada por la interacción entre nuestro equipamiento cognitivo—determinado a su vez por una historia evolutiva—y el mundo fuera de él (y también *dentro* de él, pues el mismo equipamiento cognitivo puede representarse a sí). Así, una hipótesis plausible es que la representación mental es no-convencional: no hicimos ningún acuerdo para tomar a tales o cuales estados mentales—que son, con mucha probabilidad, sucesos físicos—como representaciones de otros sucesos físicos. Dado que la representación mental es *básica*—toda otra forma de representación (quizá excepto por las proposiciones) es derivada de esta—y asumiendo que al menos en un sentido importante es *privada*—asumiendo que los agentes cognitivos tienen un punto de vista *en primera persona* de sus representaciones mentales, único de cada uno de ellos—tenemos que la representación mental es *básica, privada y no-convencional*.<sup>129</sup> Estas características arrojan cierta luz, creo, sobre la representación mental, y *a fortiori* sobre la noción general de *representación*.

Una segunda función de las convenciones representacionales es (2) establecer *qué* se busca representar—cuál es el representado intencional. Y, finalmente, una tercera función de las convenciones representacionales es (3) establecer *cómo* se busca representar. Por supuesto, parte de esto ya es logrado por la selección de *medios*: el uso de palabras establece que la representación será lingüística; el uso de fotos que será pictórica. Pero las convenciones establecen otros parámetros. Por ejemplo, establecen *estándares de precisión*—o al menos conjuntos de ellos: es porque tales estándares están en juego, que uno puede decir con verdad que “Francia es hexagonal” (en el famoso ejemplo de Austin, citado en Longworth 2012: §2.2) en ciertos contextos; aunque un geógrafo dictando una conferencia que dijera eso bien podría ser tachado de decir falsedades (o de

---

129 Tengo algunas sospechas al tomar a la representación mental como *totalmente* privada, pues eso pone barreras *a priori* para posibles descubrimientos futuros de las ciencias de la mente-cerebro. (Nada nos dice que en un futuro no podamos *percibir* los pensamientos de otras personas gracias al desarrollo de la tecnología, por ejemplo.) Los filósofos ya deberíamos haber aprendido a no poner tales restricciones a lo que el mundo podría decirnos una vez que hayamos desarrollado los medios para entenderlo (teorías, experimentos y tecnología), pero aquí seguiré la corriente, pues mi tema no es la representación mental.

ser inexacto, o en el peor de los casos, de decir sinsentidos). Es también por ello que un círculo con líneas puede representar a un ser humano, aunque el grado de parecido entre ambos tipos de cosas sea mínimo.

Esto nos lleva a un segundo componente de la tercera función de las convenciones, además de los estándares de precisión. Es *la manera en que se evalúan tales estándares*. Por poner un ejemplo particularmente claro, regularmente se cree que las fotografías o pinturas realistas retratan una escena al *parecerse* a ella (Greenberg 2011: §§2.1, 2.2). Así, dada cierta imagen *i*, escena *e* y manera en que *e* podría ser,  $A(e)$ , podría haber una convención *C* que permitiría interpretar a *i* como representando a *e* como siendo  $A(e)$ , tal que los estándares de precisión determinados por *C* permitirían decir que *i* se *parece lo suficiente* a  $A(e)$  como para representar a *e*. Ahora, no es claro que este sea el criterio que de hecho usamos (Greenberg 2011: §§2.3, 2.4). Según Greenberg (cap. 3), la manera en que se evalúa si una representación pictórica retrata a su referente intencional es mediante *proyecciones geométricas*, más que parecidos.

No juzgaré aquí entre ambas posturas. Lo que me interesa notar es cómo las convenciones representacionales incluyen no sólo estándares de precisión—con los que se evalúa qué tanto una representación gobernada por tales convenciones tiene que “ajustarse” (en un sentido que haremos más preciso abajo) a su representado intencional para poder decir que lo representa. También incluyen criterios que dictan cómo se evalúa si la representación satisface esos estándares. En la representación pictórica, como vimos, dos competidores eran el parecido y la proyección geométrica. Pero en la representación lingüística también hay competencias. Por ejemplo, para que un término singular refiera a su referente intencional, según Kripke (2005), tiene que haber una cadena causal desde el referente hasta cada usuario del término. Pero según Frege (1892/1997), el referente debe satisfacer ciertas propiedades asociadas esencialmente con el término. (Estoy caricaturizando un poco, por supuesto; por el momento, sólo uso a ambas teorías como ejemplos.) Lo mismo podría pasar con las oraciones declarativas: para que una de ellas represente a su representado intencional—que, asumo, es un hecho o estado-de-cosas—el significado de la oración, según los correspondentistas, tiene que *corresponder* al hecho (de manera suficiente para satisfacer el estándar de precisión). Pero según los convencionalistas, basta que la oración satisfaga las convenciones que los hablantes tienen en el contexto en el que se emite la oración.

Creo que esto nos da una buena base—seguramente no exhaustiva, pero suficiente para un primer acercamiento exploratorio—para concluir que las convenciones representacionales tienen tres funciones: (1) hacer que un objeto *cuenta* como representación, (2) hacer que un objeto cuente como el representado intencional de la representación en cuestión, e (3) implicar (a) estándares de precisión para el “ajuste” de la representación, y (b) maneras en que tal estándar ha de satisfacerse en usos concretos de la representación en cuestión.

Así, las convenciones en cuestión codifican tanto la *meta-semántica*—aquellos hechos que hacen que un objeto sea una representación de otro—como la *semántica*—aquellos hechos que determinan cuándo el objeto representa al otro—de las representaciones.

### 6.3. la semántica general: condiciones de adecuación representacional

La experiencia del ente deviene esencialmente representación. *Representación* deviene la categoría más general para determinar la aprehensión de cualquier cosa que concierna o interese en una relación cualquiera. Todo el discurso poscartesiano e incluso posthegeliano, si no justamente el conjunto del discurso moderno, recurre a esa categoría para designar las modificaciones del sujeto en su relación con un objeto. La gran cuestión, la cuestión matricial, es entonces para esta época la del valor de la representación, la de su verdad o *adecuación a lo que representa*.

—Jacques Derrida, “Envío”

¿Qué es esta noción de “ajuste”? Ahora le voy a quitar las comillas. La idea es que las funciones (1)-(3) de las convenciones representacionales determinan *condiciones de adecuación* de las representaciones. El “ajuste” de una representación a su representado es en realidad la *satisfacción de las condiciones de adecuación* de la representación.

Empecemos por ejemplos básicos. Un nombre propio se “ajusta” a su representado intencional, o satisface sus condiciones de adecuación, cuando se *refiere* a él. Una oración declarativa satisface sus condiciones de adecuación *cuando es verdadera*. Un sistema de ecuaciones matemáticas satisface sus condiciones de adecuación *cuando describe a la clase de sistemas* que son su referente intencional—aunque lo haga de manera *idealizada*, pues tales idealizaciones se hacen claras en la convención representacional. Una fotografía satisface sus condiciones de adecuación cuando describe el evento que es su “escenario intencional”.

Como vemos, la noción de *condición de adecuación* es una generalización de las nociones de *condición de referencia* o *condición de verdad*. Vamos a continuar con esta generalización. Regularmente se toma a las condiciones de verdad—de adecuación—de una oración declarativa  $S$  como el conjunto de mundos donde  $S$  es verdadera. Así, podríamos tomar las condiciones de adecuación de los nombres propios millianos—sus condiciones de referencia—como el conjunto de mundos donde su referente es idéntico a sí mismo, es decir, donde existe (asumiendo que los nombres millianos son *rígidos*). Las condiciones de adecuación de un modelo matemático de un tipo de sistema físico podrían ser la clase de mundos donde existen sistemas que son descritos (al nivel especificado de detalle e idealización) por tal modelo. Las condiciones de adecuación de una fotografía serían el conjunto de mundos donde hay un evento como el que la fotografía pretende describir (dependiendo del detalle de la foto y las convenciones representacionales en juego, tal conjunto puede ser de mayor o menor tamaño).

Esta es la generalización del acercamiento estándar, que está profundamente influido por la semántica tarskiana para la lógica de primer orden, ya sea clásica o modal. Así, la teoría de las condiciones de adecuación para las representaciones, sería lo mismo que lo que Greenberg llama “semántica general” (2011: §0.2) que es simplemente la idea de generalizar las semánticas al estilo Tarski y de condiciones de verdad para representaciones que no sean necesariamente lingüísticas, como las representaciones pictóricas (pero que sean públicas, gobernadas por reglas, y

convencionales). Tal generalización asignará, *sistemática y productivamente*,<sup>130</sup> un valor semántico a cada representación (pública, gobernada por reglas, y convencional) de tal manera que las reglas reflejen el constreñimiento de que *las representaciones consideradas para tales valores semánticos se usen para transmitir información sobre esos valores semánticos, bajo las convenciones relevantes*.

La idea es más o menos la siguiente. Una semántica para un sistema representacional  $S$  (e.g., el sistema del lenguaje natural, o el de representación pictórica), bajo una convención  $C$ , será algo muy cercano a la teoría de modelos tarskiana. Simplemente, para cada sistema representacional  $S$ , la semántica asignará estructuras con la forma de un par  $\langle D, I \rangle$ , donde  $D$  es el dominio de objetos que el sistema  $S$  puede representar, e  $I$  es la función de interpretación.  $D$  puede tener un simple conjunto, o tener más estructura—como las estructuras relacionales que en el capítulo 1 vimos para el lenguaje modal, y que propuse modificar de la manera especificada en el capítulo 3.  $I$  codifica a las convenciones representacionales para  $S$ , de tal manera que esas convenciones sean las que determinan las condiciones de adecuación intencionales para  $S$ . Por ejemplo, son nuestras convenciones lingüísticas las que determinan las condiciones de referencia de un nombre: esas convenciones pueden incluir que el referente satisfaga cierta descripción o cúmulo de ellas (si el descriptivismo está en lo correcto), o que cada usuario del nombre esté al final de cierta cadena causal conectando sus usos del nombre con el referente (si la teoría causal está en lo correcto), o quizá una mezcla de las dos condiciones, o quizá otra condición distinta. Por poner otro ejemplo, si la teoría del retratar (“*depict*”) como parecido (“*resemblance*”) está en lo correcto (Blumson *manuscrito*), la convención que gobierna a los sistemas simbólicos de los retratos es, básicamente, que las representaciones en cuestión se *parezcan* a sus referentes—diferentes convenciones, o aplicaciones de ellas en diferentes contextos, pueden resultar en diferentes grados de parecido necesarios para la referencia. Esto haría que la función  $I$ , para una representación en un sistema de *depiction*, estuviera relacionado con sus valores en  $D$  (es decir, para  $r$  una representación de tal sistema,  $I(r) \in D$ ) mediante una regla que se extrae de las convenciones que determinan las condiciones de adecuación.

Así, por ejemplo, un fragmento de tal *semántica* para un fragmento de un lenguaje  $\mathcal{L}$ , sería el siguiente:

Cuando los términos singulares de  $\mathcal{L}$  son  $t_1, \dots, t_n$ , el referente de  $t_i$  es  $o_i$  sii \_\_\_\_ (aquí se ponen las condiciones que un término tiene que satisfacer para referirse a su referente intencional),

Cuando los predicados de  $\mathcal{L}$  son  $P_1, \dots, P_n$ , el referente de  $P_i$  es  $\Pi_i$  sii \_\_\_\_ (aquí se ponen las condiciones que un predicado tiene que satisfacer para referirse a su referente intencional),

Etcétera.<sup>131</sup>

130 La idea es que la teoría asignará valores semánticos a las representaciones mediante la aplicación de reglas recursivas. Así, en principio, un conjunto recursivamente enumerable de reglas permitirá asignar contenidos semánticos a un conjunto potencialmente infinito de representaciones (probablemente, de manera composicional: es decir, el valor semántico de una representación compleja dependerá funcionalmente del valor de sus componentes más el modo de arreglo de estos).

131 Inicialmente parece que tal semántica inducida va a ser, en la terminología de Fine (2007), *intrinsicista*—es decir, la semántica dará, para cada término del lenguaje, su valor semántico sin hacer referencia a los demás términos del lenguaje. Hay que subrayar *no* necesita serlo. La semántica bien podría resultar *relacionista*, al estilo

No me quiero comprometer con la idea de que la semántica general *tenga* que tener una forma tarskiana—aunque ciertamente la idea es altamente atractiva. Existen varias alternativas a la semántica tarskiana: en la nota 131, por ejemplo, recordé la alternativa Fineana; pero también podríamos tomar la alternativa de Katz (2014). La hipótesis de la semántica general, como yo la acepto, sólo se compromete con la idea de que una teoría semántica puede ser suficientemente general como para tener, como especies, a la semántica del lenguaje natural—pero también a la semántica de las representaciones pictóricas, por ejemplo—y en general *a toda representación gobernada por principios*. (Nótese como ni siquiera me comprometo con que la representación mental no sea capaz de tener una semántica, por más “privada”, básica y no-convencional que sea. No sé de impedimento *a priori* alguno para tener algo de confianza en que se puede dar tal semántica para las representaciones mentales, aunque no estén gobernadas por convenciones.)

Con lo que no me molesta comprometerme es con la idea de que las representaciones *van a tener una semántica u otra*—como yo lo veo, una función esencial de las representaciones es, bueno, *representar*: describir o “dirigirse” al mundo. (Si tuviera más seguridad en la creencia de que existen verdades analíticas, diría con gusto que tomo a esto como una verdad analítica.) Y dado eso, no veo en absoluto ningún impedimento de principio para que haya una teoría formal y sistemática que pueda describir la asignación, como valores, de aquello que cada representación busca representar. En breve, me parece que la misma idea de representación involucra la idea de que tal cosa se usa *para representar a una (otra) cosa*, y no sé de una buena razón por la cual debería haber impedimentos para capturar este hecho en una teoría semántica formal y sistemática.<sup>132</sup>

Arriba mencioné que el acercamiento estándar define a las condiciones de verdad de una oración como una clase de mundos posibles: aquellos donde la oración es, intuitivamente, verdad. Por supuesto, esta no es la única manera en que se ha propuesto entender al contenido semántico de una oración declarativa.<sup>133</sup> Sin embargo, por propósitos expositivos, voy a desarrollar lo que sigue bajo las asunciones estándares.

---

de Fine. Para verlo, simplemente notemos que en las condiciones de referencia arriba enunciadas podrían incluirse condiciones relacionales, al estilo de Fine (2007: §§1.G, 2.F). Esto es bueno, pues justamente el punto de movernos hacia la teoría general de la representación es evitar compromisos teóricos que traigan problemas para la fundamentación epistemológica del conocimiento modal y de esencias—mientras que dejamos que tales problemas sean decididos en las áreas particulares de la investigación.

Véase también Valdivia 2014a para una discusión crítica de la propuesta de Fine.

132 Véase, por ejemplo, Greenberg 2011: §1.1 para la respuesta a varias críticas contra la idea de que la representación pictórica no puede tener una semántica; véase Blumson 2010 para una semántica bidimensional para la representación pictórica.

133 Por ejemplo, Yablo (*manuscrito*) asume que un contenido semántico de una oración declarativa es su *tema*, o “*subject matter*”, que se define con el acercamiento lewisiano a la lógica relevante, y apelando a ideología hiperintensional: al operador *porque*.

## 64. representaciones fiables

[W]e can say that a dynamical theory is approximately true just if the modeling geometric structure approximates (in suitable respects) to the structure to be modeled: *a basic case is where trajectories in the model closely track trajectories encoding physically real behaviors* (or, at least, track them for long enough).

—Peter Smith, *Explaining Chaos*

## 64.1. de la rigidez...

Una vez que hemos esbozado la estructura posible de la teoría general de la representación, y hemos notado que *no necesita* tener los altos compromisos teóricos con filosofías particulares del lenguaje o de la mente, podemos ver que una epistemología modal se puede fundamentar aquí. (También lo hice, en parte, en el capítulo anterior, defendiendo EXTERNISMO MÍNIMO como una tesis externista que no necesita comprometerse con la negación de varios tipos de internismos.) Han quedado detalles por afinar, pero hemos evitado compromisos problemáticos. Los detalles, y adiciones posteriores, pueden implicar compromisos sustanciales. Pero el centro de la epistemología modal quedará fundamentado—o esa, al menos, es mi intención—sobre compromisos relativamente reducidos: un tipo de externismo amigable con algunas formas de internismo, un fiabilismo que aquí terminaré de afinar, la hipótesis de que toda representación, en cuanto que lo es, debe poder evaluarse en términos de sus condiciones de adecuación, y el postulado de que tenemos una heurística o proceso mental para *simular* (puede resultar que tengamos más de uno).

En lo que sigue, hasta el final de la tesis, voy a revisar las condiciones centrales de la epistemología modal que defenderé. Aquí asumo el externismo y fiabilismo defendidos en el capítulo anterior. También doy por sentado que contamos con una semántica para las representaciones en cuestión—que simplemente necesita entregarnos, por los medios que resulten ser los adecuados, las condiciones de adecuación de tales representaciones. Con estas tres asunciones basta para fundamentar una epistemología modal plausible e informativa.

Para motivar mi postura, comencemos con una de las ideas más importantes en la filosofía del lenguaje de la segunda mitad del siglo XX: la *rigidez*. De acuerdo con Kripke (1980), los nombres propios (entre otros términos) del lenguaje natural son *rígidos*: si un nombre  $n$  refiere a un objeto  $x$ , entonces  $n$  refiere a  $x$  en todo mundo posible donde  $x$  exista, y no refiere a ninguna cosa distinta de  $x$  en los mundos donde  $x$  no exista.<sup>134</sup> De manera equivalente,  $n$  es rígido sii, dado

134 Esta caracterización, así explícitamente puesta, no aparece en *El Nombrar y la Necesidad*; pero Kaplan (1989: 569-70) cita una carta de Kripke donde aparece. Kaplan nos dice: “In a letter (asking that I take his remarks into account in these afterthoughts), Kripke states that the notion of rigid designation he intended is that ‘a designator  $d$  of an object  $x$  is rigid, if it designates  $x$  with respect to all possible worlds where  $x$  exists, and *never designates an object other than  $x$  with respect to any possible world.*’ This definition is designed to be neutral with regard to the question whether a designator can designate an object at a world in which the object doesn't exist. It was motivated, he says, by the desire to avoid getting bogged down in irrelevant discussions of the existence



que refiere a  $x$ , entonces para cualquier  $w$  y cualquier  $y$ :  $n$  refiere a  $y$  en  $w$  sii  $y = x$ .

Ahora bien, esto todavía es neutral acerca de lo que sucede en mundos donde  $x$  no exista. Como hace notar Stanley (1997: 556), tenemos tres opciones para estos mundos: que  $n$  no refiera a nada en ellos, que  $n$  refiera a  $x$  incluso en esos mundos donde  $x$  no existe, o que  $n$  refiera a  $x$  en algunos pero no todos de ellos. La última opción no está motivada, así que la ignoramos. Los nombres que satisfacen la primera condición ( $n$  no refiere a nada en mundos donde  $x$  no existe) se conocen como *nombres persistentemente rígidos*. Los nombres que satisfacen la segunda opción ( $n$  refiere a  $x$  incluso en mundos donde  $x$  no existe) se conocen como *nombres obstinadamente rígidos*. Finalmente, los nombres que satisfacen ambas opciones: los nombres de objetos necesarios; a estos se les conoce como *nombres fuertemente rígidos*.<sup>135</sup>

Como el mismo Stanley hace notar (p. 557):

[... N]o es claro hasta qué grado las cuestiones sobre la rigidez persistente versus la rigidez obstinada son sustantivas, en lugar de discusiones meramente terminológicas disfrazadas, acerca de cómo es mejor usar la expresión 'evaluación respecto a un mundo'.\*

*Aquí voy a enfocarme en los nombres persistentemente rígidos*. Como veremos abajo, estos nombres permiten fundamentar una generalización de la rigidez bastante elegante y teóricamente servicial. Algo parecido podría hacerse con los nombres obstinadamente rígidos, aunque la presentación sería algo menos elegante y sin una clara ganancia en los recursos teóricos.

Entonces me restringiré a los nombres persistentemente rígidos (que en adelante llamaré simplemente "rígidos"). Entonces, sea  $n$  un nombre rígido—tenemos que, dado que  $n$  refiere a  $x$ , entonces para cualquier mundo  $w$ :  $x$  existe en  $w$  si y sólo si  $n$  refiere a  $x$  en  $w$ . Para ver la equivalencia, notemos que si en algún mundo  $w$ ,  $x$  existe en  $w$  pero  $n$  no se refiere a él en  $w$ , entonces  $n$  no es rígido; y si en  $w$ ,  $n$  refiere (a lo que sea) pero  $x$  no existe en  $w$ , entonces  $n$  tampoco es rígido (*i.e.*, persistentemente rígido).

Notemos que una de las razones más importantes para considerar a la rigidez fue, en el contexto de Kripke (1971), poder hacer sentido de la prueba de la necesidad de la identidad *para los nombres del lenguaje natural*.<sup>136</sup> Veamos cómo.

Si los nombres cambiaran de referente de un mundo posible a otro, habría fórmulas con la forma  $\lceil (a = b) \rightarrow \Box(a = b) \rceil$ , donde  $\lceil a \rceil$  y  $\lceil b \rceil$  son nombres propios, que serían falsas. Pues aunque fuera verdad que  $\lceil a = b \rceil$ , con lo que  $\lceil a \rceil$  se referiría a lo mismo que  $\lceil b \rceil$ , todavía podría ser falso que  $\lceil \Box(a = b) \rceil$ , pues podría haber un mundo en que cada nombre refiere a una cosa distinta del referente del otro. Es sólo cuando los términos flanqueando el signo de identidad son rígidos, que el enunciado de identidad entre ellos puede ser, si verdadero, necesario.

Pero, por supuesto, esto es a nivel del lenguaje. *A nivel metafísico*, de hecho *las cosas* son, si

---

question." Pero para aclaraciones sobre la evidencia textual acerca de lo que Kripke parecería estar diciendo en *Identidad y Necesidad* y *El Nombrar y la Necesidad*, véase la nota 8 del mismo texto de Kaplan.

135 Agradezco a Melahuac Hernández la necesidad de aclarar estas cuestiones.

\* "... it is unclear to what degree issues about persistent rigidity versus obstinate rigidity are substantive, rather than merely disguised terminological discussions about how best to use the expression "evaluation with respect to a world"."

136 Pues la prueba de Kripke sólo involucró *variables ligadas*, por lo que, sin rigidez, todavía podríamos tener la tesis de la necesidad de la identidad y enunciados  $\lceil \Box(a = b) \rceil$  contingentes (véase abajo). Agradezco a Alessandro Torza el haberme hecho notar la necesidad de aclarar esto.

idénticas, necesariamente idénticas. El punto de la rigidez es que nos permite representar este hecho mediante el lenguaje natural. ¿Cómo? Pues porque, *para efectos representacionales, un nombre rígido juega el mismo papel que su referente*. Un nombre rígido es un *sustituto perfecto* de su referente, a nivel representacional. Esto porque un nombre rígido es, como vimos arriba, tal que su referirse a  $x$  es necesariamente equivalente con la existencia de  $x$ . La falla de esta equivalencia podría, de la manera explicada arriba, dar lugar a la falla del enunciado de la necesidad de la identidad.<sup>137</sup>

En un sentido claro, entonces, los nombres rígidos son *representacionalmente fiables*: no sólo nunca cambian su referente, sino que el que refieran a una cosa u otra es *evidencia decisiva* para la existencia de su referente intencional.

Ahora voy a generalizar esta idea. Vamos por pasos.

Recordemos que la relación *RepAs* tomaba a una representación, a su representado intencional, al aspecto de este bajo el cual la representación lo representa, y a las convenciones representacionales. Ahora bien, dado que *RepAs*( $r, o, A(o), C$ ) las condiciones en las que  $r$  representa a  $o$  como  $A(o)$ —como cuando un ítem lingüístico obtiene su valor semántico—fueron nombradas arriba sus *condiciones de adecuación* (así, por ejemplo, las condiciones de adecuación de una oración son sus condiciones de verdad).

La *referencia* es, plausiblemente, la condición de adecuación de los nombres propios: un mundo  $w$  satisface las condiciones de adecuación de un nombre  $n$  sii  $n$  refiere a su referente intencional en  $w$ . Pero  $n$  no necesita ocurrir en  $w$ :  $n$  puede usarse en otro mundo  $v$  para referirse a su referente intencional en  $w$ , referente que se fija según el uso en  $v$ . Por ejemplo, en el mundo real @ podemos usar mi nombre, “Carlos”, para referirnos a mí en otro mundo posible no-real,  $w$  (como cuando decimos cosas como “Carlos podría haber estudiado derecho”, que de manera estándar se interpretan como siendo verdad sii en algún otro mundo posible (accesible) existo yo y estudio derecho ahí.) Si en ese mundo existo yo y las convenciones representacionales del uso de nombres se dan, entonces un uso de “Carlos” en @ puede referirse a mí en  $w$ .

La idea entonces es la siguiente. Definimos una función del conjunto de mundos  $W$  a su potencia:

$$A(n): W \rightarrow \text{Pot}(W)$$

tal que  $A(n)(w) = V \subseteq W$  sii el uso del nombre  $n$  en  $w$  refiere a su referente intencional (como se usa  $n$  en  $w$ ) en exactamente los mundos en  $V$  ( $w$  mismo puede ser un elemento de  $V$ ).

En general, para cualquier representación  $r$ , tal que *RepAs*( $r, o, A(o), C_w$ ) en  $w$  (para  $o, A$ , dados, donde  $C_w$  son las convenciones que fijan el uso de  $r$  en  $w$ ),

$$A(r): W \rightarrow \text{Pot}(W)$$

$$A(r)(w) = \{v \in W: v \models \text{RepAs}(r, o, A(o), C_w)\}$$

Es decir,  $A(r)(w)$  le asigna a  $r$  sus *condiciones de adecuación relativas a  $w$* —el conjunto de mundos donde representan a su representado *como  $r$  es usada en  $w$* . En particular,  $A(r)(@)$  (que abreviaré  $A_{@}(r)$ ) son las condiciones de adecuación de  $r$  como se la usa en @.

Con lo que la rigidez de un nombre  $n$  se formaliza, bajo estas ideas, como:

---

137 En el artículo de la enciclopedia Stanford (LaPorte 2006: §3) se muestran otros tres usos filosóficos de la rigidez: el necesario *a posteriori*, el contingente *a priori* y el esencialismo.

$A_{@}(n) = \{w \in W:$   
 $n$  se refiere únicamente a  $x$  en @, y  
 $x$  existe en  $w$  sii  $n$  se refiere únicamente a  $x$  en  $w$  (de acuerdo a las condi-  
 ciones del mundo real, @}

Es decir,  $n$  es rígido sii sus condiciones de adecuación en el mundo real,  $A_{@}(n)$ , contienen exactamente a los mundos  $w$  donde se da la equivalencia:  $n$  refiere a su referente del mundo real,  $x$ , en  $w$  sii  $x$  existe en  $w$ . Equivalentemente,  $n$  no es rígido sii:

para algún  $w \in A_{@}(n)$ ,  $n$  refiere a  $x$  en  $w$  pero no en @, o  
 para algún  $w \in A_{@}(n)$ ,  $n$  refiere a  $x$  en @ pero no en  $w$ .

Es decir,  $n$  no es rígido sii sus condiciones de adecuación contienen a mundos donde se rompe la equivalencia entre el referir de  $n$  a  $x$  y la existencia de  $x$ . La rigidez de  $n$  en un mundo entonces consiste en que *sea adecuado siempre y cuando refiera a su referente de ese mundo*. Con lo que la rigidez de un nombre  $n$  en @ (dado que refiere a  $x$ ) se ve como: un mundo  $w$  satisface las condiciones de adecuación@ de  $n$  sii  $n$  refiere a  $x$  en  $w$ ; es necesariamente equivalente que  $n$  refiera y que  $x$  exista, es decir: es necesariamente equivalente que  $n$  satisfaga sus condiciones de adecuación y que  $x$  exista.

Hemos visto varias formulaciones equivalentes de la rigidez de un nombre, pero me quiero quedar con la última: *es necesariamente equivalente que  $n$  satisfaga sus condiciones de adecuación y que  $x$  exista*, o:

(RIGIDEZ)

Un nombre  $n$  es rígido de  $x$  sii: para todo mundo  $w$ :  $w \in A_{@}(n)$  sii  $x$  existe en  $w$ .

#### 64.2. ...pasando por un *excursus* sobre el bidimensionalismo...

Algo importante es que, como se habrá notado, relativicé la rigidez a un mundo. Esto se hace desde los albores de la tradición kripkeana. Pues un mismo nombre se puede usar de diferentes maneras—dependiendo de cómo se individúen los nombres y sus usos. Por ejemplo, usos de “Carlos” pueden referirse en otro mundo a otra persona que haya sido bautizada con ese nombre. Fijo los mundos “base” para evitar estos casos. Como dice Kripke, evaluamos a las palabras *de la manera en que las usamos nosotros* (2005: 103). No nos importa que en algún mundo posible se utilicen sonidos fonéticamente indistinguibles a las emisiones de “nopal” para referirse a los dodecaedros—nos importa cómo se usa “nopal” en este mundo.

He, por supuesto, “empacado” demasiadas cosas en la relativización a mundos. Básicamente, *estoy pasando de largo sobre un importante punto que motiva a los acercamientos bidimensionalistas*. Veamos.

Como es bien sabido, la semántica tradicional influida por Carnap—que podríamos llamar *unidimensionalista*—representa al significado de un término con una *intensión*: una función sobre mundos posibles tal que para cada mundo posible regresa la extensión del término en ese mundo. La semántica bidimensionalista no toma a las intensiones como definidas sobre el conjunto de mundos posibles, sino sobre el producto cruz de los mundos posibles con otro conjunto, quizá distinto; de manera equivalente, define dos intensiones—una sobre el conjunto de mun-

dos, otra sobre otro conjunto, quizá el mismo—y define a las extensiones como resultado de esas dos intensiones. Esto se hace para modelar cómo la referencia de nuestros términos depende de varias cosas. Por ejemplo, el acercamiento de Stalnaker en “Assertion” (Stalnaker 1978) busca modelar cómo la referencia de nuestros términos depende de las situaciones objetivas (Stalnaker 2006); mientras que, para otro ejemplo, el acercamiento de Chalmers busca (entre varias otras cosas) modelar cómo la referencia de nuestros términos depende de nuestro conocimiento *a priori* (Chalmers 2004) y con ello, se ve comprometido con un acercamiento fuertemente internista, tanto acerca de la referencia como acerca de la epistemología (cf. Schroeter 2005, Stalnaker 2006).

Dicho esto, podemos notar que nada impide, en principio, tener un acercamiento *multi-dimensional* a la referencia. No hay nada en el aparato de funciones definidas sobre mundos posibles que impida tener una, dos o más de dos tales funciones (quizá incluso infinitas). Si queremos, podríamos tener una función para cada factor que pueda contribuir a cómo usamos el lenguaje para referirnos al mundo—una función que tenga como insumos factores epistémicos, o contextuales, como los bidimensionalismos más famosos; pero también una que tome en cuenta el día del año en que se use un término o el humor del usuario. Si no individualizamos esas funciones, es porque nos parecen *filosóficamente poco importantes*: es poco iluminador de la referencia el preguntarnos cómo es que el humor del usuario de un término influye en la referencia de ese término.

Así, el acercamiento bi- o, en general, *multi-dimensionalista*, supone una *selección* de factores que se toman como influyentes en la referencia de los términos.<sup>138</sup> Una epistemología basada en un acercamiento *multi-dimensionalista*, entonces, será al menos tan controversial como controversial sea la selección de factores de tal *multi-dimensionalismo* subyacente. Sobre todo, parece que invertiría el orden explicativo apropiado—*primero* deberíamos tener una epistemología sobre los factores que conectan a nuestras representaciones con el mundo (como las creencias y su verdad, los nombres y su referencia, etc.), y *entonces* podríamos modelar la estructura de nuestra semántica: la estructura de las condiciones en las que tales representaciones refieren, son verdaderas, o, en general, satisfacen sus condiciones de adecuación—influidas por la epistemología subyacente. En un *slogan*, la tesis es: *primero la epistemología, luego la semántica*.

Aquí busco enfocarme en las cuestiones epistémicas. Claro, simplemente estoy asumiendo que tenemos una idea suficientemente clara, como creo que de hecho la tenemos, de qué significa *fijar* la interpretación de una representación, decir cómo es que esta ha de referir. Es una vez que asumo esto, que puedo pasar a indagar en las cuestiones epistemológicas de cómo esas representaciones con esas interpretaciones ya fijadas, pueden brindarnos conocimiento o representar al mundo. Pero, ya teniendo claridad sobre esto, puedo entonces ayudarme de ello para regresar a preguntarme cómo esas cuestiones epistemológicas influyen sobre la fijación de interpretaciones.

Regresemos al punto.

---

138 Por ejemplo, Stalnaker (2006) muestra que la interpretación de Chalmers del aparato bidimensional depende de una epistemología y una filosofía de la intencionalidad que son internistas, y con compromisos altamente no-triviales (como el holismo radical, o la infabilidad del valor cognitivo, o la existencia de un lenguaje “epistémicamente puro” al cual todo nuestro lenguaje se puede reducir.)

## 64.3. ...a la representación fiable

Podemos definir nociones completamente análogas a la rigidez para varios tipos de representación además de los nombres propios. Por ejemplo, Llamas (*manuscrito*) define las siguientes nociones para *conceptos*:

Un concepto  $C$  es *transparente de* un objeto  $o \leftrightarrow_{df} \Box(C \text{ se satisface} \leftrightarrow o \text{ existe})$

$C$  es *opaco de*  $o \leftrightarrow_{df} \neg\Box(o \text{ existe} \rightarrow C \text{ se satisface}) \ \& \ \neg\Box(C \text{ se satisface} \rightarrow o \text{ existe})$

$C$  es *translúcido de*  $o \leftrightarrow_{df} \Box(o \text{ existe} \rightarrow C \text{ se satisface}) \ \& \ \neg\Box(C \text{ se satisface} \rightarrow o \text{ existe})$

$C$  es *ligeramente opaco de*  $o \leftrightarrow_{df} \Box(C \text{ se satisface} \rightarrow o \text{ existe}) \ \& \ \neg\Box(o \text{ existe} \rightarrow C \text{ se satisface})$

Por supuesto, la *transparencia* de un concepto es básicamente lo mismo que su *rigidez*, si extendemos la terminología desde nombres a conceptos.<sup>139</sup>

Recordemos RIGIDEZ: Un nombre  $n$  es rígido de  $x$  sii: para todo mundo  $w: w \in \mathcal{A}_{@}(n)$  sii  $x$  existe  $w$ . Ahora bien, si entendemos la *satisfacción* de un concepto como la manera fundamental en la que un concepto representa, entonces las *condiciones de adecuación* de un concepto es el conjunto de circunstancias en las que el concepto es satisfecho. Si el concepto es *transparente* de un objeto, entonces las condiciones de adecuación del concepto involucrarán necesariamente a ese objeto y a ningún otro. Así, análogamente a RIGIDEZ, podríamos definir:

(TRANSPARENCIA)

Un concepto  $C$  es *transparente de*  $o$  sii para todo mundo  $w: w \in \mathcal{A}_{@}(C)$  sii  $o$  existe en  $w$ .

Definiciones parecidas se pueden dar para las demás nociones. Por ejemplo:

(OPACIDAD)

Un concepto  $C$  es *opaco de*  $o$  sii

para algún  $w: w \in \mathcal{A}_{@}(C)$  y  $o$  no existe en  $w$ , y

para algún  $v: v \notin \mathcal{A}_{@}(C)$  y  $o$  existe en  $v$ .

Lo que ahora quiero hacer es generalizar todavía más, en dos frentes. *Primero*, llamaré a una representación *totalmente fiable* cuando tiene propiedades que o son la rigidez o la transparencia, o son análogas a ellas:

(FIABILIDAD TOTAL)

Una representación  $r$  es *totalmente fiable de*  $o$  sii:

para todo mundo  $w: w \in \mathcal{A}_{@}(r)$  sii  $o$  existe en  $w$ .

Una representación totalmente fiable, entonces, es aquella cuyas condiciones de adecuación están necesariamente ligadas a la existencia de su representado. Así, *segundo*, una representación totalmente fiable es aquella para la que, intuitivamente:

(FIABILIDAD TOTAL—PROBABILÍSTICA)

Una representación  $r$  es *totalmente fiable de*  $o$  sii:

$\text{Prob}(o \text{ existe} \mid \text{se satisfacen } \mathcal{A}_{@}(r)) = 1$  y

$\text{Prob}(\text{se satisfacen } \mathcal{A}_{@}(r) \mid o \text{ existe}) = 1$ .

Como dije arriba, un nombre rígido es un *sustituto perfecto* de su referente, a nivel represen-

<sup>139</sup> Se asume que  $o$  es un particular, pues para propiedades la condición necesita afinarse.

Por otro lado, Llamas asume que los operadores modales en cuestión corren sobre mundos epistémicamente posibles, a su vez entendidos de manera muy cercana, al parecer, a como los entiende Chalmers. Yo no me comprometo con ese aparato conceptual; sólo me interesa la forma de las definiciones, pues muestran de manera clara cómo se puede generalizar de nombres a conceptos.

tacional, y también es *representacionalmente fiable*: no sólo nunca cambia su referente, sino que el que refieran es *evidencia decisiva* para la existencia de su referente. Estas afirmaciones son el fundamento de FIABILIDAD TOTAL—PROBABILÍSTICA. La idea de que el que una representación totalmente fiable satisfaga sus condiciones de adecuación es evidencia decisiva para que exista su referente, se formaliza diciendo simplemente que la probabilidad de que el objeto exista, *dado* que se dan tales condiciones de adecuación, es la más alta posible: 1. Un razonamiento parecido nos lleva a la segunda componente de FIABILIDAD TOTAL—PROBABILÍSTICA. Todavía más, podríamos tener una condición relativamente intuitiva de correlación, así:

(FIABILIDAD TOTAL—CORRELACIÓN)

Una representación  $r$  es *totalmente fiable* de  $o$  sii:

$$\text{Corr}(o \text{ existe, se satisfacen } A_{@}(r)) = 1.$$

(Donde **Corr** es la correlación de Pearson, y “ $o$  existe” y “se satisfacen  $A_{@}(r)$ ” son en realidad las variables aleatorias indicadoras correspondientes).<sup>140</sup> Así, la idea sobre el papel evidencial de la satisfacción de las condiciones de adecuación de  $r$ , respecto a la existencia de  $o$ , cuando  $r$  es totalmente fiable, también puede entenderse como que las condiciones de adecuación y la existencia de su representado  $o$  tienen un grado máximo de correlación.

Todavía no he dicho mucho sobre la función de probabilidad, **Prob**, que estoy usando aquí; aunque asumo que **Corr** se ha de definir en términos de **Prob** (de la manera estándar en teoría de la probabilidad). Antes de ello, notemos las dos generalizaciones que he hecho desde la noción kripkeana de rigidez: primero, generalizo una condición análoga para representaciones diferentes a los nombres; segundo, tales condiciones pueden entenderse ya sin el rígido marco de las implicaciones y bicondicionales, ahora en el marco más flexible de la probabilidad.

Ahora diré algo más sobre **Prob**.

Comenzamos notando que una proposición necesaria debe tener una *probabilidad* de 1. Pero no toda función de probabilidad hará que una *proposición* (o “evento”—o, más exactamente, *conjunto de mundos posibles*) con probabilidad de 1 sea *metafísicamente necesaria*. Lo que ahora queremos es que una función de probabilidad que nos diga que la probabilidad de  $\langle p \rangle$  es 1 sii  $\langle p \rangle$  es necesariamente verdadera. Queremos, pues, una probabilidad que mida el “grado” de necesidad metafísica de las proposiciones: 1 para la necesidad, 0 para la imposibilidad, y números intermedios para proposiciones contingentes.

La idea se formaliza así. Sea  $W$  el conjunto de todos los mundos metafísicamente posibles. Definimos un  $\sigma$ -álgebra sobre  $W$  de la manera obvia: los puntos de  $W$  son las muestras, de modo que  $W$  es nuestro espacio muestral, y nuestro conjunto de eventos,  $E$ , es simplemente el conjunto potencia de  $W$ . Con lo que los eventos serán simplemente lo que muchos llaman *proposiciones*.<sup>141</sup> Ahora definimos una función de probabilidad, **Prob**:  $E \rightarrow (0, 1)$  que cumple los axiomas

140 La correlación de Pearson se define, para dos variables aleatorias  $A, B$ , así:

$$\text{Corr}(A, B) = \text{Covarianza}(A, B) / (\sigma_A)(\sigma_B) = E[(A - E[A])(B - E[B])] / (\sigma_A)(\sigma_B),$$

donde  $\sigma$  es la desviación estándar de cada variable y  $E$  el operador de valor esperado.

Otras medidas de correlación, por supuesto, podrían motivarse aquí.

141 Bajo cierta tradición en la filosofía del lenguaje (e.g., Stalnaker, 1978; Lewis, 1986), los conjuntos de mundos posibles son llamados *proposiciones*. No todos están de acuerdo (e.g., Salmon 1986), pero no necesito compromisos aquí. Basta notar que para ambos lados, un conjunto de mundos posibles es al menos la *extensión* de una proposición: el conjunto de mundos donde la proposición es verdadera. Así, nuestro conjunto de eventos (el con-

de Kolmogorov (donde  $e$  es cualquier evento en  $E$ ):

$$\text{K1: } \text{Prob}(e) \geq 0,$$

$$\text{K2: } \text{Prob}(W) = 1$$

$$\text{K3: } \text{Prob}(e_1 \cup e_2 \cup \dots) = \sum_{i=1}^{\infty} \text{Prob}(e_i), \quad \text{cuando los } e_i \text{ son disjuntos a pares.}$$

Ahora requerimos que la probabilidad sea regular, es decir, que:

$$\text{REGULARIDAD: } \text{Prob}(e) = 0 \text{ sii } e = \emptyset.$$

Sabemos entonces que una proposición necesaria tiene probabilidad 1 y una imposible probabilidad 0; pero ¿qué pasa con las contingentes? Su probabilidad debe estar estrictamente entre 1 y 0, pero hasta ahora no tenemos una manera de definirla. Por supuesto, sabemos (pues se sigue fácilmente de los axiomas) hechos estructurales como: si una proposición  $e$  tiene probabilidad  $r$ , entonces  $\neg e$  tendrá probabilidad de  $1-r$ ; pero todavía no tenemos una manera de asignarle probabilidades incondicionalmente a los eventos contingentes.

El acercamiento más básico, que usa una distribución *uniforme*, podría funcionar aquí directamente, si el conjunto de mundos posibles fuera finito:  $\text{Prob}(e)$  sería simplemente la cardinalidad de  $e$  (el número de mundos donde  $e$  es verdadera) dividido sobre la cardinalidad de  $W$ . Sin embargo, es implausible afirmar que existen sólo un número finito de mundos posibles (de hecho, no es difícil argumentar que deberían existir al menos tantos como números reales).<sup>142</sup>

Ahora bien, quiero notar que este problema—el de dar un acercamiento matemático que sea natural (no *ad hoc*) y sistemático para espacios eventuales infinitos—es altamente general, para nada particular a mi postura o esencialmente relacionado con la epistemología de la modalidad. Así, en principio, podría adoptar cualquier solución a este problema—cuando estuviera disponible.

Una primera idea es utilizar algo de análisis matemático no estándar y decir que  $\text{Prob}(p)$ , para  $\langle p \rangle$  contingente, va a ser un número infinitesimal (cf. Lewis 1980), donde un número infinitesimal es simplemente un número  $x$  tal que para cualquier número real  $r$ ,  $0 < x < r$ .

Sin embargo, hay un argumento bien conocido de Williamson (2007b) que muestra que, incluso usando usamos infinitesimales, algunos eventos intuitivamente posibles deben tener probabilidad 0, violando así REGULARIDAD.<sup>143</sup>

junto potencia de  $W$ ) es o el conjunto de las proposiciones mismas, o al menos el de sus *extensiones*.

142 Supongamos que tenemos una partícula puntual  $p$  en el espaciotiempo y que este es continuo. Entonces supongamos que  $p$  se mueve una micra a la izquierda (relativo a cierto marco de referencia). Entonces, al moverse una micra,  $p$  se movió también a través de una cantidad no-denumerable de puntos espaciotemporales. Para un conjunto no-denumerable de puntos  $P$  a través de los cuales  $p$  pasó, parece posible que  $p$  se detuviera justamente en  $P$  y no se moviera más. Cada una de estas posibilidades es distinta, y con la asunción de que para cada posibilidad hay un mundo posible obtenemos el resultado de que existen al menos una cantidad no-denumerable de mundos posibles.

143 El argumento se puede reconstruir así (cf. Weisberg *por aparecer*):

Sea (H1...) una secuencia infinita de volados, empezando en  $t_1$ , en la que siempre cae águila (que llamaré "Heads"; al sol lo llamaré "Tails"). Sea (H1) el volado, en  $t_1$ , en el que cae Heads. Sea (H2...) una secuencia infinita de volados, empezando en  $t_2$  ( $> t_1$ ) en la que siempre cae Heads. Se asume que los volados son eventos independientes entre sí y que la moneda no está sesgada (es decir,  $\text{Prob}(\text{Heads}) = 1/2 = \text{Prob}(\text{Tails})$ ). Entonces:

$$\begin{aligned} \text{Prob}(\text{H1} \dots) &= \text{Prob}(\text{H1} \ \& \ \text{H2} \dots), \text{ por teoría de probabilidad} \\ &= \text{Prob}(\text{H1}) \text{Prob}(\text{H2} \dots | \text{H1}), \text{ por teoría de probabilidad} \\ &= (1/2) \text{Prob}(\text{H2} \dots | \text{H1}), \text{ porque la moneda no está sesgada} \\ &= (1/2) \text{Prob}(\text{H2} \dots), \text{ por la independencia entre } (\text{H2} \dots) \text{ y } (\text{H1}) \end{aligned}$$

Varias respuestas se han ofrecido y se pueden ofrecer para este problema. Consideraré tres.

La primera es simplemente cambiar de un marco teórico con una medida desde el conjunto de los mundos posibles al intervalo unitario, a un marco teórico con probabilidad comparativa. Se puede mostrar que, en un marco de probabilidades comparativas, un análogo de REGULARIDAD es satisficible (Haverkamp & Schulz 2012).

Otra respuesta es preservar el uso de infinitesimales, pero cambiar algunos fundamentos de la concepción de la probabilidad. Hofweber (2014) argumenta que debemos rechazar el acercamiento kolmogoroviano y aceptar, en vez de ello, una variación de este usando números infinitesimales, complementados con un par de tesis sobre los fundamentos de la probabilidad.

Hofweber rechaza el acercamiento kolmogoroviano porque este permite falsear la siguiente tesis:

(CONSTREÑIMIENTO MÍNIMO) Si la probabilidad objetiva de  $p$  es 0, entonces no  $p$ . Si la probabilidad objetiva de  $p$  es 1, entonces  $p$ .

De acuerdo a la argumentación de Hofweber, CONSTREÑIMIENTO MÍNIMO es una verdad conceptual acerca de la probabilidad objetiva, así que debemos buscar un acercamiento que lo satisfaga.

Hofweber argumenta que la concepción correcta sobre la probabilidad objetiva debe incorporar tres cosas:

- (1) El marco formal para modelar a la probabilidad objetiva es un acercamiento parecido al de Kolmogorov, pero sustituye al intervalo unitario real por un sistema con infinitesimales.
- (2) NO LOCALIDAD: las probabilidades objetivas no son, en lo general, características locales de los eventos. Las probabilidades objetivas de los eventos complejos no son, en lo general, determinadas localmente por las probabilidades objetivas de los eventos que los constituyen. La medición de la probabilidad objetiva debe hacerse sobre todo el espacio eventual, no localmente (sobre eventos individuales).
- (3) FLEXIBILIDAD: No hay un sistema numérico único que deba funcionar como las medidas de la probabilidad objetiva para todos los espacios eventuales. Qué sistema numérico deba usar para medir la probabilidad objetiva depende del dominio a ser medido.

Con estos tres compromisos, argumenta Hofweber, podemos tener una teoría sobre la metafísica y representación de la probabilidad que satisfaga CONSTREÑIMIENTO MÍNIMO.

Ahora bien, CONSTREÑIMIENTO MÍNIMO es un principio más débil que REGULARIDAD. Sin embargo, dado que la solución hofweberiana bloquea el problema de Williamson, podría usarse para restaurar REGULARIDAD.

Una última salida posible sería “compactar” a los eventos, de tal manera que nos podamos quedar con un conjunto *finito* de ellos, y así poder tener probabilidad clásica bien definida. Para hacerlo, simplemente necesitaríamos una relación de equivalencia sobre el conjunto de los mundos posibles, mediante la cual pudiéramos “filtrar” a esta clase hacia una clase de equivalencia. Deberíamos asegurarnos que tal clase resultara finita, por supuesto. Además de ello, deberíamos

---

Por lo que (H1...) debe tener la mitad de la probabilidad de (H2...). Pero si estipulamos que (H1...) y (H2...) consisten en una serie de volados indiscernibles, y como tienen el mismo cardinal, parece que debería pasar que:

$$\text{Prob}(H1...) = \text{Prob}(H2...)$$

Ahora bien, incluso asumiendo infinitesimales, si  $x = 1/2x$ , entonces  $x = 0$ . Por lo que todo esto nos lleva a que  $\text{Prob}(H1...) = \text{Prob}(H2...) = 0$ .



argumentar que tal relación de equivalencia está filosóficamente bien motivada.

Cerremos esta sección. No he dicho que el acercamiento probabilístico que he motivado esté libre de problemas; o que las soluciones revisadas sean las únicas posibles, o que estén libres de problemas. El problema de tener una probabilidad natural y coherente sobre espacios eventuales infinitos *es* un problema. *Pero no es únicamente mío*. Se necesita resolverlo para tener una teoría adecuada sobre los fundamentos de la probabilidad—ya sea objetiva o subjetiva. Y no veo razón alguna para dudar de que, en principio, yo no pueda adoptar esa solución—cualquiera que resulte ser.

## 6.5. el conocimiento modal y de esencias mediante la doble fiabilidad

### 6.5.1. el fiabilismo de proceso revisitado y generalizado

Vamos a regresar al fiabilismo de proceso defendido en el capítulo anterior:

(FIABILISMO DE PROCESO-VERSIÓN SIMPLE)

Una creencia  $b$  está justificada sólo si  $b$  es producida por un proceso  $P$  que es fiable.

(FIABILIDAD VERITISTA SIMPLE)

Un proceso de producción de creencias  $P$  es fiable sii  $P$  tiende a producir suficiente valor veritista.

Ahora podemos ver cómo se generaliza:

(FIABILIDAD GENERAL SIMPLE)

Un proceso de producción de representaciones  $P$  es fiable sii  $P$  tiende a producir suficientes representaciones que satisfacen sus condiciones de adecuación.

(FIABILISMO DE PROCESO GENERAL—VERSIÓN SIMPLE)

Una representación  $r$  está justificada sólo si  $r$  es producida por un proceso  $P$  que es fiable.

La conjunción de ambas tesis nos da una primera estructura para nuestra teoría de la justificación de la creencia—y, en general, de la representación—modal, que llamaré simplemente FIABILISMO. Pero falta más.

Primero, como he venido diciendo desde el capítulo anterior, creo que las nociones epistémico-normativas—o al menos una muy buena parte—que han sido ampliamente discutidas en torno a la noción de creencia, son aplicables a la noción general de representación. Así como una creencia puede estar justificada cuando—asumiendo el veritismo (*cf.* §5.3, arriba)—es producida por un proceso que tiende a producir más creencias verdaderas que falsas, una representación puede estar justificada—asumiendo un análogo del veritismo—cuando es producida por un proceso (método, heurística, etc.) que tiende a producir más representaciones adecuadas que inadecuadas (*i.e.*, más representaciones que satisfacen sus condiciones de adecuación que representaciones que no lo hacen). Mencioné cómo ya se habían tomado pasos hacia tal idea (como en Jenkins 2008a o Moss 2013, ver también Bechtel & Abrahamsen 1991, Churchland 2001). De nuevo, me parece una generalización potencialmente fructífera. Nuestro “contacto” representa-

cional con el mundo no es únicamente mediante creencias o conceptos; una evaluación epistémica de tal “contacto” mediante otras maneras de representar puede ayudarnos a tener una mejor idea de cómo logramos, mediante los variados sistemas representacionales, armonizar nuestra mente y acciones con el ambiente.

Segundo, esto todavía deja la vaguedad del término “suficiente”. Pero como los fiabilistas han venido subrayando desde que Goldman propuso por primera vez la teoría, tal vaguedad es plausiblemente un componente de nuestros conceptos epistémicos, y variará de contexto a contexto. Incluso bajo una epistemología científica (Goldman 1992, *cf.* §§5.1, 5.5, arriba), que individúe los procesos (métodos, heurísticas, etc.) productores de representaciones fiables con base en la ciencia empírica, es dable suponer que la aplicación de los estándares normativos a esos procesos (métodos, etc.), será todavía una cuestión contextual—como parte de lo que en §5.1 llamé teoría de la atribución.

Tercero, todavía nos falta conectar FIABILISMO con la noción de representación fiable, que es a lo que ahora paso.

### 6.5.2. fiabilismo de representación

Vale recordar la noción de fiabilidad representacional total, en su formulación mediante correlaciones:

(FIABILIDAD TOTAL—CORRELACIÓN)

Una representación  $r$  es *totalmente fiable* de  $o$  sii:

$$\text{Corr}(o \text{ existe, se satisfacen } A_{@}(r)) = 1.$$

Primero, consideraciones intuitivas nos llevaron, en §6.4.3, a caracterizar la fiabilidad representacional como la correlación modal entre la existencia del representado y la satisfacción de las condiciones de adecuación de la representación. Esto nos permite considerar grados de correlación menores a 1. Así, una representación no necesita estar correlacionada al grado máximo para contar como fiable. Basta que lo esté a un grado suficiente. (Como ejemplo, una descripción de mí que no me describa en todo mundo posible donde yo existo, pero sí en la mayoría de ellos, todavía cuenta como una “buena” descripción mía—una *fiable*.) Qué cuente como *suficiente* para que una representación cuente como fiable, espejulo, es algo dependiente de contexto—o, al menos, a ser fijado convencionalmente en cada aplicación de la teoría.

Segundo, la correlación usada fue la correlación de Pearson (ver nota 140, arriba). La estructura de la teoría es compatible con usar otra correlación. Incluso, una revisión algo más profunda de la teoría podría llevarnos a dejar la correlación de lado, y buscar otra medida probabilística de la relación entre la existencia del objeto representado y la adecuación de la representación.<sup>144</sup>

Tercero, y más importante, vimos arriba que queríamos que la función de probabilidad que está a la base de todo esto, **Prob**, corriera sobre todos los mundos posibles, para así devolvernos la *probabilidad metafísica* de las proposiciones en cuestión. Vimos también que, dada la infinitud

144 Por razones que ya no puedo tratar aquí con mucho detalle, me parece que una alternativa interesante a probar sería la medida definida por la razón:

$$\text{Prob}(\text{se satisfacen } A_{@}(r) | o \text{ existe}) / \text{Prob}(\text{se satisfacen } A_{@}(r) | \neg o \text{ existe})$$

Compárese con Stegenga 2013.

del conjunto de mundos posibles, surgían problemas; pero argumenté que ese no era una problema específicamente *mío*, y que en principio una solución a este podría ser adoptada dentro de mi acercamiento.

Hemos desarrollado de manera independiente nuestro fiabilismo de proceso y nuestro fiabilismo de representaciones. Con el primero buscamos procesos (metodologías, etc.) que produzcan, de manera fiable, representaciones adecuadas: procesos cuyas representaciones resultantes suelen ser adecuadas. Con el segundo buscamos representaciones que fiablemente representen a su contenido intencional: representaciones que, en un número suficiente de mundos posibles donde su contenido existe, lo representen.

Es un error considerar a una representación aisladamente. Además de pertenecer a un sistema representacional—gobernado por convenciones representacionales—una representación es siempre resultado de un mecanismo productor. Esta es una lección fundamental de la tradición del fiabilismo procesual: debemos considerar los mecanismos productores de las creencias—y, en general, de las representaciones—para tener una visión completa sobre el perfil epistémico de las creencias—y, en general, de las representaciones.

En nuestro caso, es importante considerar las bases porque es con ellas que podemos entender mejor la *interpretación* de las representaciones (aquellas que son fijadas al menos en parte por convención). Diferentes sistemas de producción de representaciones pueden utilizar uno y el mismo objeto para representar a diferentes cosas, o para representar a una misma bajo diferentes aspectos. También, por supuesto, en cuanto haya representaciones mentales cuyo contenido no es fijado convencionalmente, ellas son mejor entendidas con relación a su “base” o mecanismo productor (en la ausencia de la consideración de bases o mecanismos, incluso tesis fiabilistas que no se proponen como fiabilismo *procesual* están expuestas a dificultades: Goldman 2009b, Sosa 2007, Williamson 2009, 2009b).

Tenemos, entonces, una doble estructura fiabilista: evaluamos la representación de las representaciones, en conjunto con la fiabilidad de su mecanismo productor. En la siguiente sección pondremos esta idea a trabajar.

Antes de ello, para la epistemología de esencias, será útil definir una noción de representación esencial: definimos el que una representación *tiende a lo esencial de x*. La idea básica es que una representación así es una que representa a  $x$  ya sea como siendo su esencia, o como siendo algo *suficientemente cercano* a su esencia. Cuando  $\lceil \square_x \rceil$  es el operador de verdad esencial de Fine (cf. §2.4, arriba), tenemos:

(REPRESENTACIÓN QUE TIENDE A LO ESENCIAL)

$r$  tiende a lo esencial de  $x$  sii:

(i)  $RepAs(r, x, A(x), C)$ , y ya sea que

(ii') o  $A$  es parte de la esencia de  $x$ , o

(ii'') o  $r$  *tiende a la esencia de x*: para toda  $p$  tal que  $\square_x(p)$ :

$Corr_r(p, \text{se satisfacen } A @ (r)) = n$ , para  $1 \geq n \gg 0$ .

Esta definición nos ayuda a permanecer en el espíritu fiabilista, que no pide representación *totalmente* fiable, sino sólo *suficientemente*. Esto nos deja todavía en el campo falibilista, y es plausible suponer que, en las aplicaciones de la teoría en contextos concretos, la indeterminación en (ii'') será resuelta. (Notemos también que (ii'') cubre el caso (ii'), pues en ese caso  $Corr_r = 1$ ).

## 6.5.3. conociendo esencias y necesidades: la teoría central

Hasta ahora tenemos:

(FIABILISMO DE PROCESO GENERAL—VERSIÓN SIMPLE)

Una representación  $r$  está justificada sólo si  $r$  es producida por un proceso  $P$  que es fiable.

(FIABILIDAD GENERAL)

Un proceso de producción de representaciones  $P$  es fiable sii  $P$  tiende a producir suficientes representaciones que son adecuadas.

(FIABILIDAD REPRESENTACIONAL)

Una representación  $r$  es *suficientemente fiable* de  $o$  sii:

$\text{Corr}_r(o \text{ existe, se satisfacen } \mathbf{A}_{@}(r)) = n, \quad \text{para } 1 \geq n \gg 0.$

Con esto ya podemos dar una teoría de la justificación para creencias modales y sobre esencias.

Decimos simplemente que:

(JUSTIFICACIÓN MODAL)

La creencia de  $S$  de que  $\Box p$  ( $\Diamond p$ ) está justificada si y sólo si:

- (i)  $S$  tiene una representación de  $\langle p \rangle$  que es suficientemente fiable,
- (ii)  $S$  tiene un proceso de producción de *simulaciones* que es fiable, y
- (iii) El proceso de simulación utiliza la representación fiable de  $\langle p \rangle$  y resulta en la creencia de  $S$  de que  $\Box p$  ( $\Diamond p$ ).

Donde el proceso de *simulación* en cuestión es cualquiera que resulte de la investigación empírica (cf. §4.1, arriba). Aquí, como antes, he estado asumiendo una versión “cruda” del fiabilismo, pero para bloquear contraejemplos puede que sea necesario añadirle otras cláusulas (e.g., una cláusula anti-derrotadores). Pero éso es cuestión del fiabilismo en general, y no particular a mi teoría.

Podemos ver que JUSTIFICACIÓN MODAL es una teoría adecuada. Para hacerlo más claro, regresaré a la metáfora de las etiquetas: por (i), la representación de  $S$  de que  $p$  se pegará a  $p$  en suficientes mundos posibles, por (ii), habrá un proceso psicológico de simulación (a ser descubierto empíricamente) que tiende a producir creencias o representaciones adecuadas, y por (iii) ese proceso utilizará la etiqueta de  $\langle p \rangle$  para “rastrearla” a través de tales simulaciones.

Ahora bien, esto nos da una teoría sobre la *justificación* de nuestras creencias *modales*. Todavía nos falta una sobre la *justificación* de nuestras creencias *esencialistas*, y revisar cómo aplicar estas ideas al *conocimiento*. Vamos por pasos.

Si  $\lceil \Diamond_x \rceil$  es el dual de  $\lceil \Box_x \rceil$ , tenemos:

(JUSTIFICACIÓN ESENCIALISTA)

La creencia de  $S$  de que  $\Box_x p$  (o de que  $\Diamond_x p$ ) está justificada si y sólo si:

- (i)  $S$  tiene una representación de  $\langle p \rangle$  que es suficientemente fiable,
- (ii)  $S$  tiene una representación que *tiende a lo esencial* de  $x$ ,
- (iii)  $S$  tiene un proceso de producción de *simulaciones* que es fiable,
- (iv) El proceso de simulación utiliza las representaciones fiable y que tiende a lo esencial de  $x$  y de  $\langle p \rangle$  (respectivamente) y resulta en la creencia de  $S$  de que  $\Box_x p$

(respectivamente, de que  $\diamond_x p$ ).

La conjunción de JUSTIFICACIÓN MODAL con JUSTIFICACIÓN ESENCIALISTA nos da mi propuesta de epistemología modal y de esencias, para el caso de la justificación. He intentado seguir, hasta donde sea posible, el espíritu fiabilista: externista, falibilista y naturalista, y sin necesidad de compromisos demasiado sustantivos.

Para el caso del conocimiento, hay dos opciones (cf. el inicio del capítulo 5, p. 107). O el conocimiento se analiza—contra toda expectativa—y en ese caso plausiblemente será analizado en términos de la justificación, o es inanalizable. Si el conocimiento es analizable, la esencia del conocimiento modal está dada ya—sólo falta agregarle las condiciones triviales de verdad y creencia (además, quizá, de alguna condición anti-Gettier). Si el conocimiento no es analizable, faltaría revisar cómo se modela el conocimiento. Aquí voy a revisar el ejemplo mejor estudiado de epistemología primitivista: la filosofía de Williamson (2009 2001: cap. 7).

Williamson toma al conocimiento como un estado fáctico en el que uno está, con respecto a una proposición, cuando no podría estar fácilmente equivocado, cuando el riesgo de equivocarse es bajo. Tales estados se llaman *seguros*. Así, Williamson caracteriza al conocimiento como un estado fáctico y seguro.

La noción de seguridad se entiende modalmente: que uno no esté fácilmente equivocado con respecto a  $\langle p \rangle$  es que en ningún mundo posible cercano uno crea falsamente que  $p$  (donde el criterio de cercanía es contextual—y fuertemente determinado, consistentemente con el primitivismo, por la noción de *conocimiento* misma). Pero entonces, para llegar al conocimiento modal o esencial bajo la epistemología primitivista de Williamson, uno simplemente enmienda JUSTIFICACIÓN MODAL o JUSTIFICACIÓN ESENCIALISTA:

(CONOCIMIENTO MODAL—WILLIAMSON)

$S$  sabe que  $\Box p$  ( $\Diamond p$ ) si y sólo si:

- (i)  $S$  tiene una representación de  $\langle p \rangle$  que es suficientemente fiable,
- (ii)  $S$  tiene un proceso de producción de *simulaciones* que produce representaciones seguras en el sentido de Williamson,
- (iii) El proceso de simulación utiliza la representación fiable de  $\langle p \rangle$  y resulta en que  $S$  sabe que  $\Box p$  ( $\Diamond p$ ).

(CONOCIMIENTO ESENCIALISTA—WILLIAMSON)

$S$  sabe que  $\Box_x p$  (o de que  $\Diamond_x p$ ) si y sólo si:

- (i)  $S$  tiene una representación de  $\langle p \rangle$  que es suficientemente fiable,
- (ii)  $S$  tiene una representación que *tiende a lo esencial* de  $x$ ,
- (iii)  $S$  tiene un proceso de producción de *simulaciones* que produce representaciones seguras en el sentido de Williamson,
- (iv) El proceso de simulación utiliza las representaciones fiables y que tienden a lo esencial de  $x$  y de  $\langle p \rangle$  (respectivamente) y resulta en que  $S$  sabe que  $\Box_x p$  ( $\Diamond_x p$ ).

Vemos, entonces, que mi teoría tiene una agradable generalidad: su estructura puede adecuarse tanto a epistemologías que descomponen al conocimiento, como a las primitivistas.

## 6.54. creando representaciones fiables

Acabo de poner los fundamentos para una epistemología modal fiabilista, externista, que sólo apela a nuestra capacidad de representar y de simular. La llamaré *representacionalismo fiabilista*. Este es *naturalista* en el sentido en que deja la cuestión de cómo caracterizar estos procesos mentales para la investigación empírica, y reduce, como lo hace el fiabilismo, la terminología normativa a terminología puramente estadística y modelo-teórica (mediante la aplicación de la semántica general, cf. §6.3, arriba), además de la terminología de la ciencia cognitiva que sea necesaria para caracterizar a los proceso productores de representaciones, y a la capacidad de simular (de la que hablé en §4.1).

Me gustaría, y creo que sería interesante hacerlo, mostrar que el representacionalismo fiabilista propuesto en la sección pasada puede verse como subyacente a otras epistemologías modales, en el sentido en que *estas resultarán de agregarle compromisos extra a la propuesta presente*. (Por ejemplo, la de Williamson resultará de tomar a la capacidad de simulación en cuestión como imaginación, que desarrolla escenarios posibles a partir de suposiciones contrafácticas. Tales escenarios y suposiciones van a ser *representaciones* mentales.) Sin embargo, ya me he extendido demasiado, y es más urgente regresar a los constreñimientos que limitaron el espacio de epistemologías posibles, y que revisé al inicio de este capítulo (p. 142). Queremos que nuestra epistemología:

- (1) sea *compatible* con el estructuralismo absolutamente primitivista, defendido en el capítulo 3 (y por ello, que *no necesite* un reduccionismo de la modalidad),
- (2) dé al menos un inicio de respuesta a los tres problemas individuados en el capítulo 4,
- (3) sin por ello caer en el escepticismo global (*i.e.*, que acepte  $\neg$ MODAL GLOBAL),
- (4) satisfaga EXTERNISMO MÍNIMO, defendido en el capítulo 5,
- (5) involucre una postura fiabilista externista, defendida en el capítulo 5, y
- (6) evada la mayor cantidad de los problemas que aquejan a las principales epistemologías de la modalidad disponibles.

Es claro que las propuestas de arriba satisfacen (4) y (5), y espero haber motivado la idea de que también satisfacen (6). Nos quedan (1), (2) y (3). (1) es claramente el caso, pues nunca necesité apelar a la supuesta reducibilidad de la modalidad. De hecho, abajo diré más cómo la propuesta del capítulo 3 puede motivar ciertas ideas epistemológicamente relevantes. Nos faltan (2) y (3), y ahora diré más sobre ello.

Para (3), el representacionalismo fiabilista claramente no implica el escepticismo global de la modalidad, en la medida en que sea plausible que (i) tenemos una capacidad de simular (cf. §4.1), (ii) esa capacidad es fiable, y (iii) podemos producir representaciones fiables. Todo esto es plausible, y como mencioné en §4.1 y al inicio del cap. 2, tanto la ciencia cognitiva como ciencias fundamentales (como la física o las matemáticas) asumen que tenemos acceso epistémico a la modalidad. Estas asunciones no son contradichas por mi epistemología—de hecho, muy al contrario, la motivación de moverse desde creencias y conceptos a la noción más general de *representación*, es que así podemos juzgar las propiedades epistémicas de cosas como representaciones sensomotoras, simulaciones mentales o modelos matemáticos. Todo esto muestra que estoy lejos del escepticismo.

Queda (2). Voy a ir problema por problema, viendo cómo la propuesta de este capítulo resuelve, o al menos motiva un inicio de solución, para cada uno. Empezamos:

(PRIMER PROBLEMA)

¿Qué tan *lejos* puede estar un mundo posible, si es que hemos de tener creencias justificadas acerca de él?

Respuesta: tan lejos como sea compatible con un grado de fiabilidad representacional *suficiente* para el conocimiento. Tal grado, de nuevo, es plausiblemente determinado contextualmente, o al menos en los usos concretos de la teoría.

Seguimos:

(SEGUNDO PROBLEMA)

¿Qué habilidades cognitivas están involucradas en el logro de conocimiento modal? Los juicios filosóficos acerca de posibilidades remotas, ¿usan las mismas habilidades que los juicios sobre habilidades cercanas?

Respuesta: A lo segundo, sí. A lo primero: las habilidades que se usan al *representar*, y una o varias habilidades de *simular*. Las especificaciones de estas se dejan para la ciencia cognitiva; pero aquí se asume que tienen cierta estructura, la cual se explota para, junto con herramientas modelo-teóricas y probabilísticas, proponer la teoría.

Finalmente:

(TERCER PROBLEMA)

Al imaginar, concebir, ... etc., cierto hecho  $p$ , y con ello creer que  $\langle p \rangle$  es posible ¿qué otros hechos, y con cuánto detalle, necesitamos imaginar, concebir, ... etc., si es que nuestra creencia de que  $\langle p \rangle$  es posible ha de estar justificada?

Respuesta corta: el nivel de detalle necesario para dar *suficiente* fiabilidad.

Respuesta larga: esta pregunta invita una pregunta sobre la noción de representación fiable, y cerraré este apartado diciendo lo que pienso sobre esta.

¿Cómo podemos tener representaciones fiables? Aunque de hecho *las tenemos*, como (de nuevo) lo muestra el éxito de la física, por ejemplo (y como asume la ciencia cognitiva), todavía podemos proponer algunas heurísticas.

*Primero*, así como tenemos postulados acerca de regularidades que se extienden al pasado y al futuro, así parece que podemos tener postulados que se extienden a otros mundos posibles. De nuevo, el acercamiento estándar a la dinámica de los sistemas en física (clásica y cuántica), suele asumir ello: se postulan leyes que determinan espacios de posibilidad (cf. el inicio del cap. 2, arriba). Esto sugiere que, para al menos algunos contextos, podemos seguir la pista de la física y representar a los objetos de tal manera que podamos extender, para esas representaciones, regularidades que se dan de hecho, y que se obtengan variando parámetros *admisibles* (donde tal noción de admisibilidad dependerá de cada caso).

*Segundo*, la idea de variar parámetros *admisibles* también puede darnos pistas para representar objetos abstractos—si sabemos que ciertas estructuras abstractas son posibles, entonces, variando parámetros admisibles en su definición, podríamos representarnos otras estructuras de manera fiable (cf. Bricker 1991; para discusión, cf. Almog 1991).

*Tercero*, puede resultar que una cosa se reduzca a otra, o se fundamente en otra, o que haya una identificación teórica (como cuando el calor se identifica con la energía cinética promedio de

las moléculas). En estos casos, las leyes sobre las cosas fundamentales restringen las leyes sobre las derivadas. Si esto es así, podemos constreñir nuestras representaciones con estas leyes. (El caos reina cuando asumimos que las leyes mismas son contingentes, y que los objetos en cuestión pueden existir sin leyes.)

*Cuarto*, los objetos derivados también pueden tener sus posibilidades *qua* derivados (para propósitos epistémicos, es mejor representar a una silla como silla, no como un conjunto de partículas interactuando, si es que queremos representar las posibilidades de la silla.) Los objetos derivados estarán gobernados por leyes de alto nivel (*e.g.*, las leyes químicas, o de la dinámica macroscópica). Es buena idea constreñir nuestras representaciones con estas leyes.

Estas primeras cuatro heurísticas subrayan la importancia de las leyes para la representación fiable. De hecho, si seguimos la propuesta del capítulo 3, arriba, el conocimiento de la modalidad es más conocimiento de constreñimientos modales—que pueden resultar ser leyes. ¿Cómo conocemos leyes o constreñimientos? Una manera muy importante es por abducción—como con la teoría de Lewis de las leyes, buscamos el sistema más informativo y compacto para hacer sentido de las regularidades (en este caso, de las regularidades en el espacio modal), y con él postulamos las leyes como axiomas. No significa que las leyes sólo sean eso (si uno es primitivista de las leyes, como yo me veo inclinado a serlo, las leyes son más que axiomas en un sistema epistémico). Pero así podemos lograr conocimiento de ellas. (Véase también Biggs 2011).

*Quinto*, suponiendo que le encontramos lugar a la noción de esencia, podemos intentar representar a  $x$  como siendo su esencia. No sabemos si hay esencias, pero en la medida en que la noción juegue un papel explicativo, podemos postular que las hay. Para intentar representar a  $x$  como su esencia, nos preguntamos cuáles propiedades de  $x$  son las que lo definen, sin las cuales  $x$  no podría existir—y, quizá, las propiedades que explican las otras propiedades de  $x$ . Eso no necesita ser accesible a la intuición *a priori*—muy probablemente, necesitaremos apelar al conocimiento *a posteriori*.

*Sexto*: podemos establecer criterios de invarianza. Seleccionamos las “mejores” representaciones de  $x$ —lo cual habrá de definirse en los casos concretos, relativo a fines determinados, con marcos evaluativos dados, y con las representaciones disponibles—y entonces buscamos representar a  $x$  como aquello que es invariante entre ellas. Esto puede resultar en una buena guía para la esencia de  $x$ .

Todo esto nos da más razón para dudar de que el conocimiento modal y de esencias sea fructíferamente concebido como *a priori*: regularidades y representaciones adecuadas son típicamente influidas por el conocimiento *a posteriori* de una manera mucho más esencial que para permitirnos la mera posesión de conceptos (o representaciones).

Voy a cerrar la tesis explorando esta cuestión. Pero antes, recordemos que, para Kripke (1971), el conocimiento modal *a posteriori* se explicaba con el siguiente esquema:

- (1) Sabemos *a priori* que  $p \rightarrow \Box p$ ,
- (2) Sabemos *a posteriori* que  $p$ ,
- (3) Sabemos *a posteriori* que  $\Box p$ .

Como nuestro conocimiento necesita de conocimiento *a posteriori*, esto hace que, en estos casos, nuestro conocimiento de que  $\Box p$  sea *a posteriori*. Sin embargo, se asume (1) con base en casos como la necesidad de la identidad o tesis esencialistas basadas en la intuición. Seguramente, dado que la necesidad de la identidad es una verdad lógica, es una verdad *a priori*. Así que el caso



de identificaciones teóricas seguramente nos da necesidades que siguen el modelo kripkeano. Si es que hay tesis esencialistas que son conocidas por la mera intuición, esto nos daría otro ejemplo de (1).

Se ha cuestionado que todo el conocimiento *a posteriori* de necesidades funcione bajo este modelo (Bird 2008), al dar casos de conocimiento de condicionales con la forma de  $p \rightarrow \Box p$  que se conocen *a posteriori*. Pero creo que lo dicho arriba da motivos para generalizar más. Mucho—yo diría que, excepto por tesis lógicas y matemáticas, quizá *todo*—conocimiento de condicionales con la forma de  $p \rightarrow \Box p$  va a depender esencialmente de factores *a posteriori*—en la medida, al menos, en que conozcamos tales condicionales con base en las seis heurísticas que propuse arriba, pues casi todas ellas apelan esencialmente a factores que, en la mayoría de los casos, sólo nos son accesibles *a posteriori*.

Ahora veamos algunas consecuencias para la “*big picture*” del acercamiento de este capítulo. Con estas cierro la tesis.

## 6.6. después de la destrucción del Triángulo Dorado: Quine (otra vez), y la metafísica a oscuras

What is promised [by the foundationalist-internalist picture] is a way to factor our theorizing about the world into two parts: first, we form, *a priori*, a conception of the logical space of possibilities (stipulating what we shall mean by the words we use to describe the possibilities). Then we gather evidence to discover where in that space of possibilities the actual world is located. [...] *The central problem with this picture is that the meanings of our terms and the contents of our thoughts are determined by facts about our causal relations with things in the world, and with their properties and the relations between them, and so there is no general way, guaranteed to be neutral, to describe the space of possibilities in which we are trying to locate the world, or to characterize the evidence that would do the work of locating it in that space.*

—Robert Stalnaker, *Our Knowledge of the Internal World*

Los procesos productores y su funcionamiento no tienen por qué estar siempre disponibles a nuestro acceso consciente; al menos no de la manera inmediata que se necesitaría si uno adoptara una tesis internista fuerte. Pero en el capítulo pasado motivé justamente su opuesto: un externismo mínimo, bajo el cual las propiedades epistémicas pueden ser fijadas por factores a los que no tenemos un acceso especial, o sobre los cuales podemos no tener control.

Esta imagen de la epistemología y la representación es obviamente compatible con el *externismo del contenido semántico*—la idea de que el contenido semántico de nuestras representaciones, incluyendo palabras y conceptos, puede ser fijado por factores externos a nuestro acceso *a priori*. Cuando lo que nuestras representaciones representan no está fijado por nuestra mente consciente, no hay garantía alguna de que conozcamos, por pura introspección (entendimiento

de conceptos, intuición, competencia lingüística, conocimiento de las convenciones con que fijamos que algo representará, etcétera) lo que nuestras representaciones “pescan” en el mundo.

Si factores externos a nuestra conciencia fijan la semántica de nuestras representaciones, además de sus propiedades epistémicas, eso significa que la imagen de cuño esencialmente cartesiano, que toma a la mente y la semántica de lo que ella se representa como accesibles *a priori*, ha de ser desechada. La imagen externista que he venido defendiendo desde el capítulo anterior, y que aquí estoy aplicando, eliminan así la idea de que podemos saber *que* representamos al mundo, y saber *cómo* lo hacemos (con qué medios y de qué manera lo representamos como siendo), por análisis *a priori*.

Esto va contra la tendencia internista de otro acercamiento, muy popular, a la epistemología modal—el chalmersiano (Chalmers 2002, 2004). Bajo la imagen de Chalmers, aunque no siempre sabemos *a priori* a qué refieren nuestras palabras en el mundo—que él captura con su segunda intensión—sí tenemos acceso *a priori* a un aspecto del significado de nuestras palabras—el capturado por la primera intensión. La idea es que tenemos un “contenido estrecho” para cada palabra o concepto, una preconcepción de a qué estamos dispuestos a etiquetar con tal palabra o concepto, que es estable, independiente de cómo resulte el mundo. Cuando consideramos hipótesis sobre cómo podría de hecho resultar el mundo—cuando consideramos un escenario como real—ponemos en juego tal contenido estrecho para “decidir” a qué referirían nuestras palabras. (Por ejemplo, si resultara que vivimos en un mundo donde el líquido que llena los mares y los ríos y sale de las llaves es XYZ, estaríamos dispuestos, según Chalmers, a llamar a tal líquido “agua”.) La secuencia de tales “decisiones” de en dónde poner cada etiqueta de nuestras representaciones en el mundo fija su pegamento—es decir, su primera intensión.

El aparato de Chalmers ayuda, según él, a restaurar el *Triángulo Dorado*—la unión esencial entre el significado, la razón y la modalidad que la tradición de la semántica anti-descriptivista (la de Kripke, Putnam, Donnellan, Kaplan y Stalnaker) rompió con sus casos de necesidades *a posteriori* y contingencias *a priori*. El significado se une con la modalidad porque se obtiene abstractando a partir de un espacio de escenarios posibles—el espacio de hipótesis sobre cómo el mundo de hecho podría resultar ser—y la modalidad se une con la razón porque tales escenarios se definen como hipótesis compatibles con lo que conocemos *a priori*.

Restaurado el Triángulo Dorado, Chalmers y otros neo-descriptivistas creen haber fundamentado, ahora sí, la posibilidad de la metafísica apriorista, que busca conocer los contenidos de la estructura modal del mundo *a priori*, de una manera en principio independiente de la ciencia empírica.

Es una consecuencia de la postura epistemológica y metodológica que he venido promoviendo en esta segunda parte de la tesis, que el neodescriptivismo está errado. *El Triángulo Dorado se ha roto, y con ello la posibilidad de conocer las posibilidades del mundo de manera a priori*. La postura acerca de la representación que he venido defendiendo va mejor con un acercamiento *quineano* a la representación, con el que cerraré la tesis.

Después de que Quine rechazó el dualismo del significado de los positivistas lógicos (1952), motivado por el argumento de que la noción de analiticidad era irremediabilmente circular, propuso otra manera de entender la estructura de lo que él llamó “nuestra red de creencias”.

El argumento basado en la tesis de que la analiticidad es una noción circular, por supuesto, no necesariamente da una refutación final de la postulación de tal noción. (Podríamos simplemente

motivarla como un primitivo.) Pero no se necesita aquí. Pues bajo un entendimiento relativamente magro de la analiticidad, podemos ver por qué fundamentar la epistemología en ella—una idea que motiva la creencia en la estabilidad del Triángulo Dorado—es una idea metodológicamente errada.

Si entendemos la analiticidad como la propiedad que tiene una oración de ser verdadera meramente en virtud de lo que significan sus palabras componentes y la lógica, podemos generalizar. Digamos que una representación  $S$  que representa un hecho es *g-analítica* cuando es acerca de los contenidos de las representaciones  $r_1, \dots, r_n$ , que componen a  $S$ , y el hacedor de verdad de  $S$  es simplemente el hecho de que  $r_1, \dots, r_n$  están gobernadas por sus respectivas convenciones representacionales, más la lógica.

La *g-analiticidad* (y, en particular, la analiticidad) no ayudará en nada para fundamentar la posibilidad del conocimiento *a priori* a menos que las convenciones representacionales de nuestras representaciones sean accesibles *a priori*. Pero no lo son. Cuando aceptamos que las propiedades epistémicas de las representaciones pueden ser fijadas por factores a los que no tenemos acceso *a priori*, se vuelve difícil aceptar que las propiedades semánticas y meta-semánticas de nuestras representaciones también pueden ser fijadas *a priori*—pues la manera en que representamos al mundo está inseparablemente unida con lo que sabemos, incluso bajo la imagen internista que estamos criticando. Bajo esta imagen, un usuario de un sistema representacional es alguien que puede hacer ciertas cosas con tal sistema—tener tales o cuales intuiciones y creencias (al menos disposicionalmente), estar dispuesto a hacer tales o cuales inferencias, por ejemplo—y para hacer tales cosas se necesita un conocimiento de trasfondo. Pero cuando ese conocimiento de fondo a su vez pudo haber sido fijado por factores externistas—quizá, por ejemplo, por nuestra interacción con el ambiente mediante lo que en el capítulo 5 llamé Fuentes de Información Externa—tenemos entonces que contar como usuario de un cierto sistema representacional supone hechos que *no* son accesibles *a priori*.

Para nosotros, los teóricos post-internismo, post-racionalismo, la estructura de las representaciones es mucho menos simple que para los teóricos del contenido estrecho y otras construcciones internistas. Bajo la imagen externista, qué representan nuestros símbolos y signos y estados intencionales, y cómo lo hacen, y si es que lo hacen, está profundamente fijado por el mundo externo a nuestra cabeza y su poder de decisión, y es por ello mucho más dinámico. Si Quine propuso una red de creencias donde las más externas estaban más sujetas a la experiencia y las más internas en mucho menor grado, aquí podríamos proponer una estructura parecida para todas nuestras representaciones en general: *no* es que haya primeras intensiones que representan lo que unimos necesariamente a nuestras representaciones—sino que hay creencias acerca de esas representaciones que son mucho más centrales, que estaremos dispuestos a negar con mucha menos frecuencia que otras creencias menos centrales. *No* es que el término “agua” esté necesariamente ligado con la creencia de que el agua moja, o es un líquido incoloro e insípido, o es lo que llena los océanos y los mares—sino que estamos mucho menos dispuestos a rendir esas creencias frente al avance del conocimiento, que a rendir creencias como que el agua vale, por litro, unos 15 pesos mexicanos.

Así, cuando hemos abandonado el triángulo dorado, la estructura interna de nuestras representaciones no está determinada de una vez y para siempre *a priori*, sino que está a la espera de presiones externas que la modifican—que hacen que la manera misma en que concebimos al

mundo cambie, en respuesta a ese mundo.

El externista que está después de la fractura del triángulo dorado, entonces, no se asume como teniendo acceso privilegiado a las propiedades epistémicas o representacionales de sus representaciones. *Es hora de asumir esta lección en la epistemología de la modalidad*. Una muy importante tradición en la filosofía del lenguaje, de la mente, del conocimiento y de la ciencia la ha venido asumiendo desde la segunda mitad del siglo XX. *Es hora de que los metafísicos de la modalidad lo hagan también*.

Esta imagen general, entonces, motiva la tesis de que podríamos tener conocimiento modal—*sin saber que lo tenemos*. A diferencia del internismo chalmersiano, donde se asume que tenemos una *garantía* de que cierta manera de conocer el aspecto modal del mundo provee infalibilidad (cf. Sturgeon 2010), la imagen que he estado motivando aquí nos lleva a un falibilismo radical.

Con esto, se abre la puerta para la siguiente posibilidad: en muchos casos de disputas metafísicas sobre si tal o cual cosa es posible, puede suceder que sepamos que lo es (o que no). *Lo que no sabemos es si lo sabemos*. Según la teoría desarrollada aquí, básicamente, sabemos cuando nuestras representaciones son suficientemente fiables (en el sentido doble de haber sido producidas por mecanismos fiables, y representar fiablemente)—pero eso es compatible con que *no sepamos que lo son*.

Esto nos dice que, en un sentido preciso, podemos estar en la situación con la que inicié este capítulo: la de una persona que sabe que tiene la representación de un objeto—una “etiqueta” para él—y que intenta que tal representación se adecue a este—que se le “pegue”—pero que, como está a oscuras, no tiene manera de comprobar directamente si de hecho la etiqueta se le ha pegado.

Así nosotros. Así la metafísica de la modalidad y la esencialidad. No tenemos un acceso directo a la realidad modal y esencial, que nos permita corroborar si nuestras representaciones se adecuan a ella. No tenemos “intuiciones” ni podemos “concebir” acerca de tal realidad, de manera que el mero hecho de “intuir” o “concebir” nos muestre que estamos concibiendo cómo es ella en sí misma. *Vamos a oscuras, tanteando*. Pero lo mismo sucede con prácticamente todas nuestras otras prácticas representacionales. Eso, creo, nos indica que la metafísica no está, en principio, en un lugar peor que otras empresas cognitivas. Aunque no tengamos un acceso directo o *a priori* y *transparente* a su dominio—pues no lo tenemos a casi ningún otro dominio.

Otros verán en esto una motivación para el escepticismo y el nihilismo, o para un dionisiaco/posmoderno “libre juego de signos”. No yo. Conocer nuestra condición como agentes *sobre* el mundo comienza por reconocer las posibilidades que tenemos también como pacientes *en* él, interactuando de manera inextricable, y por ello con incertidumbre y falibilidad inevitables—pero, también, con el ajuste correcto de nuestras representaciones con el mundo.

## CONCLUSIONES

Aquí hice una investigación sobre la modalidad desde dos perspectivas: una lógico-metafísica, y una epistemológica. Parecería que las conclusiones a las que llegué son completamente independientes, pero no creo que sea así. Para cerrar la tesis, visitaré esas conclusiones, e intentaré responder si forman un todo coherente y sistemático que sugiera más caminos de investigación.

En la primera parte de la tesis, donde hice una investigación lógico-metafísica de la modalidad, mostré cómo es que la lógica modal es simplemente una estructura formal que necesita de una interpretación metafísica. En el camino, levanté algunas dudas sobre la asunción, universalmente compartida, de que es la semántica de marcos de Kripke la que debe privilegiarse. (No *argumenté* contra tal asunción, pero argumenté que ella debe fundamentarse, ante la existencia de diferentes semánticas.) Después revisé la interpretación metafísica más popular—en términos de mundos posibles. Sobre esta interpretación se suele poner la teoría de la esencia como simplemente modalidad *de re*, aunque en un capítulo posterior recordé que Fine puso contraejemplos a tal teoría. Tras haber revisado nociones generales de reducción y fundamentación, expuse y argumenté contra dos tipos de teorías: la reductiva—en la que critiqué al esencialismo, derivado del trabajo de Fine—y una fundamentadora—en la que critiqué a una versión fundamentadora del esencialismo fineano, así como a lo que llamé las teorías *no-no* de la modalidad, que incluyen a varios tipos de “mundismo” y a la teoría de las potencias de Jacobs.

Esto me llevó a explorar una postura—que llamé *estructuralismo absolutista* o *absolutamente no-reductivo*—que diera mejor cuenta de la fundamentalidad de la modalidad. Primero noté cómo podríamos tomarnos en serio la idea de que la modalidad es fundamental, al aplicar la teoría de Sider al operador de posibilidad. Luego construí un marco formal que reflejara mejor la fundamentalidad de la modalidad, al construir los mundos posibles mediante la interacción de lo que llamé “función modal”, que codifica los principios modales primitivos, y operadores de construcción que operan sobre individuos y propiedades. Mostré cómo esta postura es compatible con el esencialismo al estilo fineano, si uno deja de lado al reductivismo.

En la segunda parte me aproximé a la modalidad desde la epistemología. Primero, argumenté que no ha sido bloqueada la posibilidad del conocimiento de la modalidad, y argumenté que, de acuerdo a la ciencia cognitiva, mucho de nuestro conocimiento de hecho depende de conocimiento modal. Entonces individúe tres preguntas legítimas que, argumenté en el último capítulo, mi teoría responde. También defendí el acercamiento externista y fiabilista, mostrando que el externismo se requiere para dar cuenta de la posibilidad del conocimiento. Defendí al fiabilismo del problema de la generalidad de Conee y Feldman, mostrando que descansaba sobre presupuestos que están en una fuerte tensión.

Finalmente, exploré la postura que llamo *representacionalismo fiabilista*. Para esto, exploré la noción general de representación, y mostré cómo extender la ideología epistémica y semántica a esta. En particular, argumenté que se puede extender la teoría de modelos estándar para dar cuenta del contenido semántico de cualquier cosa que cuente como representación. Aparejando esto con la postura fiabilista, pude desarrollar una epistemología modal y de las esencias externista y fiabilista.

¿Hay alguna imagen resultante de las posturas caracterizadas y defendidas en ambas partes de la tesis, o tenemos ideas divergentes, independientes totalmente entre sí?

Primero, me gustaría notar que la metafísica modal argumentada en la primera parte, no fue propuesta con bases epistémicas—aunque, como noté en la conclusión del cuarto capítulo, tiene la agradable propiedad de no estar sujeta al problema benacerrafiano. Una segunda cosa que me gustaría notar es que la epistemología propuesta en la primera parte tampoco depende directamente de la metafísica.

Sin embargo, si nos tomamos en serio ambas ideas, puede que resulten estar sugiriendo una imagen interesante. (Al menos, yo planeo explorarla más, posteriormente.) Pues tenemos una metafísica primitivista que, de todos modos, muestra la interacción profunda entre modalidad, leyes, individuación y fundamentalidad. Y por otro, tenemos una epistemología representacionista que no exige más que capacidades de simular y representar. Tenemos, pues, un mundo fundamentalmente modal, gobernado por leyes; mientras que nuestro acceso a estas, al mundo y a su estructura modal, se puede dar a los variados mecanismos productores de representaciones, que tienden a ajustarse a estas leyes.

O, como dice Omnés en el epígrafe que abre el sexto capítulo: *The laws that structure the representation we form of Reality are an image of its own order—las leyes que estructuran la representación que nos formamos de la Realidad son una imagen de su propio orden.* Quizá esta sea la frase que mejor describe la motivación detrás del variado y alargado trayecto que pasamos durante esta tesis. Espero haberle hecho suficiente justicia. Y espero tener oportunidad de seguir explorando sus implicaciones en trabajos futuros.

## REFERENCIAS

- ADAMS, Robert (1974) "Theories of Actuality". *Noûs*, **8**, #3: 211-231.
- \_\_\_\_\_ (1979) "Primitive Thisness and Primitive Identity". *Journal of Philosophy*, **76**, #1: 5-26.
- ALMOG, Joseph (1991) "The Plenitude of Structures and the Scarcity of Possibilities". *Journal of Philosophy*, **88**, #11: 620-622.
- \_\_\_\_\_ (2010) "Nature Without Essence". *Journal of Philosophy*, **107**, #7: 360-383.
- ALSTON, William (1995) "How to Think about Reliability". *Philosophical Topics*: 1-29.
- ALVARADO MARAMBIO, Tomás (2007) "Bidimensionalismo Epistémico y el Espacio Ontológico Modal". *Crítica*, **39**, #117: 3-18.
- ARMSTRONG, David (1980) *Nominalism and Realism. Volume I: Universals and Scientific Realism. Volume II: A Theory of Universals*. Cambridge University Press.
- \_\_\_\_\_ (1989) *A Combinatorial Theory of Possibility*. Cambridge University Press.
- AU, T.K. (1983) "Chinese and English Counterfactuals: the Sapir-Whorf hypothesis revisited", *Cognition* **15**: 155-187.
- BEALER, George (2002) "Modal Epistemology and the Rationalist Renaissance". Ensayo 1 de Gendler & Hawthorne (eds.) 2002.
- BECHTEL, William & ABRAHAMSEN, Adele (1991) "Beyond the Exclusively Propositional Era". En James Fetzer (ed.), *Epistemology and Cognition*: 121-151. Kluwer Academic Publishers.
- BECKETT, Samuel (1956) *Waiting for Godot*. Faber and Faber.
- BEEBEE, James R. (2004) "The Generality Problem, Statistical Relevance, and the Tri-level Hypothesis". *Noûs* **38**, #1: 263-294.
- BEILOCK, Sian & LYONS, Ian (2009) "Expertise and the Mental Simulation of Action". Ensayo 2 en Markman *et al.* 2009: 21-34.
- BENACERRAF, Paul (1973) "Mathematical Truth", *Journal of Philosophy*, **70**: 661-679.
- BENNETT, Karen (2010) "Construction Area (No Hard Hat Required)". *Philosophical Studies* **154**, #1: 79-104.
- \_\_\_\_\_ (2011) "By Our Bootstraps". *Philosophical Perspectives*, **25**: Metaphysics: 27-41.
- BIGGS, Stephen (2011) "Abduction and Modality". *Philosophy and Phenomenological Research*, **83**, #1: 283-326.
- BILLE, Philip (2005) "A Survey of Tree Edit Distance and Related Problems". *Theoretical Computer Science* **337**: 217-239.
- BIRD, Alexander (2008) "Remarks on Our Knowledge of Modal Facts". *Norsk Filosofisk Tidsskrift*, **43**, #1: 53-59.
- \_\_\_\_\_ (2010) "A Posteriori Knowledge of Natural Kind Essences: A Defence". *Philosophical Topics*, **35**: 293-312.
- BLACKBURN, Patrick (ed.) (2007) *Handbook of Modal Logic* (col. *Studies in Logic and Practical Reasoning*, Vol. 3). Elsevier.
- BLACKBURN, Patrick & VAN BENTHEM, Johan (2007) "Modal Logic: A Semantic Perspective", ensayo 1 en (Blackburn 2007).
- BLACKBURN, Patrick & DE RIJKE, Maarten & VENEMA, Yde (2001) *Modal Logic*. Cambridge Tracts in Theoretical Computer Science.
- BLUMSON, Ben (2010) "Pictures, Perspectives, and Possibility". *Philosophical Studies*, **149**: 135-

- 151.
- \_\_\_\_\_ (manuscrito) "Depiction and Convention".
- BOLINSKA, Agnes (2013) "Epistemic Representation, Informativeness and the Aim of Faithful Representation". *Synthese*, **190**, #2: 219-234.
- BRAÜNER, Torben & GHILARDI, Silvio (2007) "First-Order Modal Logic", ensayo 9 en (Blackburn 2007).
- BRICKER, Philip (1991) "Plenitude of Possible Structures". *The Journal of Philosophy*, **88**, #11: 607-619.
- BRINGSJORD, Selmer (1985) "Are There Set-theoretical Possible Worlds?". *Analysis*, **45**: 64.
- BUCKNER, Randy & CARROLL, Daniel C. (2006) "Self-projection and the Brain". *TRENDS in Cognitive Sciences*, **11**, #2: 49-57.
- BUENO, Otávio (2009) "Models of Reduction". *Principia*, **13**, #3: 269-282.
- BUENO, Otávio & FRENCH, Steven (2011) "How Theories Represent". *British Journal for the Philosophy of Science*, **62**: 857-894.
- BURGE, Tyler (1979) "Individualism and the Mental". En French, Uehling, & Wettstein (eds.): *Midwest Studies in Philosophy*, **IV**: 73–121. University of Minnesota Press.
- BUTTERWORTH, Brian (2001) "What Makes a Prodigy?". *Nature Neuroscience*, **4**, #1: 11–12.
- \_\_\_\_\_ (2005a) "The Development of Arithmetical Abilities". *Journal of Child Psychology and Psychiatry*, **46**, #1: 3–18.
- \_\_\_\_\_ (2005b) "Developmental Dyscalculia", *Handbook of Mathematical Cognition*: 455-467.
- BYRNE, Alex & PRYOR, James (2006) "Bad Intentions". Ensayo 3 en García-Carpintero & Maciá 2006.
- BYRNE, Ruth (2002) "Mental Models and Counterfactual Thoughts About What Might Have Been". *TRENDS in Cognitive Sciences*, **6**, #10: 426-431.
- \_\_\_\_\_ (2005) *The Rational Imagination: How people create alternatives to reality*. MIT Press.
- CHAKRAVARTY, Anjan (2010) "Informational Versus Functional Theories of Scientific Representation". *Synthese* **172**: 197-213.
- CHALMERS, David (1996) *The Conscious Mind*. Oxford University Press.
- \_\_\_\_\_ (2002) "Does Conceivability Entail Possibility?". Ensayo 3 de Gendler & Hawthorne (eds.) 2002.
- \_\_\_\_\_ (2004) "The Foundations of Two-Dimensional Semantics". Ensayo 4 de García-Carpintero & Maciá (eds), 2006.
- CHENG, Edward (2009) "A Practical Solution to the Reference Class Problem". *Columbia Law Review*, **109**: 2081-2105.
- CHURCHLAND, Paul (2001) "What Happens to Reliabilism When it is Liberated From the Propositional Attitudes?". *Philosophical Topics*, **29**, #1: 91-112.
- CLARK, Michael & LIGGINS, David (2012) "Recent Work on Grounding". Por aparecer en *Analysis*.
- CONNOR, Earl & FELDMAN, Richard (1998) "The Generality Problem for Reliabilism". *Philosophical Studies*, **89**, #1: 1-29.
- \_\_\_\_\_ (2004a) "Afterword", en E. Conee & R. Feldman, *Evidentialism: essays in epistemology*: 159–165. Oxford Clarendon Press.
- \_\_\_\_\_ (2004) "Internalism Defended", en E. Conee & R. Feldman, *Evidentialism: essays in epistemology*: 53–80. Oxford Clarendon Press.
- CORREIA, Fabrice (2006) "Generic Essence, Objectual Essence, And Modality". *Noûs*, **40**, #4: 753-



- 767.
- \_\_\_\_\_ (2007) "(Finean) Essence and (Priorean) Modality". *Dialectica*, **61**, #1: 63-84.
- \_\_\_\_\_ (2012) "On the Reduction of Necessity to Essence". *Philosophy and Phenomenological Research*, **84**, #3: 639–653.
- DASGUPTA, Shamik (2009) "Individuals: An Essay in Revisionary Metaphysics". *Philosophical Studies*, **145**: 37-67.
- DECETY, Jean & STEVENS, Jennifer (2009) "Action Representation and its Role in Social Interaction", ensayo 1 en Markman *et al.* 2009: 3-20.
- DEHAENE, Stanislas (1997) *The Number Sense: How mind creates mathematics*. Oxford University Press.
- \_\_\_\_\_ & PIAZZA, Manuela & PINEL, Philippe & COHEN, Laurent (2003) "Three Parietal Circuits for Number Processing", *Cognitive Neuropsychology*, **20**: 487–506.
- DELEUZE, Gilles (1994) *Difference and Repetition*. Columbia University Press. (Traducción de Paul Patton de *Différence et Répétition*. Presses Universitaires de France, 1968.)
- DELANDA, Manuel (2002) *Intensive Science and Virtual Philosophy*. Continuum.
- DEROSSET, Louis (2012) "Possible Worlds for Modal Primitivists". *Journal of Philosophical Logic*: 1-23.
- DERRIDA, Jacques (1996) "Envío". Traducción de Patricio Peñalver, en Derrida, Jacques, *La Deconstrucción en las Fronteras de la Filosofía*. Paidós.
- DIVERS, John & MELIA, Joseph (2002) "The Analytic Limit of Genuine Modal Realism". *Mind*, **111**, #441: 15-36.
- DONNELLAN, Keith (1966) "Reference and Definite Descriptions". *The Philosophical Review*, **77**: 281-304.
- DOWELL, J.L. (2008) "Serious Metaphysics and the Vindication of Reductions". *Philosophical Studies*, **139**: 91-110.
- DRETSKE, Fred (2003) "Externalism and Self-Knowledge". Ensayo 6 de Susana Nuccetelli (ed.), *New Essays on Semantic Externalism and Self-Knowledge*. MIT Press.
- DUNAWAY, Billy & EDMONDS, Anna & MANLEY, David (2012) "The Folk Probably Do Think What You Think They Think". Por aparecer en *Australasian Journal of Philosophy*.
- FIELD, Hartry (1989) *Realism, Mathematics, and Modality*. Oxford: Blackwell.
- \_\_\_\_\_ (2009) "Epistemology Without Metaphysics", *Philosophical Studies*, **143**, #2: 249-290.
- FINE, Kit (1974) "An Incomplete Logic Containing S4". *Theoria*, **40** : 23–29.
- \_\_\_\_\_ (1977) "Prior on the Construction of Possible Worlds and Instants". Reimpreso como el ensayo 4 de (Fine 2005).
- \_\_\_\_\_ (1991) "The Study of Ontology". *Noûs* **25**: 263-294.
- \_\_\_\_\_ (1994) "Essence and Modality", *Philosophical Perspectives*, **8**: 1-16.
- \_\_\_\_\_ (1995) "The Logic of Essence". *Journal of Philosophical Logic*, **24**: 241-273.
- \_\_\_\_\_ (1997) "Senses of Essence", ensayo 5 de Walter Sinnott-Armstrong (ed.): *Modality, Morality and Belief: Essays in honor of Ruth Barcan Marcus*. Cambridge University Press.
- \_\_\_\_\_ (2000) "Semantics for the Logic of Essence", *Journal of Philosophical Logic*, **29**: 543–584.
- \_\_\_\_\_ (2001) "The Question of Realism". *Philosophers' Imprint* **1**, #1: 1–30.
- \_\_\_\_\_ (2002) "The Varieties of Modality". Ensayo 6 de Gendler & Hawthorne (eds.) 2002.
- \_\_\_\_\_ (2005) *Modality and Tense*. Oxford University Press.

- \_\_\_\_\_ (2007a) *Semantic Relationism*. Blackwell Publishing.
- \_\_\_\_\_ (2007b) "Response to Fabrice Correia". *Dialectica*, **61**, #1: 85–88.
- \_\_\_\_\_ (2010a) "Towards a Theory of Part". *Journal of Philosophy*, **107**, #11: 559–589.
- \_\_\_\_\_ (2010b) "Some Puzzles of Ground". *Notre Dame Journal of Formal Logic* **51** #1: 97–118.
- \_\_\_\_\_ (2012a) "The Pure Logic of Ground". *Review of Symbolic Logic* **5** #1: 1–25.
- \_\_\_\_\_ (2012b) "A Guide to Ground". Capítulo 1 de F. Correia & B. Schnieder (eds.): *Metaphysical Grounding: Understanding the Structure of Reality*. Cambridge University Press.
- FLORIDI, Luciano (2004) "On the Logical Unsolvability of the Gettier Problem", *Synthese*, **142**: 61–79.
- FODOR, Jerry (1983) *Modularity of Mind: An essay on faculty psychology*. MIT Press.
- \_\_\_\_\_ (2000) "Is Science Biologically Possible?", ensayo 16 de su *In Critical Condition*: 189–202. MIT Press.
- FOLEY, Richard (2002) "Review of Timothy Williamson's *Knowledge and its Limits*". *Mind*, **111**, #443: 718–726.
- FREGE, Gottlob (1892/1997) "On Sense and Reference". En Michael Beaney (ed.), *The Frege Reader*: 151–171. Blackwell Publishing.
- GARCÍA-CARPINTERO, Manuel & MACIÀ, Josep (eds.) (2002) *Two-Dimensional Semantics*. Oxford University Press.
- GEIRSSON, Heimir (2005) "Conceivability and Defeasible Modal Justification", *Philosophical Studies*, **122**: 279–304.
- GENDLER, Tamar Szabó & HAWTHORNE, John (eds.) (2002) *Conceivability and Possibility*. Oxford University Press.
- GETTIER, Edmund (1963) "Is Justified True Belief Knowledge?". *Analysis* **23**, #6: 121–123.
- GIBBONS, John (2006) "Access Externalism". *Mind*, **115**, #457: 19–39.
- GIERE, Ronald (1985) "Constructive Realism". En *Images of Science: Essays on Realism and Empiricism with a Reply by Bas C. Van Fraassen*; editado por Paul M. Churchland & Clifford A. Hooker. University of Chicago Press.
- GILOVICH, T. & WANG, R. F. & REGAN, D. & NISHINA, S. (2003) "Regrets of Action and Inaction Across Cultures". *Journal of Cross-Cultural Psychology*, **34**: 61–71.
- GOLDMAN, Alvin (1976) "Discrimination and Perceptual Knowledge". *Journal of Philosophy* **73**: 771–791.
- \_\_\_\_\_ (1979) "What is Justified Belief?". En G. Pappas (ed.): *Justification and Knowledge*. Reidel. Reimpreso en Alvin Goldman: *Liaisons: Philosophy Meets the Cognitive and Social Sciences*. MIT Press.
- \_\_\_\_\_ (1986) *Epistemology and Cognition*. Harvard University Press.
- \_\_\_\_\_ (1988) "Strong and Weak Justification" *Philosophical Perspectives*, **2**: *Epistemology*: 51–69.
- \_\_\_\_\_ (1992) "Epistemic Folkways and Scientific Epistemology". En Alvin Goldman, *Liaisons: Philosophy Meets the Cognitive and Social Sciences*: 155–175. MIT Press.
- \_\_\_\_\_ (1999) *Knowledge in a Social World*. Oxford University Press.
- \_\_\_\_\_ (2001) "The Unity of the Epistemic Virtues". En Abrol Fairweather & Linda Zagzebski (eds): *Virtue Epistemology: Essays on Epistemic Virtue and Responsibility*: 30–48. Oxford University Press.

- \_\_\_\_\_ (2008) "Reliabilism", *The Stanford Encyclopedia of Philosophy* (Spring 2011 Edition), Edward N. Zalta (ed.). <<http://plato.stanford.edu/archives/spr2011/entries/reliabilism/>>.
- \_\_\_\_\_ (2009) "Internalism, Externalism, and the Architecture of Justification". *Journal of Philosophy*, **106**: 306-38.
- \_\_\_\_\_ (2009b) "Recursive Tracking versus Process Reliabilism". *Philosophy and Phenomenology Research*, **79**, #1: 223-230.
- \_\_\_\_\_ & Olsson, Eric (2009) "Reliabilism and the Value of Knowledge", en A. Haddock & A. Millar & D. H. Pritchard (eds.): *Epistemic Value*. Oxford University Press.
- GORMAN, Michael (2005) "The Essential and the Accidental". *Ratio*, **18**: 276-289.
- GOULD, Stephen Jay & LEWONTIN, Richard C. (1979) "The Spandrels of San Marco and the Panglossian Paradigm: A critique of the adaptationist programme", *Proceedings of the Royal Society of London, Series B: Biological Sciences*, **205**, #1161: *The Evolution of Adaptation by Natural Selection*: 581-598.
- GREENBERG, Gabriel (2011) *The Semiotic Spectrum*. Tesis doctoral, departamento de filosofía de la universidad Rutgers, Estados Unidos. URL: <[http://gjgreenberg.bol.ucla.edu/greenberg\\_the\\_semiotic\\_spectrum.pdf](http://gjgreenberg.bol.ucla.edu/greenberg_the_semiotic_spectrum.pdf)> (accedido en Diciembre, 2013).
- GRICE, Paul (1957) "Meaning". *Philosophical Review*, **66**, #3.
- HACKING, Ian (1983) *Representing and Intervening: Introductory Topics in the Philosophy of Natural Science*. Cambridge University Press.
- HAMKINS, Joel (2010) "The Set-theoretical Multiverse". *Philosophy and Model Theory Conference*, Paris, Junio de 2010.
- \_\_\_\_\_ & LÖWE, Benedikt (2008) "The Modal Logic of Forcing". *Transactions of the American Mathematical Society*, **360**, #4: 1793-1818.
- HALE, Bob (1996) "Absolute Necessities". *Philosophical Perspectives*, **10**, *Metaphysics*: 93-117.
- \_\_\_\_\_ (2013) *Necessary Beings*. Oxford University Press.
- HAWKE, Peter (2011) "Van Inwagen's Modal Skepticism", *Philosophical Studies*, **153**: 351-364.
- HAUSER, Marc & MACNEILAGE, Pogen & WARE, Molly (1996) "Numerical Representations in Primates". *Proceedings of the National Academy of Sciences of the USA*, **93**: 1514-1517.
- HAVERKAMP, Nick & SCHULZ, Moritz (2012) "A Note on Comparative Probability". *Erkenntnis*, **76**, #3: 395-402.
- HÁYEK, Alan (2007) "Interpretations of Probability", *The Stanford Encyclopedia of Philosophy*, (Winter 2012 Edition), Edward N. Zalta (ed.), <<http://plato.stanford.edu/archives/win2012/entries/probability-interpret/>>.
- \_\_\_\_\_ (2007b) "The Reference Class Problem is Your Problem, Too". *Synthese*, **156**: 563-585.
- HEIDEGGER, Martin (1927/1951) *Ser y Tiempo*. Fondo de Cultura Económica. (Traducción de José Gaos de *Sein und Zeit*, 1927.)
- HELLER, Mark (1995) "The Simple Solution to the Generality Problem", *Noûs*, **29**, #4: 501-515.
- \_\_\_\_\_ (1998a) "Property Counterparts in Ersatz Worlds". *Journal of Philosophy*, **95**, #6: 293-316.
- \_\_\_\_\_ (1998b) "Five Layers of Interpretation for Possible Worlds". *Philosophical Studies*, **90**: 205-214.
- HETHERINGTON, Stephen (2012) "The Significance of Fallibilism Within Gettier's Challenge: A Case Study", *Philosophia*, **40**: 539-547.

- HITTMAIR-DELAZER, Margarette & SEMENZA, C. & DENES, G. (1994) "Concepts and Facts in Calculation", *Brain*, **117**: 715-728.
- HITTMAIR-DELAZER, Margarette & SAILER, U. & BENKE, T. (1995) "Impaired Arithmetic Facts but Intact Conceptual Knowledge—a Single Case Study of Dyscalculia", *Cortex*, **31**: 139-147.
- HUBBARD, Edward M. & PIAZZA, Manuela & PINEL, Philippe & DEHAENE, Stanislas (2005) "Interactions Between Number and Space in Parietal Cortex", *Nature Reviews: Neuroscience*, **6**: 435-448.
- IZARD, Véronique & SANN, Coralie & SPELKE, Elizabeth & STRERI, Arlette (2009) "Newborn Infants Perceive Abstract Numbers", *Proceedings of the National Academy of Sciences of the USA*, **106**, #25: 10382–10385.
- JACOBS, Jonathan. (2010). "A Powers Theory of Modality: or, How I Learned to Stop Worrying and Reject Possible Worlds". *Philosophical Studies* **151**: 227–248.
- JACKSON, Frank (1998) *From Metaphysics to Ethics: A defense of conceptual analysis*. Oxford University Press.
- \_\_\_\_\_ & PETTIT, Philip (1995) "Moral Functionalism and Moral Motivation", *Philosophical Quarterly*, **45**, #178: 20-40.
- JANSANA, Ramón (1990) *Una Introducción a la Lógica Modal*. Tecnos.
- JENKINS, Carrie S. (2008a) *Grounding Concepts: An empirical basis for arithmetical knowledge*. Oxford University Press.
- \_\_\_\_\_ (2008b) "Modal Knowledge, Counterfactual Knowledge, and the Role of Experience". *Philosophical Quarterly*, **58**, #233: 693-701.
- KAHNEMAN, Daniel & TVERSKY, Amos (1982) "The Simulation Heuristic", en *Judgment Under Uncertainty: Heuristics and Biases* (Daniel Kahneman, ed.): 201–208. Cambridge University Press.
- KAHNEMAN, Daniel & MILLER, D. T. (1986) "Norm Theory: comparing reality to its alternatives", *Psychological Review*, **93**: 136–153.
- KAPLAN, David (1975) "How to Russell a Frege-Church". *Journal of Philosophy* **72**, #19: 716–729.
- KATZ, Jerrold (2014) *Sentido, Referencia y Filosofía*. Facultad de Filosofía y Letras, UNAM—Ítaca. (Traducción de Lourdes Valdivia de *Sense, Reference, and Philosophy*, 2004, Oxford University Press.)
- KLEIN, Peter (2007) "Human Knowledge and the Infinite Progress of Reasoning", *Philosophical Studies*, **134**: 1-17.
- KORNBLITH, Hilary (2002) *Knowledge and its Place in Nature*. Oxford University Press.
- KOSLICKI, Kathrin (2012) "Essence, Necessity, and Explanation". En T. Tahko (ed.): *Contemporary Aristotelian Metaphysics*. Cambridge University Press.
- KRIPKE, Saul (1971) "Identidad y Necesidad". *Cuadernos de Crítica*, #7. Instituto de Investigaciones Filosóficas, UNAM. (Traducción de Margarita Valdés de "Identity and Necessity", reimpresso en Jaegwon Kim & Ernest Sosa (eds.) (1999) *Metaphysics: An Anthology*: 72-89. Blackwell).
- \_\_\_\_\_ (1975) "Outline of a Theory of Truth". *Journal of Philosophy*, **72**: 690–716.
- \_\_\_\_\_ (2005) *El Nombrar y la Necesidad*. Instituto de Investigaciones Filosóficas, UNAM. (Traducción de Margarita Valdés de *Naming and Necessity*, 1980, Harvard University Press).
- LADYMAN, James & ROSS, Don (2007) *Every Thing Must Go: Metaphysics Naturalized* (con David Spurrett & John Collier). Oxford University Press.
- LAPORTE, Joseph (2006) "Rigid Designators". *The Stanford Encyclopedia of Philosophy* (Sum-

- mer 2011 Edition), Edward N. Zalta (ed.),  
 URL = <<http://plato.stanford.edu/archives/sum2011/entries/rigid-designators/>>.
- LEUNBERGER, Stephan (2006). "A New Problem of Descriptive Power". *Journal of Philosophy*, **103**, #3: 145- 158.
- LEWIS, David (1966). "An Argument for the Identity Theory". *Journal of Philosophy*, **63**: 17- 25.
- \_\_\_\_\_ (1969) *Convention: A Philosophical Study*. Harvard University Press.
- \_\_\_\_\_ (1970) "How to Define Theoretical Terms", *Journal of Philosophy*, **67**: 427-46.
- \_\_\_\_\_ (1973) *Counterfactuals*. Blackwell Publishers.
- \_\_\_\_\_ (1979) "Counterfactual Dependence and Time's Arrow". *Noûs* **13**, #4: 455-476.
- \_\_\_\_\_ (1980) "A Subjectivist's Guide to Objective Probability". En Richard C. Jeffrey (ed.) *Studies in Inductive Logic and Probability*. University of California Press.
- \_\_\_\_\_ (1984) "New Work for the Theory of Universals". Reimpreso como el primer ensayo en (Lewis 1994b).
- \_\_\_\_\_ (1984) "Putnam's Paradox". Reimpreso como el segundo ensayo de (Lewis 1999).
- \_\_\_\_\_ (1984) *Philosophical Papers, Vol. I*. Oxford University Press.
- \_\_\_\_\_ (1986) *On the Plurality of Worlds*. Oxford University Press.
- \_\_\_\_\_ (1994a) "Reduction of Mind" Reimpreso como el ensayo 18 en su (1994b).
- \_\_\_\_\_ (1994b) *Papers in Metaphysics and Epistemology*. Cambridge University Press.
- LINNEBO, Øystein (2009). "Bad Company Tamed". *Synthese*, **170**: 371-391.
- LLAMAS, Erick (*manuscrito*) "Descriptions, Essence and the Mind-Body Problem".
- LONGWORTH, Guy (2012) "John Langshaw Austin". *The Stanford Encyclopedia of Philosophy* (Winter 2013 Edition), Edward N. Zalta (ed.), URL = <<http://plato.stanford.edu/archives/win2013/entries/austin-jl/>>.
- LOWE, Jonathan (2013) "What is The Source of Our Knowledge of Modal Truths?". *Mind*, **121**, #484: 919-950.
- LYCAN, William (1979) "The Trouble with Possible Worlds". In M.J. Loux (ed.), *The Possible and the Actual*: 274-316. Cornell University Press.
- \_\_\_\_\_ & SHAPIRO, Stewart (1986) "Actuality and Essence", *Midwest Studies in Philosophy XI: Studies in Essentialism*: 343-377.
- LYON, Aidan & COLYVAN, Mark (2008) "The Explanatory Power of Phase Spaces". *Philosophia Mathematica*, **3**, #16: 227-243.
- MACHERY, Edouard & MALLON, Ron & NICHOLS, Shaun & STICH, Stephen (2004) "Semantics, Cross-Cultural Style". *Cognition*, **92**, #3: B1-B12.
- MACKIE, Penelope (2006) *How Things Might Have Been*. Oxford University Press.
- MAGUIRE, Barry (2013) "Defending David Lewis's Modal Reduction". Por aparecer en *Philosophical Studies*.
- MANDEL, David & HILTON, Denis & CATELLANI, Patrizia (eds.) (2005) *The Psychology of Counterfactual Reasoning*. Routledge.
- MARKMAN, Keith & KLEIN, William & SUHR, Julie (eds.) (2009) *Handbook of Imagination and Mental Simulation*. Psychology Press.
- MARKMAN, Keith & McMULLEN, Matthew (2005) "Reflective and Evaluative Modes of Mental Simulation", ensayo 5 en (Mandel *et al.*, 2005): 77-93.
- MARTÍNEZ, Manolo (2013) "Ideal Negative Conceivability and the Halting Problem". *Erkenntnis*,

- 78, #5: 979-990.
- MAUDLIN, Tim (2007) *The Metaphysics Within Physics*. Oxford University Press.
- MCGEE, Vann (2006) "There Are Many Things". En Judith Jarvis Thomson y Alex Byrne (eds.), *Content and Modality: Themes From the Philosophy of Robert Stalnaker*. Oxford University Press.
- MCMICHAEL, Alan (1983) "A Problem for Actualism About Possible Worlds". *The Philosophical Review*, **92**, #1: 49-66.
- MELIA, Joseph (2001) "Reducing Possibilities to Language". *Analysis*, **61**, #1: 19-29.
- MENZEL, Christopher (2012) "Sets and Worlds Again". *Analysis*, **72**, #2: 304-9.
- MOCHIZUKI, Shinichi (2012) "Inter-Universal Teichmüller Theory, IV: log-Volume Computations and Set-Theoretic Foundations". Manuscrito disponible en: <<http://www.kurims.kyoto-u.ac.jp/~motizuki/Inter-universal%20Teichmuller%20Theory%20IV.pdf>> (accedido en diciembre de 2013).
- MOSS, Sarah (2013) "Epistemology Formalized". *Philosophical Review*, **122**, #1: 1-43.
- NANAY, Bence (2010a) "Neither Moralists, nor Scientists: We are counterfactually reasoning animals", *Behavioral and Brain Sciences*, **33**, #4: 347-348.
- \_\_\_\_\_ (2010b) "Morality or Modality? What Does the Attribution of Intentionality Depend on?", *Canadian Journal of Philosophy*, **40**: 28-40.
- NIMTZ, Christian (2004) "Two-Dimensional and Natural Kind Terms". *Synthese*, **138**: 125-148.
- NORTH, Jill (2009) "The 'Structure' of Physics: A Case Study". *Journal of Philosophy*, **106**,: 57-88.
- \_\_\_\_\_ (2012) "The Structure of a Quantum World". En *The Wave Function* (editado por Alyssa Ney & David Albert). Oxford University Press.
- NOZICK, Robert (2001) *Invariances: The Structure of the Objective World*. Harvard University Press.
- NUCCETELLI, Susana (2003) "Introduction". En Susana Nuccetelli (ed.), *New Essays on Semantic Externalism and Self-Knowledge*. MIT Press.
- OMNÉS, Roland (1999) *Quantum Philosophy*. Princeton University Press.
- OPPENHEIM, Paul & PUTNAM, Hillary (1958) "The Unity of Science as a Working Hypothesis". En H. Feigl, M. Scriven, & G. Maxwell (eds.): *Minnesota Studies in the Philosophy of Science, Volume II*: 3-36. University of Minnesota Press.
- PAPPAS, George (2005) "Internalist vs. Externalist Conceptions of Epistemic Justification". *The Stanford Encyclopedia of Philosophy* (Fall 2013 Edition), Edward N. Zalta (ed.), URL = <<http://plato.stanford.edu/archives/fall2013/entries/justep-intext/>>.
- PEACOCKE, Christopher (1992) *A Study of Concepts*. MIT Press.
- \_\_\_\_\_ (1999) *Being Known*. Oxford University Press.
- PIAZZA, Manuela & GIACOMINI, E., & LEBIHAN, D. & DEHAENE, Stanislas (2003) "Single-trial Classification of Parallel Pre-attentive and Serial Attentive Processes Using Functional Magnetic Resonance Imaging", *Proceedings of the Royal Society B*, **270**, #1521: 1237-1245.
- PLANTINGA, Alvin (1974) *The Nature of Necessity*. Oxford University Press.
- \_\_\_\_\_ (1976) "Actualism and Possible Worlds". *Theoria* #42: 139-160.
- \_\_\_\_\_ (1979) "De Essentia". *Grazer Philosophische Studien* **7**: 101-121.
- PRIEST, Graham (2006) *In Contradiction: A Study of the Transconsistent (Second Edition)*. Oxford University Press.
- PUTNAM, Hillary (1975) "The Meaning of 'Meaning'". En su *Philosophical Papers, Vol. II: Mind, Language, and Reality*. Cambridge University Press.
- QUINE, W.v.O. (1952) "Two Dogmas of Empiricism". Reimpreso como el ensayo 2 de su *From a*

- Logical Point of View*. Harvard University Press.
- \_\_\_\_\_ (1953a) "Three Grades of Modal Involvement". Reimpreso como el ensayo 13 de su *The Ways of Paradox* (1966): 156-154. Random House.
- \_\_\_\_\_ (1953b) "Reference and Modality". Reimpreso como el ensayo 9 de su *From a Logical Point of View*. Harvard University Press.
- \_\_\_\_\_ "Propositional Objects". *Crítica*, **2**, #5: .
- RAMACHANDRAN, V.S. & HUBBARD, E.M. (2001) "Synaesthesia: A Window Into Perception, Thought and Language". *Journal of Consciousness Studies*, **8**, #12: 3-34.
- RAYO, Agustín (2008) "Vague Representation". *Mind*, **117**: 329-373.
- \_\_\_\_\_ (2012) "An Actualists' Guide to Quantifying In". *Crítica*, **44**, #132: 3-34.
- RICKLES, Dean (2008) *Symmetry, Structure, and Spacetime*. Elsevier.
- RESTALL, Greg (1996) "Truthmakers, Entailment and Necessity". *Australasian Journal of Philosophy*, **74**, #2: 331-340.
- ROBERTSON, Teresa (2000) "Essentialism: Origin and Order". *Mind*, **109**, #434: 299-307.
- \_\_\_\_\_ (2007) "Essential vs. Accidental Properties". *The Stanford Encyclopedia of Philosophy* (Fall 2008 Edition), Edward N. Zalta (ed.), URL=<<http://plato.stanford.edu/archives/fall2008/entries/essential-accidental/>>.
- ROMERO, Carlos (2010) *Lógica Modal, Paradojas de Indiscernibilidad y Teorías Esencialistas*. Tesis de Licenciatura. Facultad de Filosofía y Letras, UNAM.
- \_\_\_\_\_ (2014) "Identidad, Posibilidad y Esencia". Ensayo 3 de Valdivia 2014b.
- ROSEN, Gideon (1990) "Modal Fictionalism". *Mind*, **99**, #395: 327-354.
- \_\_\_\_\_ (2010). "Metaphysical Dependence: Grounding and Reduction". En Bob Hale y Aviv Hoffman, (eds.): *Modality: Metaphysics, Logic and Epistemology*. Oxford University Press.
- ROY, Tony (1993) "Worlds and Modality". *The Philosophical Review*, **102**, #3: 335-361.
- SALMON, Nathan (1986) *Frege's Puzzle*. Ridgeview Publishing.
- \_\_\_\_\_ (1987) "Existence". Reeditado como el ensayo 1 en Salmon 2005.
- \_\_\_\_\_ (1989) "The Logic of What Might Have Been". Reeditado como el ensayo 7 en Salmon 2005.
- \_\_\_\_\_ (1998) "Nonexistence". Reeditado como el ensayo 2 en Salmon 2005.
- \_\_\_\_\_ (2005) *Metaphysics, Mathematics and Meaning*. Oxford University Press.
- SCHAFFER, Jonathan (2009) "On What Grounds What". En David Chalmers, David Manley y Ryan Wasserman, (eds.): *Metametaphysics: New Essays on the Foundations of Ontology*. Oxford University Press.
- \_\_\_\_\_ (2010) "Spacetime the One Substance", *Philosophical Studies*, **145**, #1: 131-148.
- SCHNIEDER, Benjamin. (2004). "Once More: Bradleyan Regresses". En Herbert Hochberg & Kevin Mulligan (eds.): *On Relations and Predicates*: 219-56. Ontos-Verlag.
- SCHROETER, Laura (2005) "Considering Empty Worlds as Actual". *Australasian Journal of Philosophy*, **83**, #3: 331-47.
- SEARLE, John (1958) "Proper Names". *Mind*, **67**: 166-173.
- \_\_\_\_\_ (1997) *The Construction of Social Reality*. Free Press.
- SEBENS, Charles (2013) "A Laws-First Introduction to Quantum Field Theory". Manuscrito, disponible en: <<https://sites.google.com/site/csebens/>> (accedido Diciembre 2013).
- SEGURA, Susana & MORRIS, Michael (2005) "Scenario Simulations in Learning", ensayo 6 en (Mandel *et al.*, 2005): 94-109.
- SHALKOWSKI, Scott (1994) "The Ontological Ground of Alethic Modality". *The Philosophical*

- Review*, **103**, #4: 669-688.
- \_\_\_\_\_ (1997) "Essentialism and Absolute Necessity". *Acta Analytica*, **12**, #4: 41-56.
- SIDELLE, Alan (1989) *Necessity, Essence, and Individuation: A Defense of Conventionalism*. Cornell University Press.
- SIDER, Theodore (2002) "The Ersatz Pluriverse", *Journal of Philosophy*, **99**: 279-315.
- \_\_\_\_\_ (2003) "Reductive Theories of Modality". En *The Oxford Handbook of Metaphysics* (M. J. Loux & D. W. Zimmerman, eds.): 180-208. Oxford University Press.
- \_\_\_\_\_ (2012) *Writing the Book of the World*. Oxford University Press.
- SMITH, Peter (1998) *Explaining Chaos*. Cambridge University Press.
- SOBER, Elliot (1996) "Evolution and Optimality – Feathers, Bowling Balls, and the Thesis of Adaptationism". *Philosophical Exchange*, **26**, 41-57. Reimpreso en David Hull & Michael Ruse (eds.): *The Philosophy of Biology*. Oxford University Press, 1998. (La cita es a la reimpresión.)
- SOSA, Ernest (1991) *Knowledge in Perspective*. Cambridge: Cambridge University Press.
- \_\_\_\_\_ (2003) "The Place of Truth in Epistemology". En M. DePaul & L. Zagzebski (eds.): *Intellectual Virtue: Perspectives from ethics and epistemology*: 155-179. Oxford University Press.
- \_\_\_\_\_ (2007) *A Virtue Epistemology: Apt Belief and Reflective Knowledge, Vol. I*. Oxford University Press.
- SPEAKS, Jeffrey (2010) "Epistemic Two-Dimensionalism and the Epistemic Argument". *Australasian Journal of Philosophy*, **88**, #1: 59-78.
- SPELLMAN, Barbara & KINCANNON, Alexandra & STOSE, Stephen (2005) "The Relation Between Counterfactual and Causal Reasoning", ensayo 2 en (Mandel *et al.* 2005): 29-43.
- STALNAKER, Robert (1976) "Possible Worlds". *Noûs*, **10**, #1: 65-75.
- \_\_\_\_\_ (1978) "Assertion". *Syntax and Semantics*, **9**: 315-32.
- \_\_\_\_\_ (2006) "Assertion Revisited: On the Interpretation of Two-Dimensional Modal Semantics". Ensayo 13 de García & Macià (eds.), 2006.
- \_\_\_\_\_ (2008) *Our Knowledge of the Internal World*. Oxford University Press.
- STANLEY, Jason (1997) "Names and Rigid Designation". Ensayo 22 de Bob Hale & Crispin Wright (eds.): *A Companion to the Philosophy of Language*, 1997. Blackwell Publishers.
- STICH, Stephen (1990) *The Fragmentation of Reason: Preface to a pragmatic theory of cognitive evaluation*. MIT Press.
- STURGEON, Scott (2010) "Apriorism About Modality". Ensayo 13 de Hale & Hoffman (eds.) *Modality: Metaphysics, Logic and Epistemology*, 2010. Oxford University Press.
- STREVEN, Michael (2004) "The Explanatory Role of the Notion of Representation". Manuscrito, URL: <<http://www.strevens.org/research/expln/mimeoexpln.pdf>> (accedido en Diciembre, 2013).
- SUDDENDORF, Thomas & ADDIS, Donna Rose & CORBALLIS, Michael (2009) "Mental Time Travel and the Shaping of the Human Mind", *Philosophical Transactions of the Royal Society, B*, **364**: 1317-1324.
- SVEINSDÓTTIR, Ásta (2008) "Essentiality Conferred". *Philosophical Studies*, **140**: 135-148.
- \_\_\_\_\_ (2013) "Knowledge of Essence: The Conferralist Story". *Philosophical Studies*, **168**, #1: 315-32.
- SZPUNAR, Karl K. & McDERMOTT, Kathleen B. (2009) "Episodic Future Thought: Remembering the Past to Imagine the Future", ensayo 8 en (Markman *et al.* 2009): 119-129.
- THOMASSON, S. K. (1974) "An Incompleteness Theorem in Modal Logic". *Theoria*, **40**: 150-158.



- VAIDYA, Anand (2007) "The Epistemology of Modality". *The Stanford Encyclopedia of Philosophy* (Winter 2011 Edition), Edward N. Zalta (ed.), URL = <http://plato.stanford.edu/archives/win2011/entries/modality-epistemology/>.
- VALDIVIA, Lourdes (2014a) "La Identidad del Significado y la Creencia". Ensayo 7 de Valdivia (ed.) 2014b.
- \_\_\_\_\_ (ed.) (2014b) *La Identidad: su Semántica y su Metafísica. Una Aproximación Desde la Filosofía Analítica*. Facultad de Filosofía y Letras, UNAM.
- VAN BENTHEM, Johan (1984) "Correspondence Theory". En *Handbook of Philosophical Logic*, Second Edition, Vol. 3: 325-409.
- VAN CLEVE, James (2003) "Is Knowledge Easy—Or Impossible? Externalism as the Only Alternative to Skepticism". En Stephen Luper (ed.), *The Skeptics: Contemporary Essays*. Ashgate.
- VAN INWAGEN, Peter (1998) "Modal Epistemology", *Philosophical Studies*, **92**: 67–84.
- VAIDYA, Anand (2007) "The Epistemology of Modality". *The Stanford Encyclopedia of Philosophy* (Winter 2011 Edition), Edward N. Zalta (ed.). <http://plato.stanford.edu/archives/win2011/entries/modality-epistemology/>.
- WALLACE, David (2010) "Decoherence and Ontology (or, How I Learned to Stop Worrying and Love FAPP)". Ensayo 1 de Saunders, Simon & Barrett, Jonathan & Kent, Adrian & Wallace, David (eds.): *Many Worlds? Everett, Quantum Theory, and Reality*. Oxford University Press.
- WANG, Jennifer (2013) "From Combinatorialism to Primitivism". *Australasian Journal of Philosophy*, **91** #3: 535-554.
- WEATHERSON, Brian (2003) "What Good Are Counterexamples?". *Philosophical Studies*, **115**, #1: 1-31.
- WEISBERG, Jonatha (*por aparecer*) "You've Come a Long Way, Bayesians". Disponible en <http://www.utm.utoronto.ca/~weisber3/articles/JPL40-Bayesianism.pdf> (accedido en enero 2014).
- WHITTLE, Bruno (2010) "There Are Brute Necessities". *The Philosophical Quarterly*, **60**, #238: 149–159.
- WILLIAMS, Michael (1991) *Unnatural Doubts: Epistemological Realism and the Basis of Scepticism*. Oxford University Press.
- WILLIAMS, J. R. G. (2010) "Fundamental and Derivative Truths". *Mind*, **119**: 103-141.
- WILLIAMSON, Timothy (2000) *Knowledge and its Limits*. Oxford University Press.
- \_\_\_\_\_ (2002) "Peacocke's Theory of Modality". *Philosophy and Phenomenological Research*, **64**, #3: 649-654.
- \_\_\_\_\_ (2007) *The Philosophy of Philosophy*. Routledge.
- \_\_\_\_\_ (2007b) "How Probable is an Infinite Sequence of Heads?". *Analysis*, **67**, #3: 173-180.
- \_\_\_\_\_ (2009) "Probability and Danger". *The Amherst Lecture in Philosophy*, **4**: 1–35. <http://www.amherstlecture.org/williamson2009/>.
- \_\_\_\_\_ (2009b) "Reply to Alvin Goldman". En Patrick Greenough & Duncan Pritchard (eds.), *Williamson on Knowledge*: 305-312. Oxford University Press.
- \_\_\_\_\_ (2010) "Reclaiming the Imagination". *The New York Times*, Agosto 15, 2010. <http://opinionator.blogs.nytimes.com/2010/08/15/reclaiming-the-imagination/>
- \_\_\_\_\_ (2013a) *Modal Logic as Metaphysics*. Oxford University Press.
- \_\_\_\_\_ (2013b) "How Deep is the Distinction Between A Priori and A Posteriori Knowledge?". Ensayo 13 de Albert Casullo & Joshua Thurow (eds.): *The A Priori in Philosophy*. Oxford University Press.

- WONG, Elaine & GALINSKY, Adam & KRAY, Laura (2009) "The Counterfactual Mind-Set: A decade of research". Ensayo 11 en (Markman *et al.* 2009): 161-174.
- WUNDERLICH, Mark (2003) "Vector Reliability". *Synthese* **136**, #2: 237-262.
- WYNN, Karen (1992) "Addition and Subtraction by Human Infants". *Nature*, **358**: 749-750.
- YABLO, Stephen (1993) "Is Conceivability a Guide to Possibility?". *Philosophy and Phenomenological Research*, **5**, #1: 1-42.
- \_\_\_\_\_ (2002) "Coulda, Woulda, Shoulda". Ensayo 13 de Gendler & Hawthorne (eds.) 2002.
- \_\_\_\_\_ (*manuscrito*) "Semantic Arithmetic". Disponible en su sitio:  
<<http://www.mit.edu/~yablo/semantearithmetic.pdf>>
- ZAGZEBSKI, Linda (1994) "The Inescapability of Gettier Problems". *The Philosophical Quarterly* **44**, #174: 65-73.
- \_\_\_\_\_ (1996) *Virtues of the Mind: An inquiry into the nature of virtue and the ethical foundations of knowledge*. Cambridge University Press.
- ZALTA, Edward (2006) "Essence and Modality". *Mind*, **115**, #459: 659-694.