

LÓGICA MODAL, PARADOJAS DE
INDISCERNIBILIDAD Y TEORÍAS
ESENCIALISTAS

TESIS

QUE PARA OBTENER EL TÍTULO DE:
LICENCIADO EN FILOSOFIA

PRESENTA:

Carlos Alberto Romero Castillo

ASESORA DE TESIS:

Dra. Lourdes Valdivia Dounce

México, D. F. Ciudad Universitaria, 2010



Universidad Nacional
Autónoma de México



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.

Agradecimientos

El primer agradecimiento es para mi familia: mis abuelos, tías y primos, y en especial para mi hermano y mi hermana, que han sido una compañía esencial desde que nacieron; y, por supuesto, a mi mamá y mi papá –una conjunción de cariño y rectitud cuya ejemplaridad apenas comienzo a comprender, conjunción siempre guiada por su amor y su experiencia de vida.

También me gusta agradecer a la Doctora Lourdes Valdivia, mi asesora de tesis y coordinadora del seminario *Metafísica del Significado* (el proyecto de CONACyT 57331). La influencia de Lourdes recorre toda esta tesis, y le estoy tremendamente agradecido por el apoyo que me ha brindado al escribirla; a tal punto que esto que presento sería otra cosa si ella no hubiera estado ahí para asesorarme. Lourdes ha sido no sólo una profesora remarcable en cuanto a atención y habilidad para entrenar una nueva generación de filósofos, sino una verdadera amiga y maestra que goza de sabiduría y buena voluntad. El proyecto *Metafísica del Significado* ha venido funcionando como ejemplo de lo que la inteligencia y la buena fe pueden lograr cuando se proponen la organización por propia cuenta. Quiero agradecer de nuevo a Lourdes por haber logrado establecer este grupo; por supuesto, también quiero agradecer a cada uno de los miembros de éste por seguir contribuyendo con él y brindando no sólo un ambiente que da pie a sanas sesiones argumentativas, sino también a la amistad y al compañerismo.

Otro enorme ejemplo de organización con buena voluntad e inteligencia es el seminario *Lógicaclara*, creado para reflexionar filosóficamente sobre la didáctica, que después aplicamos en los cursos intersemestrales de lógica iniciados por César López. Quiero agradecer a los miembros de este seminario por haber sido mis compañeros y amigos (y haberme soportado tantas horas en el ágora).

Un agradecimiento que no podría dejar pasar es a mis sinodales: Dr. Raymundo Morado Estrada, Dr. Mario Gómez Torrente, Mtra. Natalia Luna Luna y Dra. Maria Inés Pazos; quienes gentilmente me apoyaron para finalizar esta tesis.

También están las amigas y los amigos que me he ganado o en la facultad –por otra motivación que el leer y discutir argumentos–, o en otros momentos de mi vida, amistades todas cuya compañía me ha hecho aprender y comprender cosas de la vida que uno no comprende en la biblioteca. Quisiera agradecerles la compañía, la diversión y el aprendizaje que he logrado junto a ellas y ellos en

diferentes etapas.

Blaise Pascal distinguió el Dios de Abraham y Jacob frente al Dios de los geómetras y los filósofos. Mientras que las pruebas lógicas ofrecidas por grandes genios de la filosofía *more geometrico* siguen siendo discutidas, no puedo dejar de preguntarme (Kant llamó *Metaphysica naturalis* a algo como este impulso, a este hablar donde deberíamos callar) cómo debo agradecer a ese mismo Dios de Abraham y Jacob, cuya existencia seguimos sin demostrar axiomáticamente.

Finalmente, se cuenta que Albert Einstein agradeció alguna vez el fuerte impacto que el café que tomaba por las tardes tuvo en el desarrollo de sus ideas. Me siento comprometido a hacer un agradecimiento análogo e igualmente pintoresco (si es que algo de este estilo tiene sentido) por la existencia de la cafeína y su poder de inspiración intelectual, aún si los resultados que aquí presento no pueden compararse con los del alemán.

Tabla de Contenidos

AGRADECIMIENTOS	3
0. INTRODUCCIÓN	
0.1 El problema del cambio	4
0.2. El problema de la modalidad <i>de re</i>	5
0.3. Combinando respuestas	8
0.4. Lógica modal y teorías de la modalidad <i>de re</i>	10
0.5. Paradojas	12
0.6. Esta tesis	13
0.7. Cerrando la introducción	15
1. LOS PRINCIPIOS Y LA LÓGICA	
1.1. Los principios metafísicos	16
1.2. La lógica	22
1.3. Recuento	26
2. LA PARADOJA DE INDISCERNIBILIDAD	
2.1. Preliminares	28
2.2. La paradoja de Roderick Chisholm	28
2.3. La paradoja de Penelope Mackie	33
2.4. El argumento de Saul Kripke por la esencialidad de <i>ser ficticio</i> ...	35
2.5 Recuento	39
3. TEORÍAS ESENCIALISTAS	
3.1. Preliminares	41
3.2. Formulando el esencialismo	41
3.3. ¿Evitamos la Paradoja apelando a una teoría esencialista?	43
3.3.1. Nuestra mejor teoría esencialista... ..	45
3.3.2. ...Y sus problemas	48
4. LA ESTRATEGIA $\neg 4$	
4.1. Preliminares	52
4.2. Rechazando el axioma 4	54
4.3. Del esencialismo a $\neg 4$	57
CONCLUSIONES	60
REFERENCIAS	62

O. Introducción

O.1. El problema del cambio

Esta tesis...

...

...Al momento de comenzar a escribir esto, me di cuenta de que había derramado el café sobre el teclado de mi computadora. Así que, después de notar que había quedado inutilizado, me dispuse a cambiarlo por un teclado nuevo. Pero hace meses que también me vi obligado a cambiar el disco duro por uno nuevo, pues el viejo había quedado lleno. De hecho, poco después de comprar mi computadora, cambié su unidad DVD por una Blue-Ray. Ayer que llevé mi *laptop* con el técnico y él le cambió el teclado, me pregunté lo siguiente: después del cambio de la unidad DVD por una Blue-Ray, después del cambio de disco duro por uno nuevo y después del cambio del teclado por uno nuevo, ¿Cuántos otros cambios podría hacer sin verme orillado a decir que tengo una computadora diferente de la que compré?

Es un hecho ordinario que las cosas cambian. Por ejemplo, sabemos que obras arquitectónicas del pasado han sido restauradas una y otra vez. Claro que han permanecido –al menos la mayoría de ellas– en el lugar donde fueron construidas, pero también es cierto que un enorme número de sus piezas ha sido intercambiado por piezas nuevas. Y podemos imaginar que en el futuro otro enorme número de ellas será así intercambiado. De hecho, parece que nada nos impide imaginar que *todas* las piezas originales sean cambiadas por piezas nuevas –piedra por piedra, mosaico por mosaico, tabla por tabla.

La pregunta que nos invita a la perplejidad es la siguiente: ¿Cuántas piedras (mosaicos, tablas...) podemos cambiar antes de decir que nos encontramos ante un objeto distinto al original? ¿Cuál es el *límite* que no debemos cruzar al cambiar un objeto, antes de encontrarnos con otro objeto?¹

De hecho, podemos dar un caso general, que abstrae las características comunes de los ejemplos de la computadora y los monumentos –las

1 Aquí debemos notar una dificultad para dar cabal explicación de nuestras ideas. Al decir que un objeto “desaparece” quiero decir que ese objeto en particular se destruyó. Claro que después de cambiarle todas las piezas a mi computadora todavía habrá un objeto que lleve en mi mochila, que esté en mi escritorio, etcétera. La pregunta es si ese objeto es la misma computadora que la primera, exactamente el mismo x tal que $x =$ mi computadora original. La dificultad para exponer la idea es que es difícil decir que un objeto “desaparece” o que “pierde su identidad”, pues parecemos estar diciendo que un mismo objeto x pasa a ser otro objeto y ; o que x ya no está donde estaba pero sí en otro lado. Pero esto no es lo que quiero expresar, sino más bien que donde estaba x ahora está otra cosa, y ; que x simplemente dejó de existir; que fue destruido y en su lugar apareció (vino a la existencia) otro objeto.

características comunes que nos importan si lo que queremos hacer es una investigación filosófica sobre el cambio. Supongamos entonces que un objeto O consiste de n piezas. Supongamos que en un momento, llamémoslo t_1 , le quitamos una pieza P_1 a O , la que sustituimos por una pieza P_1' . Imaginemos que en un momento posterior, t_2 , le quitamos a O otra pieza, P_2 , que sustituimos por la pieza P_2' . Así que podemos preguntarnos qué pasará en el momento t_n , cuando le hayamos quitado a O todas sus piezas originales para sustituirlas por otras n piezas diferentes: ¿Seguiremos teniendo enfrente de nosotros al mismo O original, o a otro objeto completamente distinto, O^* ? Muchos nos sentiremos tentados a decir que tendremos a otro objeto distinto. Pero también es difícil decir cuál número m , menor a n , es tal que: al intercambiar exactamente m piezas de O , tenemos otro objeto completamente distinto a O . Es decir, es difícil decir cuál es el número que marca el límite entre O y otro objeto distinto.

No sólo tenemos este problema. Imaginemos que un ingeniero amigo mío está armando su propia computadora, y acepta gustosamente las antiguas piezas de la mía. Las repara y, con ellas, arma su computadora. Así que tiene una computadora con el antiguo teclado, la unidad DVD y el disco duro de la mía. Después de unos meses, también le regalo la memoria RAM y hasta la tarjeta madre. Yo, por mi lado, compro una nueva memoria RAM y una nueva tarjeta madre. Después de varios meses —quizá hasta un par de años—, le he regalado ya todas las piezas de mi computadora original, piezas que yo, por supuesto, he sustituido por nuevas.

La pregunta ahora es: ¿Cuál, si es que alguna de las dos, es la computadora que compré originalmente: la que yo tengo o la que tiene mi amigo?

El caso general es el siguiente: supongamos que un objeto O tiene n piezas en un momento t_0 . En un momento t_1 , le quitamos una pieza P_1 , que sustituimos por una pieza P_1' . Además, tenemos otro objeto, Q . En t_2 , le agregamos P_1 (la pieza original de O) a Q . Y así hasta que, en t_m , todas las piezas originales de O le han sido colocadas a Q ; y cada pieza original de O (cada P_i) ha sido sustituida por otra pieza P_i' . La pregunta, entonces, es: en t_m , ¿Es O , Q , o ninguno de los dos, el objeto original que teníamos en t_0 ?

El problema de cuánto podemos cambiar un objeto antes de encontrarnos con un objeto distinto es un antiguo problema filosófico. Thomas Hobbes lo presentó en su tratado *De Corpore* (HOBBS 1655) con un ejemplo que remitía a la antigüedad: el famoso ejemplo de la *barca de Teseo*. Supongamos que la barca de Teseo consistía en tres grandes piezas, que llamaremos, respectivamente, A , B y C . Supongamos que, por otro lado, tenemos otras tres grandes piezas (D , E y F). Y luego imaginemos que sustituimos A por D , B por E y C por F . Así que donde teníamos la barca de Teseo habrá un objeto constituido por D , E y F . Llamemos a este objeto, sin prejuizar por el momento si es o no el objeto

original, *barco uno*. Imaginemos entonces que con las piezas originales de la barca de Teseo construimos otro barco, *barco dos*, que entonces llega a estar constituido de *A*, *B* y *C*. Preguntamos entonces cuál, si es que alguno, entre *barco uno* y *barco dos* es la mismísima barca de Teseo. Por extraño que parezca, cualquiera de las cuatro respuestas posibles (*barco uno* = *barca de Teseo*; *barco dos* = *barca de Teseo*; ninguno es la barca de Teseo; ambos *barco uno* y *barco dos* son, en algún sentido, la misma barca de Teseo) puede fundamentarse en una teoría filosófica sobre el cambio y la identidad –teoría que puede ser, por supuesto, discutida y argumentada.

Por el momento podemos dejar pendiente la cuestión de qué tipo de teoría filosófica aceptaríamos al responder de una u otra manera a este problema del cambio a través del tiempo, para pasar a una problemática parecida que surge, por así decir, en otra dimensión –una dimensión donde no necesariamente hablamos del tiempo, sino de circunstancias posibles y de diferentes propiedades de los objetos en ellas.

0.2. El problema de la modalidad *de re*

En esta tesis defenderé que cierta concepción acerca de los objetos es correcta. Esta concepción es la base de cierto tipo de sistemas lógicos (de los cuales hablaré más adelante), y suele ser conocida como modalidad *de re* (por el latín “*res*”, *cosa*). La idea de la modalidad *de re*, *grosso modo*, es que los objetos pueden o podrían ser diferentes a como de hecho son.² Defenderé a la modalidad *de re* frente a un tipo de paradojas que surgen cuando intentamos hacer una teoría filosófica sobre ella. Ahora debemos ver con algo más de detalle cuál es la idea detrás de esta teoría filosófica.

Vimos en la sección anterior qué pasa cuando nos preguntamos sobre los cambios que un objeto sufre en el tiempo, y qué pasa cuando pasamos partes de un objeto a otro. Pero en lugar de preguntarnos qué pasa cuando realizamos cambios físicos en objetos actuales, podemos preguntarnos qué pasa cuando esos mismos cambios *podrían haberse dado*. Es decir: en lugar de preguntar si la computadora con el teclado, el disco duro y la unidad distintos es la misma computadora que la que compré, preguntamos si yo *podría* realizar esos cambios y seguir teniendo esa misma computadora. En lugar de preguntarnos

2 Podríamos entender a la modalidad *de re* como un tipo de “cambio” que se da *no* en el tiempo –o, más bien, *no exclusivamente* en el tiempo–, sino en *circunstancias posibles*. (Claramente, el presente, pasado y futuro reales son circunstancias posibles, pero no toda circunstancia posible es real, *actual*). Parece que el concepto de *cambio* implica –o al menos está muy relacionado– con el tiempo, y es justamente por este riesgo de confusión que evitaré hablar de “cambio”. *Modalidad de re* es un concepto más general, que incluye tanto a todo el tiempo real como a tiempos posibles que no son reales.

cuál (si es que alguna) de las dos computadoras —la que yo tengo ahora o la que tiene mi amigo con las piezas originales— es la misma computadora que compré, decimos que podría haber realizado esos intercambios, y entonces preguntar cuál *sería* mi computadora. Veremos que aquí surge una problemática filosófica muy amplia. En esta tesis, básicamente, examinaré cierto acercamiento a esta problemática, acercamiento conocido como esencialismo. El esencialismo está fundado sobre la idea de que los objetos sí podrían cambiar y seguir permaneciendo los mismos —es decir, está fundado sobre la modalidad *de re*—, aunque afirma que los objetos *no* podrían cambiar completamente.

Vista a vuelo de pájaro, la estructura de mi tesis es la siguiente: primero, haré explícitos los principios sobre los que estaremos volviendo en el desarrollo de la argumentación. Entonces presentaré un tipo de problemas que surgen cuando aceptamos estos principios y los aplicamos a la teorización filosófica sobre la modalidad *de re*. Finalmente, expondré cómo el esencialismo, salida propuesta para este tipo de problemas, sufre de problemas análogos. Estos problemas se evitan si rechazamos un alegado principio de cierto sistema lógico. Así que concluiré, junto con el filósofo Nathan Salmon, que debemos rechazar tal principio.

El lector deberá concederme un poco más de su tiempo antes de que exponga con algo de más profundidad qué es lo que haré en esta tesis. Primero me gustaría revisar con un poco más de detalle la problemática general en la que esta tesis se sumerge. Así que regresemos a la exposición de los problemas.

Podemos notar cómo surge el problema de la modalidad *de re* haciéndonos una pregunta. Con el ejemplo de las obras arquitectónicas, la pregunta sería así: ¿Qué tantas partes *podría* (incluso si no las cambia de hecho) cambiar el Templo Mayor, antes de que fuera un edificio diferente? ¿Dónde está el punto tal que, a partir de él, hemos llegado a otro objeto?

De hecho, podemos extender la pregunta, hablando no sólo de *partes*, sino de *propiedades*. (*Grosso modo*, una propiedad es una característica de un objeto: la propiedad *ser humano* es una propiedad que yo tengo (pues yo soy un ser humano), así como la propiedad *ser estudiante de filosofía*.) Y hay filósofos que han argumentado que existen distintos tipos de propiedades: unas más “naturales”, unas más apegadas a la identidad del objeto, unas que son propiedades de otras propiedades, etcétera.

Así que podemos hacer dos *tipos* de preguntas: ¿Cuántas propiedades podría cambiar mi computadora antes de “desaparecer del mapa”, antes de que nosotros nos encontremos con otro objeto distinto? ¿Cuántas propiedades podría intercambiar mi computadora con otro objeto, y seguir siendo el mismo objeto? Y ¿Cuáles, o de qué tipo, podría cambiar y seguir siendo la misma computadora? ¿Cuáles, o de qué tipo, podría intercambiar y seguir siendo la

misma computadora?

Llamemos al primer tipo de pregunta la *pregunta cuantitativa*; llamemos a la segunda pregunta la *pregunta cualitativa*. (En adelante, para abreviar, hablaremos sólo de *cambio* de propiedades, pero eso incluye al *intercambio* de propiedades.)

Hay varias respuestas posibles para ambos tipos de preguntas y, como en el caso del tiempo y las partes materiales, cada una de ellas puede fundamentarse filosóficamente. Al dar una respuesta, estaremos haciendo *metafísica*: haremos una investigación filosófica sobre los objetos del mundo. Nuestra investigación filosófica será una investigación de metafísica *modal*: al inquirir sobre las *posibilidades* de los objetos, inquirimos sobre los *modos* en los que ellos podrían ser o haber sido. Así pues, la metafísica modal es la rama de la filosofía que investiga sobre las posibilidades de los objetos —la modalidad *de re*.

Dijimos que había varias respuestas posibles. Y las tres respuestas posibles a la pregunta *cuantitativa* son las siguientes: (i) un objeto puede cambiar *todas* sus propiedades y seguir siendo el mismo; (ii) un objeto puede cambiar *algunas pero no todas* sus propiedades si es que ha de retener su identidad; (iii) un objeto no puede cambiar ninguna propiedad: el cambiar una sola implica que es otro objeto. Pronto veremos algunos ejemplos, mientras tanto revisemos las respuestas posibles a la pregunta cualitativa.

Las respuestas posibles a la pregunta *cualitativa* son dos: (a) hay un cierto tipo de propiedades, llamémoslas *esenciales*, tal que si un objeto pierde alguna de sus propiedades *esenciales*, entonces ya no tenemos al mismo objeto; (b) no hay ningún tipo de propiedades tales que, si un objeto pierde alguna (o incluso todas) de ellas, deja de ser el mismo. Daremos ejemplos de estas respuestas un poco más adelante.

Notemos que las respuestas (i)-(iii) y (a), (b), pueden combinarse. Es decir, podemos dar una teoría filosófica acerca del cambio modal de los objetos que combine una respuesta a la pregunta cuantitativa con una a la pregunta cualitativa. Y ahora, antes de ver cómo serían estas combinaciones, debemos notar otra cosa.

El reemplazo del *hardware* de una computadora por *hardware* más nuevo, la restauración de monumentos, son hechos de la vida cotidiana. Pasa todos los días que las cosas cambien, así como pasa todos los días que digamos enunciados del tipo: *Si hubiera pasado que X, entonces...* o *Si este objeto fuera de tal o cual manera, entonces...* o que preguntemos cosas como *¿Qué hubiera pasado si Y?* Tenemos hechos de la vida cotidiana y creencias sobre estos hechos. Creemos, en la vida cotidiana, que una computadora podría tener una tecla diferente y seguir siendo la misma —y si creemos esto es porque creemos que es *un hecho* que una cosa puede o podría cambiar y permanecer ella misma.

Pero no siempre tenemos una teoría bien estructurada sobre este tipo de hechos.

Los filósofos suelen llamar *intuiciones* a este tipo de creencias de la vida cotidiana sobre hechos que podemos investigar filosóficamente. Tenemos la *intuición* de que muchas cosas podrían cambiar y retener su identidad: tenemos la intuición sobre la posibilidad de *esta misma* computadora con una tecla diferente; tenemos la intuición de que es posible que *el mismo* Templo Mayor tenga un par de piedras menos. Lo que los filósofos suelen hacer con estas intuiciones es *sistematizarlas*, tratar de acomodarlas en una teoría que cubra muchos (si es posible, *todos*) los casos del mismo tipo. El proceso de sistematización de las intuiciones en una teoría es un proceso complejo, pero suele incluir –en mayor o menor medida– la construcción de modelos lógico-matemáticos en el marco de los cuales estructurar esta teoría.

Hace unos párrafos vimos que la investigación filosófica sobre los objetos y sus diferencias posibles es llamada *metafísica modal* (o *metafísica de la modalidad*). El marco lógico-matemático dentro del cual se suele trabajar al hacer metafísica modal es la *lógica modal*. Esta lógica es la sistematización del comportamiento formal de las nociones de *posibilidad* y *necesidad*.

El problema central de esta tesis es cómo, bajo la aplicación de la lógica modal, los filósofos intentan dar una respuesta satisfactoria a las dos preguntas que hemos venido persiguiendo, que nombramos las preguntas *cuantitativa* y *cualitativa*: ¿qué tanto puede un objeto cambiar? Y: ¿cuáles podrían ser las propiedades que cambien?

Y el problema con ese problema central es que es un problema enorme. No podría esperar tratarlo a cabalidad en esta tesis. De hecho, investigaré sólo una pequeña sección de él: una respuesta a él conocida como *esencialismo*. Y de las respuestas es justamente de lo que hablaremos inmediatamente.

0.3. Combinando respuestas

Habíamos dejado pendiente la cuestión de cómo podíamos combinar las respuestas a las dos preguntas fundamentales, y prometí dar algunos ejemplos. Recordemos que teníamos las siguientes respuestas: para la pregunta cuantitativa, (i)-(iii); y para la cualitativa (a) y (b), como sigue:

- (i) un objeto puede cambiar *todas* sus propiedades y seguir siendo el mismo;
- (ii) un objeto puede cambiar *algunas pero no todas* sus propiedades si es que ha de retener su identidad;
- (iii) un objeto no puede cambiar *ninguna* propiedad: el cambiar una sola implica que “desaparece del mapa”;
- (a) hay un cierto tipo de propiedades, llamémoslas *esenciales*, tal que si un objeto pierde alguna de sus propiedades *Esenciales*, entonces “desaparece

del mapa”;

(b) no hay ningún tipo de propiedades tales que, si un objeto pierde alguna (o incluso todas) de ellas, deja de ser el mismo.

Ahora veamos un poco más a fondo las cinco respuestas.

Una respuesta del tipo (i) diría, en el caso de mi computadora, que ésta puede cambiar *todas* sus propiedades y aún así seguir siendo la misma. Por ejemplo, puede dejar de tener la propiedad de tener a su teclado actual como parte, puede dejar de tener la propiedad de tener el mismo botón de apagado que tiene ahora, y así sucesivamente. Llamemos “Chompi” a mi computadora. Si Chompi puede dejar de tener *todas* sus propiedades, entonces también puede dejar de tener la propiedad *ser una computadora*. Así que la respuesta (i) implica que mi computadora podría no ser una computadora –y aún así seguir siendo *el mismo objeto*: Chompi. Pero si puede cambiar todas sus propiedades, puede cambiar la propiedad *ser una computadora* por la propiedad *ser un ser humano*. Incluso podría adquirir la propiedad *ser un objeto matemático*. Esto nos puede parecer al menos extravagante, pero es lo que la respuesta (i) permite: si un mismo objeto puede cambiar o intercambiar todas sus propiedades, entonces puede intercambiar cualquier propiedad con otro objeto (que, suponemos, también puede cambiar o intercambiar todas sus propiedades). Así que Chompi y Felipe Calderón pueden intercambiar las propiedades *ser una computadora* y *ser un ser humano*. De hecho, permite algo todavía peor: mi computadora, llamémosla *Chompi*, puede dejar de tener la propiedad *ser Chompi*. Es decir, puede dejar de tener la propiedad *ser idéntica a sí misma*. Esto es contradictorio: es una ley de la lógica que cualquier objeto es idéntico a sí mismo. Así que el tipo de respuesta (i), sin restricciones, es lógicamente inconsistente.

Por suerte, los que responden a la pregunta cuantitativa con el tipo de respuesta (i) pueden librarse de la contradicción, restringiendo su postura pero conservando su espíritu. El que sostiene esta postura puede decir que un objeto puede perder todas sus propiedades, *excepto* por la propiedad *ser idéntico consigo mismo* –y, por supuesto, excepto también por todas aquellas propiedades que sean tales que: el no poseerlas implica *lógicamente* el no poseer la propiedad *ser idéntico consigo mismo*. Las dificultades con esta posición son al menos dos: *primero*, decir cuáles propiedades son del tipo que: el no tenerlas implica el no ser idéntico consigo mismo; *segundo*, explicar cómo la extravagancia que postula (el que mi computadora podría no haber sido una computadora, sino, por ejemplo, un planeta) nos ayuda a explicar el cambio modal sin traernos problemas más graves a bordo. La teoría que defiende una respuesta del tipo (i) se conoce como teoría de la identidad desnuda (*bare identity theory*) o, a veces, *hacceitismo*.³ En esencia, la idea es que toda

3 Del latín “hacceitas”, *esteidad*, la propiedad *ser esta cosa y no otra*, propiedad que es

propiedad que no implique lógicamente la identidad de un objeto x consigo mismo, es una propiedad que x puede o no tener. Esta postura ha sido expuesta y problematizada por gente como David KAPLAN (1975), Graeme FORBES (1994) y Max Black (el hablante B en BLACK 1952); y ha sido defendida por gente como Robert ADAMS (1979) y Penelope MACKIE (2008, especialmente el capítulo 9).

La teoría de las identidades desnudas parece poder combinarse con ambas respuestas a la pregunta cualitativa, siempre que haga ciertas restricciones. Una teoría del tipo (i a) diría que un objeto puede cambiar todas sus propiedades, *excepto* por las propiedades esenciales. Las propiedades esenciales serían todas y sólo aquéllas que sean tales que: el no tenerlas, implique lógicamente la diferencia consigo mismo. Una teoría de tipo (i b) diría que una misma cosa podría cambiar todas sus propiedades, incluso la propiedad de la autoidentidad. Al enfrentarse con la respuesta de que la lógica pura afirma que todo es idéntico consigo mismo, el defensor de la teoría (i b) se pondría a trabajar en dar un nuevo fundamento lógico a su teoría: trabajaría en hacer coherente la idea de que *la identidad es contingente*. Es decir, trataría de hacernos inteligible cómo es que *un mismo* objeto, digamos O , *podría* haber sido *otra* u *otras* cosas. Esta teoría se conoce como la teoría de la *identidad contingente*. (Véase GIBBARD 1975 y YABLO 1987 para teorías de identidad contingente).

La respuesta de tipo (ii), que dice que un objeto puede cambiar *algunas pero no todas* sus propiedades, también tiene el problema de decirnos qué propiedades sí puede cambiar y cuáles no, relativo a cada tipo de objetos. Pero puede dar el comienzo de una respuesta, si se combina con la respuesta (a). Una respuesta del tipo (ii a) dirá que hay cierto tipo de propiedades, las propiedades *esenciales*, tal que un objeto no puede perder ninguna propiedad esencial si es que ha de seguir siendo el mismo. La postura que defiende esta respuesta se conoce como *esencialismo*. (A diferencia de la postura que defiende (i a), el esencialismo defenderá que hay propiedades esenciales tales que: el no poseerlas *no* implica lógicamente el poseer la propiedad de la autoidentidad). Aunque las propiedades esenciales han sido definidas de varias maneras, parece ser que generalmente se las ha tomado como las propiedades más “ligadas” a la identidad (o existencia) del objeto. Podemos dar una definición de esta postura en el marco de la lógica modal —de hecho, una sección de esta tesis se centra en este tema—, lo que queda todavía por responder es cuáles propiedades han de contar como esenciales para qué tipo de objetos.

La respuesta de tipo (ii) también puede combinarse con la respuesta de tipo (b), pero parecería una respuesta al menos extraña. La respuesta (ii b) diría que

justamente la identidad de x consigo mismo. Mi hacciedad es mi propiedad *ser Carlos*. El filósofo Robert ADAMS (1979) fue quien reintrodujo esta terminología, siguiendo a Duns Escoto.

un objeto puede cambiar sólo algunas de sus propiedades, pero no hay ningunas propiedades más “especiales” cuya pérdida implique también la pérdida de la identidad (excepto, claro, por las propiedades tales que el perderlas implique *por pura lógica* la pérdida de la identidad). Es decir, la espuesta (ii b) sería una respuesta *anti-esencialista* que no admitiría que un objeto podría haber sido completamente diferente: aunque sí hay un límite, el límite no es “uniforme”: no hay propiedades esenciales para cada tipo de objetos, sin embargo, estos objetos no podrían ser completamente diferentes.

La respuesta de tipo (iii), que afirma que un objeto no puede cambiar *ninguna* propiedad, suele conocerse como *esencialismo extremo*, *esencialismo máximo* o *superesencialismo*. Parece que la combinación de (iii) con cualquiera de (a) o (b) resulta en la misma teoría. Pues si decimos que un objeto no puede cambiar ninguna propiedad, no importa si las que no puede cambiar sean propiedades de un cierto tipo o no: simplemente no puede cambiar *ninguna*. El problema con el superesencialismo está en explicar por qué tenemos la intuición de que las cosas de hecho pueden cambiar y seguir siendo las mismas.

Varios filósofos han sostenido varios tipos de superesencialismo. Por ejemplo, según muchos intérpretes, Georg Wilhelm von Leibniz sostuvo una teoría superesencialista en su *Discurso de metafísica* (LEIBNIZ 1686). Roderick Chisholm sostuvo el *esencialismo mereológico*, que dice que ningún objeto puede perder ninguna de sus partes sin perder también su identidad (por ejemplo en CHISHOLM 1979, apéndice B; 1989, capítulo 7). Y han habido filósofos que ven, en la filosofía de David Lewis, una teoría superesencialista. Para ver la razón de esto, antes diremos más sobre la lógica modal.

0.4. Lógica modal y teorías de la modalidad *de re*

Como vimos, el marco formal en el que la metafísica modal suele inscribirse es la lógica modal. Aunque las principales motivaciones originales para investigar esta lógica fueron otras, los filósofos pronto tomaron en cuenta el potencial que este formalismo posee para poder inscribir en él una investigación de metafísica modal.

La lógica modal que los filósofos suelen utilizar muestra cómo se comportan lógicamente lo necesario y lo posible, tomando “necesario” y “posible” como nuevos operadores lógicos. La interpretación semántica más común de esta lógica es la llamada *semántica de marcos*. A su vez, al dar una interpretación de ésta semántica, que sea útil para la metafísica modal, los filósofos asumen que estos marcos consisten en *mundos posibles* relacionados entre sí.

La utilidad primaria de postular mundos posibles es ofrecer una imagen más tratable de los conceptos modales. Al decir que algo es *posible*, digamos un hecho p , decimos que p es el caso en *algún mundo posible* (o en algún mundo

posible debidamente relacionado con el nuestro, pero por el momento ignoraremos esta complicación). Al decir que *p es necesario*, eso lo interpretamos como que *p es verdadero en todo mundo posible* (o en todo mundo posible debidamente relacionado con el nuestro). Así que lo necesario y lo posible resultan ser conceptos formalizados, en lógica modal, relativos a mundos posibles.

Ahora bien, también podemos decir que un objeto *x* tiene posiblemente una propiedad (por ejemplo, mi computadora posiblemente tiene un botón diferente de apagado.) Usemos las letras *F, G...* para abreviar propiedades. Así, si decimos que *x es posiblemente F*, eso lo interpretaremos como que *x es F en algún mundo posible*. Si decimos que *x es necesariamente G*, eso lo interpretaremos como que *x es G en todo mundo posible*. (Por supuesto, si un objeto tiene una propiedad necesariamente, entonces también la tiene posiblemente. Pero el caso inverso no siempre se cumple.)

Supongamos entonces que mi computadora posiblemente tiene un botón de apagado diferente. Si decimos que eso se traduce como que mi computadora tiene un botón de apagado diferente en un mundo posible, lo que queremos decir es que mi *misma* computadora tiene, en *otro* mundo posible, un botón de apagado diferente. Y así para cualquier objeto *x* del que prediquemos una posibilidad. Técnicamente, decimos que debe haber una *identidad transmundo* entre ese *x* de otro mundo posible y nuestro *x* actual. El cómo pueda ello ser posible —el cómo pueda ser posible que nuestro *x* actual y *x* en otro mundo posible sean el mismo *x*— es una problema que ha sido discutido por muchos filósofos (por ejemplo KRIPKE 2005, especialmente las páginas 45-50; LEWIS 1986, capítulo 4, sección 4.3; PLANTINGA 1974, capítulo 6; QUINE 1982; VAN INWAGEN 1985). En esta tesis tendré que asumir que la identidad transmundo es el caso.

Hasta aquí podemos reconstruir lo que he venido exponiendo así: (1) tenemos modalidad *de re*; (2) queremos hacer una teoría filosófica sobre este cambio modal; (3) para hacerlo, utilizamos la lógica modal. Agregaré ahora que (4) al hacer un uso filosófico de la lógica modal, nos encontramos con más problemas, además del de la identidad transmundo.

No todos los filósofos estuvieron contentos con la lógica modal estándar y la suposición de que hay identidad transmundo. David Lewis argumentó que, en lugar de postular axiomas, constantes lógicas y cláusulas semánticas extra para la lógica clásica (con lo cual obtenemos la lógica modal), uno puede quedarse con la lógica clásica y construir *sobre ella* una nueva teoría modal. Dicho laxamente, no postulamos una lógica extra, sino que hacemos una teoría que restrinja nuestra lógica clásica, que en principio habla sobre *todo*, para hablar *sólo* de modalidad (LEWIS 1968, 1983a y 1986).

La teoría que Lewis postuló se llama *teoría de las contrapartes*. La diferencia

con la lógica modal es que ésta no necesita que *un mismo* individuo exista en varios mundos posibles. Más bien, dice que un mismo individuo existe *sólo* en un mundo posible. Por ejemplo, mi computadora (“Chompi”) sólo existe en nuestro mundo actual. Pero si podemos decir que *Chompi posiblemente tiene una tecla de encendido diferente*, es porque podemos decir que hay *otro* objeto, en *otro* mundo, que *se parece mucho* a mi computadora y que tiene una tecla de encendido diferente. Decimos que ese otro objeto es la *contraparte* de Chompi. Hablando más generalmente, decimos que *x es posiblemente F* si *x tiene una contraparte en otro mundo y esa contraparte es F en ese mundo*.

Las contrapartes, entonces, son otras cosas *en virtud de las cuales* las cosas actuales tienen *sus* posibilidades. Pero suena extraño decir que el *Templo Mayor* podría haber tenido un par de piedras menos en virtud de que *otro* objeto tiene un par de piedras menos. Esto ha llevado a algunos filósofos a decir que la teoría de las contrapartes no es otra cosa que un superesencialismo. Los filósofos en cuestión dirían algo que parafrasearé así: según la teoría de las contrapartes, uno *no* dice que el *Templo Mayor* mismo en otro mundo tenga un par de piedras menos. Así que, estrictamente hablando, según la teoría de las contrapartes, el *Templo Mayor mismo* no podría haber tenido ninguna piedra menos (ni otra propiedad diferente). Así que llegamos al superesencialismo. (SALMON 1986, 1988 defiende una posición en esta líneas).

Es un debate si esto es así. Pues un teórico de las contrapartes puede replicar lo siguiente (siguiendo a HAZEN 1979 y a FORBES 1984, 2008): No, la teoría de las contrapartes no es un superesencialismo. Pues no dice que un objeto *no pueda* cambiar *ninguna* de sus propiedades. De hecho, acepta –con el sentido común– que muchos objetos pueden cambiar varias de sus propiedades. Sólo que analiza esto de distinta manera: en lugar de decir que *el mismo objeto* en otros mundos posibles tiene diferentes propiedades, dice que otro objeto en otro mundo posible tiene esas propiedades diferentes, y es en virtud de parecerse lo suficiente al objeto actual en cuestión que éste tiene *posiblemente* propiedades diferentes. Aceptamos los mismos hechos. No aceptamos el mismo análisis. Pero los hechos permanecen ahí.

Esto es un debate muy extenso, y por ello no puedo tratarlo en mi tesis.

Hasta aquí, sólo he presentado un panorama muy general de las posturas más defendidas sobre el tema de la modalidad *de re*, para ubicar la importancia y lugar que ocupa la discusión que expongo y argumento en mi tesis: la discusión sobre el esencialismo y la modalidad *de re*. Ahora entraré en materia: diré algo más sobre los problemas y los argumentos que trataré en mi tesis.

0.5. Paradojas

Vimos que tenemos (i) la *intuición* de la existencia de modos posibles en que las

cosas podrían ser (modalidad *de re*); (ii) *sistemas* (o teorías) lógicos para sistematizar filosóficamente esta intuición y (iii) *debates* tanto sobre la teoría apropiada para sistematizar la intuición de la modalidad *de re* como sobre el sistema formal en el que hemos de basarnos. Pero aún si aceptamos el sistema de la lógica modal estándar, los problemas no paran aquí.

Notamos que había tres respuestas posibles a la pregunta cuantitativa: todos, algunos pero no todos, y ninguno: un objeto puede cambiar todas sus propiedades, o algunas pero no todas, o ninguna en absoluto. Identidad desnuda, esencialismo, supersencialismo. Tres respuestas posibles, mutuamente excluyentes y conjuntamente exhaustivas, a una misma pregunta. Además, teníamos dos respuestas posibles a la pregunta cualitativa: propiedades esenciales y cualesquiera. Consideramos las combinaciones posibles de las respuestas a ambas preguntas, calificando a cada respuesta a la pregunta cuantitativa con una respuesta a la pregunta cualitativa. La pregunta ahora es: ¿Por qué aceptar combinación de respuestas cualitativa y cuantitativa, pero no las otras?

Sigueindo a Russell en su famoso "On denoting" (RUSSELL 1905, p. 484), podemos poner a prueba a las teorías por su capacidad de manejar paradojas. En este caso, las teorías de la modalidad *de re* (dadas por las combinaciones de una respuesta cuantitativa y una cualitativa), habrán de enfrentarse a una paradoja que surge al intentar sistematizar teóricamente la intuición sobre la modalidad *de re*, sistematización que incluya *todos* los casos de modalidad *de re* —incluyendo casos *extremos*.

Recordemos un ejemplo. Al reparar mi computadora, me pregunté cuántos cambios podría hacerle antes de hacerla desaparecer, antes de encontrarme con un nuevo objeto. Esta pregunta puede llevar a imaginar casos más extravagantes, como el siguiente. Supongamos que yo podría cambiarle una molécula a mi computadora, y mi computadora permanecer la misma. Y supongamos que si yo cambiara una molécula y mi computadora permaneciera la misma, entonces yo podría cambiar otra molécula más y mi computadora permanecería la misma. Entonces es así para n número de moléculas: si yo pudiera cambiar $n-1$ moléculas de mi computadora y mi computadora permaneciera la misma, entonces yo podría cambiar n moléculas de mi computadora y ésta permanecería la misma. Al final, podría colocar un montón de nuevas moléculas en donde estaban las anteriores, de las cuales ya no permanece ninguna, y la computadora permanecería la misma.

Y esto cuenta como *paradójico*, pues: de principios razonables, llegamos a una conclusión completamente extravagante. El cómo se resuelva la paradoja nos da una de las tres respuestas posibles a la pregunta cuantitativa. Así que cada uno de los tres teóricos toma a la paradoja como una motivación para su postura.

El teórico de la identidad desnuda dice algo como: “Claro, la paradoja nos muestra que uno puede cambiar todas las propiedades, asumiendo lo que asumimos de inicio. Dar un punto de parada para el cambio sería *ad hoc*, capricho filosófico. Así que la teoría de la identidad desnuda es verdadera”.

El esencialista dice algo como: “Claro, la paradoja nos muestra que uno puede cambiar algunas propiedades, pero no todas. Pues si cambiamos todas, tenemos otro objeto, pero sabemos que podemos cambiar algunas. Las que no podemos cambiar se llaman *propiedades esenciales*. Así que el esencialismo es verdadero”.

El superesencialista dice algo como: “Claro, la paradoja nos muestra que uno no puede cambiar ninguna propiedad. Pues si uno cambia alguna, no hay manera de postular un punto de parada en la posible cadena de cambios que no sea *ad hoc*, pero tampoco podemos decir que uno puede cambiar todas las propiedades. Así que el superesencialismo es verdadero”.

0.6. Esta tesis

Las tres respuestas han sido defendidas por filósofos muy importantes, pero yo creo que la respuesta que casa mejor con nuestras creencias cotidianas es el esencialismo. Es por eso que en esta tesis mi preocupación central será el esencialismo.

Temo no poder decir mucho sobre identidad contingente, superesencialismo, teoría de la identidad desnuda y teoría de las contrapartes. Cada uno de éstos temas podría merecer, con toda justicia, una discusión que tomara toda una tesis. Pero yo sólo puedo hacer una tesis a la vez. Así que mi investigación tendrá un carácter *tentativo*: defender que el esencialismo es una postura abierta y plausible; aunque aquí no puedo defender que es la *única* opción plausible, ni la establecida.

Como vimos, el esencialismo está en el centro de una discusión más grande, la discusión sobre la modalidad *de re*. Y esta discusión suele apelar muy fuertemente a la lógica modal. Lo que hemos revisado es suficiente (espero) para dar sentido a lo que me propongo hacer en esta tesis.

Primero, una breve exposición de qué es la lógica modal, junto con los principios que estarán jugando un papel central en el desarrollo de la argumentación. Después expondré con más detalle algunas variaciones de la paradoja que acabamos de revisar, paradoja que ha servido de motivación para el esencialismo. Esto nos llevará a preguntarnos qué es el esencialismo, por lo cual dedico un capítulo a discutir este tema. Mostraré cómo el esencialismo, a menos de convertirse en superesencialismo, también está sujeto al mismo tipo de paradoja que intentaba resolver. Así que tenemos que investigar más profundamente en dónde está la fuente de esta paradoja. En el capítulo 4

expongo un filósofo, Nathan Salmon, que ha argumentado que la fuente del problema está en la base lógica sobre la cual se formaliza la metafísica modal. Salmon rechaza un axioma de la lógica modal, y ese rechazo cierra definitivamente la entrada a la paradoja —si aceptamos que existen, además de mundos posibles, mundos *imposibles*. Concluiré que el rechazo del axioma modal en cuestión nos permite tener un esencialismo coherente y, aún más, plausible. Con esto, tendremos una salida plausible para la paradoja, y el inicio de una explicación filosófica de la modalidad *de re*.

Sin embargo, todavía quedarán abiertas varias soluciones posibles a la paradoja: o aceptar la interpretación estándar de la lógica modal y con ella o hacer teoría de las identidades desnudas o hacer teoría superesencialista; o rechazar la lógica modal estándar y así o hacer teoría de las contrapartes o teoría de la identidad contingente o buscar una explicación cuyo armazón formal esté en alguna lógica de primer orden no-clásica. Todas estas opciones han sido defendidas de maneras ingeniosas. Esto, por supuesto, ha derivado en debates enormes que no podría intentar tratar aquí, y es por eso que me he enfocado en la opción esencialista que se inscribe en la lógica modal clásica.

Será útil hacer un diagrama del panorama que tengo en mente. En negritas están los temas en los que profundizaré en la tesis.



Diagrama 1. Esquema de la tesis.

0.7. Cerrando la introducción

Esta tesis (o una contraparte suya) comenzó como una defensa de una teoría

esencialista en particular: *el esencialismo del origen* (dicho laxamente, esta teoría postula que el origen de los objetos materiales les es esencial a ellos.) Los defensores de esta teoría se han enfrentado con que el argumento más común para establecerla tiene una premisa que nos lleva a un problema del tipo del barco de Teseo. Además, muchos críticos han notado que la argumentación de los esencialistas-del-origen no establece *sólo* su teoría, sino que esa misma argumentación puede utilizarse para establecer muchos otros tipos de esencialismos (algunos de hecho muy descabellados.) Así que el esencialismo del origen sigue siendo problemático, teniendo como justificación más fuerte a la intuición.

Fue justamente al leer sobre este esencialismo que pude notar lo que parece ser la dialéctica más general en la metafísica de la modalidad. Y era obvio: la metafísica de la modalidad se aplica a explicar la modalidad *de re*. Tendría que haber, entonces, discusiones en las respuestas que se ofrecen para explicar este "cambio" a través de las posibilidades, y discusiones sobre cómo formularlas, en términos lógicos. Ya con una teoría en mano, uno vería cuáles son sus costos y beneficios. Éste es, finalmente, el interés central de mi tesis. Si no espero decir algo nuevo, por lo menos espero lograr un panorama útil que sirva de punto de partida para otra investigación —ya sea mía, de alguna contraparte mía, o de algún individuo posible (quizá actual).

(Una nota metodológica: las traducciones que aparecen en este trabajo son casi todas mías. Cuando sea ese el caso, pego el texto original citado a pie de página, llamando a estas notas con dagas: "†").

Capítulo 1. Los principios y la lógica

1.1. Los principios metafísicos

Asumiré que los principios metafísicos incluyen tanto las verdades de la lógica y las matemáticas como las verdades conceptuales. Estos principios incluyen lecturas *de re* de principios lógicos, así como la muy importante intuición de que existe la modalidad *de re*. Los llamo principios *metafísicos* porque, justamente, hablan de los *objetos* y de cómo se comportan *modalmente* (i.e., a través de sus posibilidades).

Los principios relevantes para plantear el problema central de mi tesis son los principios que gobiernan a la identidad y la diferencia, así como los que gobiernan a la posibilidad y a la necesidad. Por brevedad, englobo bajo **LOG=** varias proposiciones, que expresan todas las propiedades lógicas de la identidad, cuál de las tres se use será evidente en el contexto, aunque a veces lo diré de forma explícita. Lo mismo para **NEC=**, la necesidad de la identidad y la diferencia. Asumiré, también, que los nombres propios y las constantes son designadores rígidos, en el sentido de Kripke (2005); además de que los individuos *tienen* las propiedades que tienen, y en el *modo* en el que las tengan, de manera independiente de cómo se los describa; aunque no diré más sobre cuestiones semánticas. Ahora presentaré la lista completa de los principios a utilizar, después de ello haré algunos comentarios que considero pertinentes.

LOG=

Reflexividad: $(\Box \forall x)[\Box(x=x)]$

Simetría: $(\Box \forall x)(\Box \forall y)[\Box(x=y \Rightarrow y=x)]$

Transitividad: $(\Box \forall x)(\Box \forall y)(\Box \forall z)[\Box(x=y \wedge y=z) \Rightarrow x=z]$

NEC=

Necesidad de la identidad: $(\Box \forall x)(\Box \forall y)[\Box(x=y \Rightarrow \Box x=y)]$

Necesidad de la diferencia: $(\Box \forall x)(\Box \forall y)[\Box(x \neq y \Rightarrow \Box x \neq y)]$

LEIB

$(\Box \forall x)(\Box \forall y)(\Box \forall F \in X)[\Box(Fx \Leftrightarrow Fy) \Leftrightarrow x=y]$

MOD DE RE

$(\Box \forall x \in C)[\Box(\Diamond Fx \wedge \Diamond \neg Fx)]$

CONC-->POS

$(\forall p)[\Box(\text{Conceivable}(p) \Rightarrow \Diamond p)]$

Principio heurístico: **REPMUN**

La sintaxis modal utilizada en la argumentación metafísica debe interpretarse con la teoría de modelos basados en marcos, la cual a su vez debe ser interpretada filosóficamente como una teoría que postula un sistema de mundos posibles.

AXIOMAS MODALES

S5-UC = $\{\varphi: \varphi \text{ es un axioma de la lógica clásica de primer orden con identidad}\} \wedge \mathcal{K} \wedge \mathcal{N} \wedge \mathcal{T} \wedge \mathcal{4} \wedge \mathcal{5} \wedge \mathcal{UC}$:

$$\mathcal{K} = (\Box(\varphi \Rightarrow \psi) \Rightarrow (\Box\varphi \Rightarrow \Box\psi))$$

\mathcal{N} = si φ es un teorema de la lógica clásica de primer orden con identidad, $\Box\varphi$ es un teorema de la lógica modal

$$\mathcal{T} = (\Box\varphi \Rightarrow \varphi)$$

$$\mathcal{4} = (\Box\varphi \Rightarrow \Box\Box\varphi)$$

$$\mathcal{5} = (\Diamond\varphi \Rightarrow \Box\Diamond\varphi)$$

$\mathcal{UC} = \{\forall x\varphi: \varphi \text{ es un axioma de la lógica clásica de primer orden con identidad}\}$

Ahora haré comentarios sobre cada uno de los principios. Primero, tenemos dos grandes principios:

La identidad es una relación de equivalencia (LOG=)

Reflexividad: $(\Box \forall x) [\Box (x=x)]$

Simetría: $(\Box \forall x)(\Box \forall y) [\Box (x=y \Rightarrow y=x)]$

Transitividad: $(\Box \forall x)(\Box \forall y)(\Box \forall z) [\Box [(x=y \wedge y=z) \Rightarrow x=z]]$

La identidad y la diferencia, si se sostienen entre x e y , lo hacen necesariamente (NEC=)

Necesidad de la identidad: $(\Box \forall x)(\Box \forall y) \Box (\Diamond x=y \Rightarrow \Box x=y)$

Necesidad de la diferencia: $(\Box \forall x)(\Box \forall y) \Box (\Diamond x \neq y \Rightarrow \Box x \neq y)$

Hay que notar que he modalizado *de re* y *de dicto* cada uno de los principios. Es decir, he colocado operadores modales (de necesidad y de posibilidad) tanto antes de los cuantificadores como dentro del alcance de éstos. La manera en que aparecen los operadores modales expresa que tanto las *fórmulas* como las *propiedades* –que estas fórmulas atribuyen a las variables– son necesarias. Así que estamos hablando tanto de necesidad lingüística como objetual, metafísica.

Me parece que la modalización *de dicto* no es discutible entre quienes aceptan (la enorme mayoría de los teóricos) la lógica clásica de la identidad, pero la

modalidad *de re* sí podría ser más discutible, por cuestiones como referencia o contextos opacos. Es el momento apropiado para remarcar que en esta tesis no discutiré cuestiones de semántica, sino que me limito a suponer que podemos *hablar* de propiedades modales de los objetos (como *ser posiblemente una computadora*), y que estas propiedades son independientes de cómo se los describa. De nuevo, también asumiré que los nombres propios y las constantes son designadores rígidos. La controversia de mi tesis no es la semántica del lenguaje natural para la modalidad *de re*, sino precisamente la modalidad *de re*.

Así que asumiré que cada *objeto* necesariamente se comportan según la lógica clásica de la identidad. Y así para la necesidad de la identidad y la diferencia.

La identidad y la diferencia están necesariamente ligadas a las propiedades, o “Ley de Leibniz” en su versión modalizada (LEIB)

$$(\Box \forall x)(\Box \forall y)\Box[(\forall F : Fx \Leftrightarrow Fy) \Leftrightarrow x = y]$$

Dos notas sobre la Ley de Leibniz: Primero, una de sus “direcciones” ha sido más discutida que la otra (la que va de la derecha a la izquierda): *si* dos objetos comparten todas sus propiedades, *entonces* son idénticos. Esta consecuencia de **LEIB** suele conocerse como *Identidad de los indiscernibles*. Hay quienes afirman, *contra* Identidad de los indiscernibles, que dos objetos pueden compartir *todas* sus propiedades y aún así ser diferentes. Se dice entonces que hay una “diferencia desnuda” entre esos dos objetos, pues su diferencia no está fundada en sus cualidades.

Hemos revisado, de pasada, esta opción en la introducción de esta tesis (sección 1.3). Como dije en la misma introducción, no puedo ser más específico sobre esta teoría. Quedará como una opción a trabajarse en otro momento, frente a la opción que trataré aquí y que sí acepta **LEIB**. Por ahora, sobre este punto, basta hacer notar que **LEIB** establece un vínculo necesario entre las propiedades de los objetos y su identidad.

Segunda nota sobre **LEIB**: no deberíamos incluir *toda* propiedad en el rango de F . Debemos restringir ese rango a propiedades que no impliquen lógicamente a la identidad: por ejemplo, la propiedad *ser idéntico con x* , no puede ser una asignación de F . Pues si lo fuera, **LEIB** se trivializaría: no es informativo decir que dos objetos son idénticos si comparten todas sus propiedades, *incluyendo la identidad*. Quienes nieguen *Identidad de los indiscernibles* (y *a fortiori*, **LEIB**), dirán que el afirmar que *la identidad de los objetos está fundada en sus propiedades* es trivial, si aceptamos que esas propiedades incluyen a la misma identidad.⁴

4 Como en BLACK 1952, p. 155, (en la edición original de *Mind*) en donde B pone esta restricción: “Si quieres tener un principio interesante que defender, debes interpretar ‘propiedad’ más estrechamente –de cualquiera manera, lo suficiente para que ‘identidad’ y ‘diferencia’ no cuenten como propiedades”. [“If you want to have an interesting principle

Para encauzar la discusión hacia el esencialismo, **LEIB** debe restringirse. Así que haré la siguiente restricción para las F relevantes:

LEIB

$$(\Box \forall x)(\Box \forall y)(\Box \forall F \in X) \Box [(Fx \Leftrightarrow Fy) \Leftrightarrow x=y]$$

Donde cualquier propiedad F es miembro de la clase X si y sólo si F no implica a la identidad con un objeto específico.⁵ Es decir: si es posible que Fx , Fy y aún así $x \neq y$, entonces F es un miembro de X .

En el resto de esta tesis, la versión de arriba será la versión de **LEIB** que utilizaremos. Confío en que el contexto hará obvio cuál de las dos "direcciones" de la coimplicación (en caso de que sea sólo una) es la que se está aplicando en cada momento.

El siguiente principio por aceptar es la modalidad *de re*:

La identidad acepta variabilidad en las posibilidades de los individuos (MOD DE RE)

$$\Box \forall x \in C \Box (\Diamond Fx \wedge \Diamond \neg Fx)$$

La idea central de la modalidad *de re* es que *muchos* individuos pueden o podrían haber sido diferentes a como son actualmente pero *permaneciendo ellos mismos* ("Actualmente" debe entenderse en un sentido más amplio que el temporal: en el sentido modal en son *actuales* tanto el pasado que de hecho sucedió, el presente que de hecho sucede y el futuro que de hecho sucederá.) Claro que en esta formulación hay dos cosas por aclarar, que trataré por turnos: ¿Cuáles individuos sí podrían haber sido diferentes, y cuáles no? Y: ¿Qué tan diferentes podrían ser?

Primero ¿cuáles individuos sí pueden o podrían haber sido diferentes, y cuáles no? Por ejemplo, yo podría haber estudiado Derecho y mi computadora podría haber tenido una tecla menos. Pero el número dos no podría haber sido diferente. Esto no es decir que los individuos que podrían ser diferentes son sólo los individuos concretos, pues algunos objetos abstractos podrían haber sido diferentes también: por ejemplo, parece que mi conjunto unitario existe contingentemente: en los momentos y mundos posibles en donde yo no existo, él tampoco existe (véase PLANTINGA 1983, p. 160).⁶ Y un ejemplo de la concebilidad

to defend, you must interpret "property" more narrowly-enough so, at any rate, for "identity" and "difference" not to count as properties"]

5 Nota metodológica. Siempre que diga que "una propiedad F implica *tal y cual*", lo que quiero decir es: para cualquier x , si Fx entonces *tal y cual*. "La propiedad F implica la propiedad G " significa: para cualquier x , si Fx entonces Gx . Siempre que diga que "una propiedad F no implica *tal y cual*", lo que quiero decir es: *no* para cualquier x , si Fx entonces *tal y cual*, y análogamente para la no-implicación de propiedades.

6 Plantinga menciona que esta postura podría depender de la verdad de la postura llamada "actualismo serio", que defiende que un x sólo tiene propiedades en los mundos en los que exista. Pues si el actualismo serio es falso, entonces –por ejemplo– yo podría tener la

de concretos necesariamente existentes: si el cristianismo es cierto y por ello Jesucristo = Dios, y Dios existe necesariamente, entonces Jesucristo (un objeto concreto) existe necesariamente. Otro ejemplo: si el atomismo metafísico es verdadero, entonces parece que podríamos decir que los átomos, entes concretos, existen necesariamente.

Entonces, si es concebible que haya objetos concretos necesariamente existentes y objetos abstractos contingentemente existentes, esto quiere decir que podemos concebir que la categoría de objetos que podrían haber sido diferentes no coincide con la división concreto / abstracto. Y sin esta opción, parece difícil dar una caracterización general de la diferencia categorial entre objetos que podrían haber sido diferentes y los que no. Así que parece difícil dar una respuesta a la pregunta de cuáles individuos podrían ser diferentes y cuáles no. Por el momento, asumiré que *hay* una diferencia exhaustiva y exclusiva entre los objetos que podrían ser diferentes y los que no, sin asumir que haya algo que *explique* o *fundamente* esta diferencia. Ahora hablaré un poco sobre la terminología que usaremos para denotar esta diferencia.

¿Qué objetos hemos de llamar objetos necesarios, o *ens necessarium*? ¿Deberíamos llamar *ens necessarium* a los que necesariamente son como son actualmente, o sólo los que necesariamente existen? Esto determinará, a su vez, cómo hemos de definir a la categoría de los entes contingentes: un contingente debería ser un no-necesario.

Claramente, todo objeto actualmente existente que necesariamente sea como es actualmente, es tal que necesariamente existe. Pero no viceversa: parece que podemos concebir objetos que, aunque existan necesariamente, tienen algunas propiedades contingentes (*i.e.*, son contingentemente como son actualmente). Como dije al hablar del ejemplo de Jesucristo: si el cristianismo es verdadero, entonces Jesucristo = Dios, y dado que Dios necesariamente (o así parece) existe, entonces Jesucristo, un objeto concreto, necesariamente existe. Pero parece intuitivamente correcto decir que Jesucristo podría haber tenido un par de cabellos menos (aunque ¿Podría no haber hecho ver a los ciegos? Parece que no). Así, parece que nuestra intuición modal soporta casos de objetos necesariamente existentes que sí podrían haber cambiado alguna de sus propiedades. Tenemos entonces:

- (i) Que x exista contingentemente *no* implica que: actualmente Fx y posiblemente $\neg Fx$ (pues mi conjunto unitario es contingentemente existente, pero eso no implica que no tenga alguna otra de sus propiedades actuales en otros mundos en los que exista);
- (ii) Que x exista necesariamente no implica que: si actualmente Fx ,

propiedad *ser miembro del conjunto unitario de Carlos* incluso en mundos en los que yo no exista, lo cual resultaría en que mi conjunto unitario no es contingente como yo.

entonces necesariamente Fx (pues Jesucristo actualmente tuvo n cabellos, pero eso no parece implicar que necesariamente tenga n cabellos).

Así que la contingencia o necesidad de la existencia, no parece estar esencialmente ligada a la contingencia o necesidad de otras propiedades (excepto, claro por aquéllas que sean implicadas lógicamente por la existencia). Para motivos de exposición, llamaré *ens necessarium* a los objetos que no podrían haber sido diferentes, y llamaré **C-objetos** a los objetos que sí podrían haber sido diferentes. (Así, incluso un objeto necesariamente existente puede contar como C-objeto).⁷

El segundo punto por aclarar era qué tan diferentes podrían ser los objetos (por definición, este punto es sólo acerca de C-objetos). Aquí también hay dos cosas a aclarar.

Primero, si decimos que un C-objeto x puede o podría haber sido diferente, estamos hablando de lo que podría haber pasado *con ese mismo* x . Es decir: asumo que la modalidad *de re* incluye lo que se suele conocer como *identidad transmundo*: dada la representación estándar de la modalidad en términos de mundos posibles (de la que hablaré pronto), se dice que un mismo objeto existe en varios mundos posibles. La idea es que, por ejemplo, *yo mismo* soy el único objeto denotado por 'Carlos' en 'Carlos podría haber estudiado Derecho': yo mismo soy ese objeto que podría haber estudiado Derecho y que vive en un escenario posible en el cual su estudiar Derecho se realiza.⁸

La otra pregunta es: ¿qué tan diferentes podrían ser (de tal manera que sigan siendo *ellos mismos*)? Ésta también es una pregunta difícil de responder. Yo podría haber tenido un par de cabellos menos en mi cabeza, podría haber estudiado Derecho, pero ¿podría haber sido hijo de los Azcárraga? ¿Podría haber sido mujer? O, más radicalmente: ¿Podría haber sido un tigre? ¿Podría haber sido un objeto abstracto? Es filosóficamente problemático marcar límites no-triviales para las posibilidades de los C-objetos.

En el capítulo 3 de ésta tesis hablaré más sobre los límites no-triviales para las posibilidades de los C-objetos. Por el momento, volveremos a enunciar el principio que intenta capturar la intuición de la existencia de la modalidad *de re*. Diremos que los elementos de C son los C-objetos, y enunciamos así al principio:

7 Se podría argumentar que, en algún sentido, los *ens necessarium* sí pueden cambiar: por ejemplo, hace un minuto no pensaba en el número 17, mientras que ahora lo estoy haciendo. Por lo tanto, el número 17 ha adquirido la propiedad *ser pensado por Carlos*. Intuitivamente, sin embargo, el ser pensado por mí no es una propiedad relevante para los modos en que es el número 17. Parece que lo que se necesita aquí es una restricción a las propiedades que son relevantes para los modos de ser de las cosas, que excluya muchas de las propiedades relacionales.

8 Véase la sección 0.4 de la introducción. Agradezco a Ricardo Murillo y a Cristian Gutiérrez el haberme hecho notar la necesidad de hacer esta aclaración.

La identidad entre objetos es compatible con la variabilidad posible en sus propiedades (MOD DE RE)

$$\Box \forall x \in C \Box (\Diamond Fx \wedge \Diamond \neg Fx)$$

El siguiente principio por aceptar también es discutible. Establece un vínculo entre (a) los resultados de las habilidades cognitivas de los agentes racionales y (b) la posibilidad metafísica (y con esto, también con la necesidad metafísica). Este principio ha sido muy debatido (véase por ejemplo, los ensayos en HAWTHORNE & SZABÓ-GENDLER 2006), pero me apartaría mucho de la preocupación central de mi tesis si lo tratara más a fondo. Por el momento, asumiré que algún vínculo (suficiente, por supuesto, para los fines de mi argumentación) entre la concebibilidad y la posibilidad se obtiene, dados ciertos constreñimientos,⁹ dejando para otra investigación el problema de los constreñimientos que éste vínculo debería cumplir. El principio es el siguiente:

Si una proposición es concebible, entonces es metafísicamente posible (CONC-->POS)

$$\forall p (\text{Conceivable}(p) \Rightarrow \Diamond p)$$

La siguiente y última asunción por hacer no parece ser esencial para mi argumentación: es simplemente la asunción de que el aparato modelo-teórico de la lógica modal, interpretado filosóficamente como un sistema de *mundos posibles*, es la mejor *representación*¹⁰ disponible para la aplicación de la lógica modal alética (*i.e.*, la lógica que gobierna los conceptos *necesidad* y *posibilidad*). No inquiriré demasiado en la naturaleza de los mundos posibles *per se* (aunque una discusión en el último capítulo resultará en que no son el único tipo de “mundos”.) Oficialmente, permanezco neutral en el debate sobre si son

9 Constreñimientos tales como la diferencia que Kripke hace entre una *descripción fijadora de la referencia* y el *contenido semántico de un nombre propio*. Según Kripke, confundir ambas puede llevar a ilusiones modales. Por ejemplo, si confundimos la descripción “La estrella de la mañana”, que *fija* en este mundo la referencia de “Héspero”, con el contenido semántico de “Héspero” (que es Héspero mismo), podríamos crearnos la ilusión de que Héspero podría no haber sido idéntico a Fósforo, pues la descripción que fija la referencia de “Fósforo” es diferente (a saber, “La estrella de la tarde”). Sin embargo, dado que Héspero y Fósforo son idénticos, y dada la necesidad de la identidad, Héspero y Fósforo son necesariamente idénticos —aún si las descripciones asociadas a cada nombre son diferentes. Véase *Conceivability and Possibility* (HAWTHORNE & SZABÓ-GENDLER 2006), Introducción, sección 3, y KRIPKE 2005.

10 El lector notará que estoy hablando de *representación*, no de *análisis*. Que la así llamada “semántica de mundos posibles” no es un análisis (en el sentido de *análisis reductivo*), debería ser obvio cuando notamos el “posibles” en “mundos posibles”. ¿Cómo vamos a analizar (no-circularmente, si es que hay otra manera) esta aparición del (supuesto) *analyzandum* en el (supuesto) *analyzans*? Muchos filósofos han aceptado que la “semántica de mundos posibles” no es un análisis —sino una representación— de la posibilidad, la necesidad, y la manera en que se comportan lógicamente (e.g. SALMON 1989, p. 18 y KRIPKE 2005, prefacio, nota 18). La posibilidad es tomada como primitiva.

escenarios máximos, conjuntos máximos consistentes de proposiciones, etcétera. Sólo pongo la restricción de que la interpretación filosófica asumida debe ser coherente con la argumentación que ofrezco aquí. Digo que esta asunción no es esencial pues, por todo lo que parece, podría argumentar mi tesis sin presuponer una interpretación determinada de la sintaxis modal. Todo lo que requiero es que toda interpretación que pueda haber para el formalismo, sea en términos de mundos posibles o no (y sea lo que sea un mundo posible), sea coherente con mi argumentación. El principio, que es más bien heurístico, es el siguiente:

Principio heurístico: **Representación modal por mundos posibles (REPMUN)**

La sintaxis modal utilizada en la argumentación metafísica debe interpretarse con la teoría de modelos basados en marcos, la cual a su vez debe interpretarse filosóficamente como una teoría que postula un sistema de mundos posibles relacionados entre sí.

Esto completa la enumeración de *casí* todos los principios centrales para mi tesis (dentro de poco veremos los principios lógico-modales). Estos comentarios apenas servirán para contextualizar los principios usados. Cada uno de estos es tan filosóficamente significativo que se han escrito cientos y cientos de páginas sobre ellos. Y aunque son así, temo no poder hacer más que simplemente darlos por hecho: como se sabe, uno debe partir de algún lugar si ha de querer llegar a otro.

Otro principio central será un axioma modal alético, que expondré en el marco de una breve (y no lo más exacta posible) exposición de la lógica modal alética de primer orden. Tratamientos más profundos pueden ser encontrados en HUGHES & CRESSWELL 1996, COCHIARELLA & FREUND 2008, KRIPKE 1963 y FINE 1978.

1.2. La lógica

La lógica a la que estaremos apelando continuamente dentro de la tesis es la lógica modal alética cuantificada **S5-UC**. Ésta lógica tiene el siguiente **lenguaje**:

(i) Un número que suponemos denumerablemente infinito de variables de individuo: x, x_1, x_2, \dots (Por comodidad en la exposición, extraoficialmente podemos hacer uso de x, y, z y u como nuestras variables de individuo.)

(ii) Un número que suponemos denumerablemente infinito de constantes de individuo: c_1, c_2, \dots (En la exposición filosófica, regularmente sólo usaremos nombres propios bien conocidos.)

(iii) Un número que suponemos denumerablemente infinito de letras de predicado, cada una con una aridad fija (*i.e.*, con un número fijo de lugares de

argumento): $F^1_1, F^1_2, F^1_3, \dots, F^2_1, F^2_2, \dots, F^3_1, F^3_2, \dots$ (Por comodidad en la exposición, a veces utilizaremos cualquier letra mayúscula como letra de predicado y dejaremos implícita a la aridad de cada una de ellas, a determinarse por el número de variables individuales que le siguen.)

(iv) Los símbolos auxiliares: paréntesis, corchetes y llaves.

(v) Las constantes lógicas: cuantificadores *universal* y *existencial* (\forall y \exists , respectivamente), *negación* (\neg), *disyunción* (\vee), *conjunción* (\wedge), *implicación* (\Rightarrow), *coimplicación* (\Leftrightarrow), *identidad* ($=$), *necesidad* (\Box) y *posibilidad* (\Diamond).

Las siguientes son las **reglas de formación** que definen recursivamente al conjunto de los términos y al conjunto *FBF* de fórmulas bien formadas. Utilizando τ_1, τ_2, \dots como metavariables para términos, φ y ψ como metavariables que corren sobre *FBF*, tenemos:

(I) Todas las variables de individuo y las constantes de individuo son términos.

(II) Si F^n_m es una letra de predicado con aridad n y τ_1, \dots, τ_n son términos, entonces $F^n_m(\tau_1, \dots, \tau_n)$ es una fórmula bien formada. (Por comodidad, a veces eliminamos los paréntesis.)

(III) $\tau_n = \tau_m$ es una fórmula bien formada, para cualesquiera n, m .

(IV) Si φ es una fórmula bien formada, entonces $\neg\varphi$ es una fórmula bien formada.

(V) Si φ y ψ son ambas fórmulas bien formadas, entonces todas las siguientes también lo son: $(\varphi \vee \psi)$, $(\varphi \wedge \psi)$, $(\varphi \Rightarrow \psi)$, $(\varphi \Leftrightarrow \psi)$.

(VI) Si φ es una fórmula bien formada, entonces ambas $\forall x_n \varphi$ y $\exists x_n \varphi$, también lo son, para todo n .

(VII) Si φ es una fórmula bien formada, entonces ambas $\Box\varphi$ y $\Diamond\varphi$, también lo son.

Los siguientes son los **axiomas** de **S5-UC**:

$$\mathcal{K} = (\Box(\varphi \Rightarrow \psi) \Rightarrow (\Box\varphi \Rightarrow \Box\psi))$$

$$\mathcal{T} = (\Box\varphi \Rightarrow \varphi)$$

$$4 = (\Box\varphi \Rightarrow \Box\Box\varphi)$$

$$\mathcal{B} = (\varphi \Rightarrow \Box\Diamond\varphi)$$

$$5 = (\Diamond\varphi \Rightarrow \Box\Diamond\varphi)$$

$\mathcal{UC} = \{\forall x\varphi: \varphi \text{ es un axioma de la lógica clásica de primer orden con identidad}\}$

Como John Halleck hace notar (HALLECK 2008), **S5-UC** se suele confundir con **S5** (es decir, **S5** sin el axioma \mathcal{UC}). Pero hay una enorme diferencia: mientras que **S5** valida la fórmula Barcan, **S5-UC** no la valida. Y es importante, para los propósitos de mi tesis, el *no* utilizar un sistema que valide ésta fórmula: como Reina Hayaki ha mostrado (HAYAKI 2006), el esencialismo del origen es

inconsistente con la fórmula Barcan. En ésta tesis investigaré, entre otras cosas, a ese esencialismo, por lo que sería fútil investigarlo en el marco de una lógica que de inicio lo falsea. Esto implica aceptar que la fórmula Barcan no es válida, asunción que también es debatible (véase por ejemplo KRIPKE 1963 y MENZEL 2008, sección 2.8), pero que también tendré que hacer para entrar al debate sobre el esencialismo.

Suponemos que **S5-UC** está cerrada bajo la consecuencia lógica de primer orden, más la regla modal de necesitación, \mathcal{N} . Es decir, si φ es un teorema de la lógica de primer orden, entonces también lo es de **S5-UC**, y además:

$\mathcal{N} =$ Si φ es un teorema de la lógica de primer orden, entonces $\Box\varphi$ es un teorema de **S5-UC**.

Ahora expondré brevemente algunas nociones básicas de la **teoría de modelos** para la lógica modal alética de primer orden.

(1) Decimos que \mathcal{F} es un **marco** si $\mathcal{F} = \langle W, R \rangle$, donde W es un conjunto no-vacío y R relaciona a los miembros de W entre sí (i.e.: $W \neq \emptyset$ y $R \subseteq W \times W$). Distinguiremos un único miembro de W , al que denotaremos $@$ (intuitivamente, $@$ es el mundo actual). Interpretado filosóficamente, un marco es un sistema de mundos posibles relacionados entre sí. La relación entre los mundos es la relación de *accesibilidad* o de *posibilidad relativa*. Decir que w *accesa a* w' quiere decir algo como: *dado como son las cosas en w , esas cosas podrían haber sido como son las cosas en w'* . Es decir, la relación de accesibilidad es la relación de *ser-posible-relativo-a*.

(2) Decimos que \mathcal{F}^* es un **marco aumentado** si $\mathcal{F}^* = \langle \mathcal{F}, D, Q \rangle$, donde \mathcal{F} es un marco, D es un conjunto no-vacío y Q es una función que, a cada mundo en W , le asigna un subconjunto de D (i.e.: $D \neq \emptyset$ y $Q: W \rightarrow \wp D$). La interpretación filosófica es la siguiente: D es el conjunto de todos los objetos posibles, y $Q(w)$ es el conjunto de los objetos que existen en el mundo posible w .

(3) Decimos que \mathcal{M} es un **modelo modal** si $\mathcal{M} = \langle \mathcal{F}^*, I \rangle$, donde \mathcal{F}^* es un marco aumentado e I es la función de interpretación, que le asigna a cada mundo posible w la tupla: $\langle Q(w), I_w(F_1), \dots, I_w(c_1), \dots \rangle$, que incluye cada letra de predicado y cada constante, de la siguiente manera: (3a) $Q(w)$ es el conjunto de existentes en w ; (3b) para cada i : $I_w(F_i)$ es un conjunto de n -tuplas de elementos de D (i.e.: $I_w(F_i) \subseteq D^n$); (3c) para cada i : $I_w(c_i)$ es un elemento de D , tal que para cualesquiera mundos w^* y w^{**} , $I_{w^*}(c_i) = I_{w^{**}}(c_i)$. La interpretación filosófica es: un marco se compone de mundos relacionados entre sí por accesibilidad, el dominio de todos los objetos posibles, la asignación de los existentes en cada mundo e I , que le asigna extensiones tanto a las constantes como a los predicados y relaciones en cada mundo. Notemos que la cláusula (3c) implica que las constantes individuales son *designadores rígidos*: como la extensión de

una variable c es la misma en cada mundo posible, c denota rígidamente a su extensión. Así que podemos dejar el subíndice y hablar simplemente de $I(c)$, el objeto denotado por c . Decimos que un modelo \mathcal{M} está basado en el marco \mathcal{F} , si éste es un miembro de \mathcal{M} .

(4) Decimos que s es una **asignación en D** si s es una función que lleva términos (ya sea constantes o variables) del lenguaje a D , de la siguiente manera: (4a) si τ es una variable individual, $s(\tau)$ es un elemento de D ; (4b) si τ es una constante individual c , $s(c) = I(c)$. Intuitivamente, una asignación en D asigna individuos posibles (o actuales) a las variables y deja a las constantes con sus denotados asignados por I .

Sean F^n_m una letra de predicado con aridad n y τ_1, \dots, τ_n términos. Entonces ya podemos definir a la relación \models_s (*verdad en un modelo \mathcal{M} , un mundo w de ese modelo, bajo una asignación s en el dominio D del modelo*) de la siguiente manera:

- (i) $\langle \mathcal{M}, w \rangle \models_s F^n_m(\tau_1 \dots \tau_n)$ sii $\langle s(\tau_1), \dots, s(\tau_n) \rangle$ es elemento de $I_w(F^n_m)$.
- (ii) $\langle \mathcal{M}, w \rangle \models_s (\varphi \wedge \psi)$ sii $\langle \mathcal{M}, w \rangle \models_s \varphi$ y $\langle \mathcal{M}, w \rangle \models_s \psi$.
- (iii) $\langle \mathcal{M}, w \rangle \models_s \neg \varphi$ sii no es el caso que $\langle \mathcal{M}, w \rangle \models_s \varphi$.
- (iv) $\langle \mathcal{M}, w \rangle \models_s \tau_n = \tau_m$ sii $s(\tau_n) = s(\tau_m)$ y ambos son elementos de $Q(w)$.
- (v) $\langle \mathcal{M}, w \rangle \models_s \forall x \varphi$ sii para cada asignación r en D que difiera de s a lo más en lo que asigna a x , tal que $r(x)$ es un existente en w (i.e., $r(x) \in Q(w)$), $\langle \mathcal{M}, w \rangle \models_r \varphi$.
- (vi) $\langle \mathcal{M}, w \rangle \models_s \Box \varphi$ sii para todo w_i tal que wRw_i : $\langle \mathcal{M}, w_i \rangle \models_s \varphi$.

Notemos que he dado sólo las cláusulas para *conjunción, negación, cuantificador universal, identidad y necesidad*. Las cláusulas para las otras conectivas pueden obtenerse directamente aplicando las siguientes definiciones:

$$\begin{aligned} (\varphi \vee \psi) &=_{df} \neg(\neg\varphi \wedge \neg\psi) \\ (\varphi \Rightarrow \psi) &=_{df} (\neg\varphi \vee \psi) \\ (\varphi \Leftrightarrow \psi) &=_{df} [(\varphi \Rightarrow \psi) \wedge (\psi \Rightarrow \varphi)] \\ (\exists x_n \varphi) &=_{df} (\neg \forall x_n \neg \varphi) \\ (\Box \varphi) &=_{df} (\neg \Diamond \neg \varphi) \end{aligned}$$

Lo siguiente es definir es la relación \models (**verdad en un modelo**):

$$\mathcal{M} \models \varphi \quad \text{sii} \quad \text{para todo mundo } w_i \text{ de } \mathcal{M} \text{ y para toda asignación } s \text{ en } D: \\ \langle \mathcal{M}, w_i \rangle \models_s \varphi.$$

Es decir, una fórmula es verdadera en un modelo si y sólo si es verdadera en todo mundo del modelo, bajo cualquier asignación sobre su dominio de objetos. Intuitivamente, una fórmula es verdadera en un modelo si y sólo si no importa

qué objetos le asignemos a las variables de la fórmula (si las tiene) o a sus constantes (si las tiene), esa fórmula sigue siendo verdadera en cualquier mundo.

Decimos que *un marco \mathcal{F} valida una fórmula φ s*ij* $\mathcal{M} \models \varphi$* para todo modelo \mathcal{M} basado en \mathcal{F} . Es decir, intuitivamente: un marco valida una fórmula si y sólo si, sin importar con qué interpretación se suplemente ese marco, la fórmula resulta verdadera (en todo mundo posible).

Ahora bien, Kripke descubrió que varios sistemas axiomáticos modales, populares en la comunidad filosófica, tienen clases de modelos diferentes. Los sistemas analizados por Kripke reflejan sus diferencias en una característica de sus marcos: la relación de accesibilidad. (Recordemos que: $\mathcal{M} = \langle \mathcal{F}^*, I \rangle$, lo cual, desglosado, nos da: $\mathcal{M} = \langle W, R, D, Q, I \rangle$). Imponiendo condiciones sobre la relación R de accesibilidad entre mundos, obtenemos marcos que validan distintos axiomas. Resulta que el axioma \mathcal{K} es validado por todas las clases de marcos, el axioma \mathcal{T} por la clase de los marcos cuya relación es reflexiva, el axioma $\mathcal{4}$ por la clase de los marcos cuya relación es transitiva, el axioma \mathcal{B} por los marcos con relación simétrica, y el axioma $\mathcal{5}$ por los marcos cuya R es una relación de equivalencia (i.e., transitiva, simétrica y reflexiva). Las definiciones son las siguientes:

R es **reflexiva** s*ij* $(\forall x)(xRx)$.

R es **simétrica** s*ij* $(\forall x)(\forall y)(xRy \Rightarrow yRx)$.

R es **transitiva** s*ij* $(\forall x)(\forall y)(\forall z)[(xRy \wedge yRz) \Rightarrow (xRz)]$.

R es **equivalencia** s*ij* R es simétrica, reflexiva y transitiva.

La relación de accesibilidad y los constreñimientos sobre ella que caracterizan cada axioma serán importantes en el desarrollo de la tesis.

Finalmente, me gustaría notar que uno puede *indizar* los miembros de W , los mundos posibles, por los números naturales. (Es decir, podemos establecer una correspondencia (una biyección) entre los elementos de W y un subconjunto inicial de ω : un subconjunto del conjunto $\{0, 1, 2, 3, \dots\}$). Si tenemos n mundos en W , estableceremos una correspondencia entre W y el conjunto $\{0, 1, 2, \dots, n-1\}$. Ésto nos servirá para ponerle "nombres" a los mundos ("mundo w_1 , mundo w_2 ", etcétera), y ello nos servirá para establecer un *orden* en los mundos. Ésto, a su vez, nos servirá para poder hablar de "un mundo anterior a otro" y cosas por el estilo.

El cómo indicemos los mundos es cuestión de utilidad. Podemos decir que éste es el primero y aquél el segundo, etcétera; siempre que todos los mundos queden numerados y que ningunos dos mundos diferentes tengan el mismo índice, ni que dos índices diferentes numeren a un mismo mundo. En esta tesis

supondremos que el orden es dado por la relación de accesibilidad a partir del mundo actual, además de los cambios que estipularemos. Es decir, supondremos que el mundo actual (@) es el mundo “cero” (w_0), y que el primer mundo accesible desde el actual, en donde sucede un mínimo cambio (que estipularemos en contexto) mientras lo demás permanece constante (excepto por lo que el cambio implique lógicamente), es el mundo “uno” (w_1). El mundo “dos” (w_2) es el mundo accesible desde w_1 , en el que sucede un mínimo cambio mientras lo demás permanece constante, y así sucesivamente.

Con esto completamos la enumeración de los “requisitos técnicos” para esta tesis. En el capítulo 2 aplicaremos esta idea de indizar mundos en la exposición de una paradoja en la metafísica modal. En el capítulo 3 veremos cómo se formula una tesis metafísica sobre la base de la lógica modal, mientras que en el último capítulo veremos razones para rechazar un axioma modal.

1.3. Recuento

Para futuras referencias, repetiré sin comentario la lista de principios asumidos.

LOG=

Reflexividad: $(\Box \forall x)[\Box(x=x)]$

Simetría: $(\Box \forall x)(\Box \forall y)[\Box(x=y \Rightarrow y=x)]$

Transitividad: $(\Box \forall x)(\Box \forall y)(\Box \forall z)[\Box(x=y \wedge y=z) \Rightarrow x=z]$

NEC=

Necesidad de la identidad: $(\Box \forall x)(\Box \forall y)\Box(\Diamond x=y \Rightarrow \Box x=y)$

Necesidad de la diferencia: $(\Box \forall x)(\Box \forall y)\Box(\Diamond x \neq y \Rightarrow \Box x \neq y)$

LEIB

$(\Box \forall x)(\Box \forall y)(\Box \forall F \in X)\Box[(Fx \Leftrightarrow Fy) \Leftrightarrow x=y]$

MOD DE RE

$(\Box \forall x \in C)\Box(\Diamond Fx \wedge \Diamond \neg Fx)$

CONC-->POS

$(\forall p)\Box(\text{Conceivable}(p) \Rightarrow \Diamond p)$

Principio heurístico: **REPMUN**

La sintaxis modal utilizada en la argumentación metafísica debe interpretarse con la teoría de modelos basados en marcos, la cual a su vez debe ser interpretada filosóficamente como una teoría que postula un sistema de mundos posibles relacionados entre sí.

AXIOMAS MODALES

S5-UC = La conjunción de los axiomas de la lógica clásica de primer orden con identidad $\wedge \mathcal{K} \wedge \mathcal{N} \wedge \mathcal{T} \wedge 4 \wedge 5 \wedge \mathcal{UC}$:

$$\mathcal{K} = (\Box(\varphi \Rightarrow \psi) \Rightarrow (\Box\varphi \Rightarrow \Box\psi))$$

\mathcal{N} = si φ es un teorema de la lógica clásica de primer orden con identidad, $\Box\varphi$ es un teorema de la lógica modal

$$\mathcal{T} = (\Box\varphi \Rightarrow \varphi)$$

$$4 = (\Box\varphi \Rightarrow \Box\Box\varphi)$$

$$5 = (\Diamond\varphi \Rightarrow \Box\Diamond\varphi)$$

$\mathcal{UC} = \{\forall x\varphi: \varphi \text{ es un axioma de la lógica clásica de primer orden con identidad}\}$

Capítulo 2. La paradoja de indiscernibilidad

2.1. Preliminares

En este capítulo expondré tres paradojas que podemos englobar bajo el nombre genérico de **Paradoja de indiscernibilidad**. Esta tipo de paradojas tiene la siguiente idea de fondo:

Paradoja de indiscernibilidad

Dados los principios **LOG=**, **NEC=**, **LEIB**, **MOD DE RE** y **CONC-->POS** y la lógica **S5-UC**: dos entidades actuales x e y , que son *prima facie* diferentes, pueden llegar a intercambiar *todas* sus propiedades en algún mundo posible w^* . Cuando tratamos de responder a la pregunta sobre si *de hecho* esos x e y de w^* son posibilidades de los x e y actuales, llegamos a una contradicción con alguno o algunos de nuestros principios:

- con **LEIB** (tendremos una diferencia entre x e y que no está fundada en las propiedades de x e y); o
- con **LOG=** (negaremos que la identidad es una relación que tiene la propiedad de la transitividad); o
- con **MOD DE RE** (negaremos que los C -objetos pueden o podrían haber sido diferentes); o
- con **CONC-->POS** (negaremos que el concebir coherentemente que p , implica la posibilidad de p); o
- o con nuestra lógica **S5-UC** (negaremos alguno de sus axiomas).

Como vemos, este tipo de paradojas es un serio problema. Pues parece que los principios en cuestión son considerablemente útiles en la investigación metafísica. Si ellos implican una paradoja, no estamos justificados en usarlos hasta que descubramos la fuente de la inconsistencia: de otra manera, nada nos asegura que el razonar con ellos esté libre de problemas de invalidez o contradicción. Tenemos tres salidas: o encontrar la fuente de la paradoja en algún o algunos principio(s) y entonces rechazarlo(s); o proveernos de un(os) nuevo(s) principio(s) que suplementen a los anteriores e impidan la derivación de la paradoja; o rechazar algunos o todos los principios y apelar a unos nuevos para el razonamiento modal. En esta tesis investigaré la tercera opción.

Por lo pronto, analizaré con más detalle ejemplos de la paradoja en cuestión.

2.2. La Paradoja de Roderick Chisholm

La primer paradoja a analizar fue expuesta por Roderick Chisholm (1967) (una versión parecida también fue expuesta por Hugh Chandler en CHANDLER 1976). El

problema puede ser expuesto preguntándonos si Charles Darwin (ese hombre que escribió *On the origin of species*) pudo haber sido idéntico con Karl Marx (ese hombre que escribió *Das Kapital*).¹¹ Tenemos toda la evidencia a favor para *saber* que Marx de hecho fue diferente de Darwin. Sin embargo, parece que los principios aquí presentados implican que, de hecho, esto no es así. Veamos por qué.

Intuitivamente, Marx y Darwin son del tipo de objetos que caen bajo **MOD DE RE**. Es decir: parece que tanto Darwin como Marx podrían haber sido de maneras diferentes. ¿Qué *tan* diferentes? La intuición y los principios hasta aquí asumidos no nos dicen nada sobre ello, excepto cosas triviales como: un cambio que implique lógicamente la diferencia de Marx (Darwin) consigo mismo es un cambio que le está bloqueado. Es decir: como vimos en la introducción, debemos restringir los cambios posibles de Marx (Darwin) a los cambios que no impliquen lógicamente que Marx (Darwin) pierde su identidad, que Marx (Darwin) es otra cosa. Esto es por pura consistencia lógica: **LOG=** nos dice que todo objeto es idéntico consigo mismo, que si x e y son el mismo z , entonces $x = y$. Así que no podríamos decir que Marx (Darwin) es posiblemente diferente de sí mismo. Pues entonces diríamos que hay mundo en el que un x es idéntico a Marx (Darwin) mismo, pero diferente del Marx (Darwin) actual. Contradicción. Así que los cambios posibles están limitados por la consistencia lógica. Sin embargo, si no tenemos límites (fuera de la consistencia lógica) para “mover” a Darwin y Marx en el “espacio” de los posibles cambios, entonces podemos correr el siguiente argumento.

Primero, supongamos que Marx de hecho midió 1.75 metros exactamente y Darwin 1.80 metros exactamente. Ahora bien, si es cierto que cada uno podría haber sido de diferentes maneras (según **MOD DE RE**), entonces Marx pudo haber medido 1.76 metros y Darwin 1.79. Usando nuestro principio heurístico para representar modalidad con mundos posibles, diremos que existe un mundo w_1 (diferente al actual w_0 , al que denotaremos rígidamente con @) en el que Marx mide 1.76 metros y Darwin 1.79 metros.

(También diremos que en w_1 todos los hechos incompatibles con esta modificación en estaturas son cambiados por hechos compatibles con ellas, tales que sean los más parecidos a los hechos sustituidos: Por ejemplo, si en @ existe el hecho *El hombre que escribió Das Kapital midió 1.75 metros*, en w_1 no existirá ese hecho, sino el siguiente: *El hombre que escribió Das Kapital midió 1.76 metros.*)

Además, y muy importantemente, w_1 es *accesible* desde @: si Marx mide

¹¹ En realidad, Chisholm basa su ejemplo en Adán y Noé. Sin embargo, los escépticos de la Biblia podrían confundir el problema aquí con los problemas de la inexistencia, los nombres de entidades ficticias, términos vacíos, etcétera. Para dejar de lado esto, en mi ejemplo me refiero a entidades cuya existencia y diferencia mutua está fuera de toda duda.

posiblemente 1.76m, entonces el mundo actual podría ser tal que, en él, Marx midiera 1.76m. Y si Marx posiblemente mide 1.76m, entonces en un mundo w_1 Marx mide tanto así. (Y análogamente para Darwin). Así que w_1 es una manera posible en que @ podría haber sido. El afirmar esto está justificado por **CONC-->POS**: pues si en @ es concebible w_1 , entonces para @ es posible haber sido como w_1 .¹²

Pero si ello es posible, entonces hay un mundo posible w_2 , que es posible relativo a w_1 , en el que Marx mide 1.77m y Darwin 1.78m. Así como w_1 es accesible desde w_0 (es decir, al actual @), w_2 será accesible desde w_1 . Y así w_3 (en el que Marx mide 1.78m y Darwin 1.77m) será accesible desde w_2 . Llegaremos a w_5 , en el que Marx mide 1.80m y Darwin 1.75: en w_5 , los mismísimos *Marx* y *Darwin* han intercambiado sus estaturas. Pero aún podemos seguir postulando posibilidades, y decir que en w_6 : Marx escribe una frase menos del *Das Kapital*, pero escribe la primera frase (actual) del *On the origin of species*; y Darwin escribe una frase menos del *On the origin of species*, pero escribe la primera frase del *Das Kapital*. Llegaremos a un mundo de acuerdo al cual: Marx escribirá todo el *On the origin of species* y Darwin todo el *Das Kapital*. Y aún podemos seguir postulando posibilidades, a través de las cuales Marx y Darwin (nuestros *mismísimos* y *bien conocidos Marx* y *Darwin*, esos hombres de carne y hueso que murieron hace muchos años) vayan intercambiando propiedades. ¿Qué lo impide? Pero esto nos obligará a llegar a un mundo en el que Darwin y Marx habrán intercambiado *todas* sus propiedades. Llegaremos así al mundo w^* , el mundo problemático. Pues nuestros principios nos permiten postular lo siguiente:

(i) Una serie de mundos empezando en el mundo @ y terminando en un mundo w_n , una serie de mundos $\langle @, \dots, w_n \rangle$ tal que: de un mundo a otro, Marx y Darwin intercambian una propiedad, y en w_n han intercambiado todas sus propiedades. (Por **CONC-->POS**: podemos concebir que @ sea como w_1 , que w_1 sea como w_2 , ... y que w_{n-1} sea como w_n . Así que cada elemento de $\langle @, \dots, w_n \rangle$ será una posible manera de ser para @, un mundo posible).

(ii) En esa serie, cada uno de los mundos es accesible desde el anterior (únicamente @ no tiene uno anterior): para todos los índices i, j : si $i < j$, entonces $w_i R w_j$ (w_j es *posible relativo* a w_i , o w_i *accesa* a w_j). (Por (i) y la cláusula semántica de la lógica modal para *posibilidad*: si w_j es una manera en que w_i podría haber sido, entonces w_j es posible relativo a w_i ; es decir, w_j es accesible desde w_i).

(iii) La relación de accesibilidad es transitiva: para todos los índices i, j, k : si $w_i R w_j$ y $w_j R w_k$, entonces $w_i R w_k$. (Por el axioma 4 de la lógica **S5-UC**: $\Box\phi \Rightarrow \Box\Box\phi$.)

12 Agradezco a Cristian Gutiérrez el haberme hecho notar esta justificación.

Dados (i), (ii) y (iii), tendremos que cada mundo en la tupla $\langle @, \dots, w_n \rangle$ no sólo es posible relativo al mundo inmediatamente anterior, sino también al primer mundo (el actual, @). Esto implica que w_n es posible relativo a @ —es decir, que las cosas como son en w_n es una manera en que las cosas del mundo actual podrían haber sido. Y como habíamos estipulado que en w_n : Marx tiene todas las propiedades que Darwin tiene en @, y Darwin tiene todas las propiedades que Marx tiene en @, llegamos a concluir que: *es posible que Marx tenga todas las propiedades que Darwin tiene en @, y es posible que Darwin tenga todas las propiedades que Marx tiene en @*. Ésta es la Paradoja de Chisholm, que se ilustra en el diagrama 2.

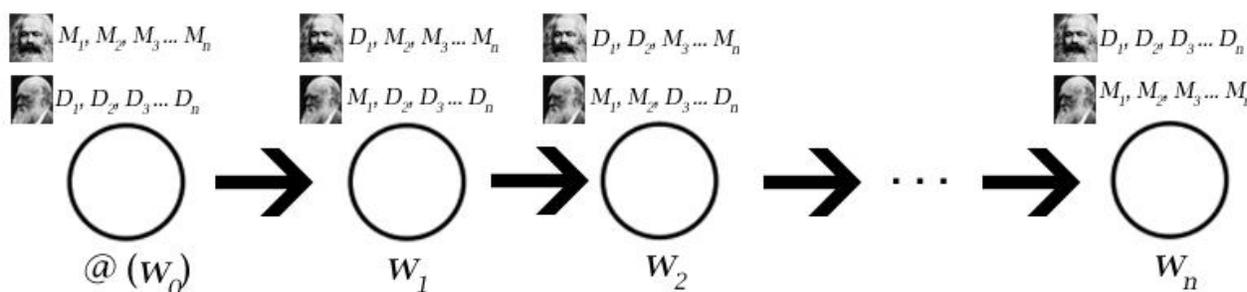


Diagrama 2. La paradoja de Chisholm.

Así que el mundo que verifica la extraña aserción anterior es w_n . (Este mundo es el mundo w^* del que hablamos en los preeliminares de éste capítulo). Es cierto que lo hemos estipulado, pero esa estipulación no fue gratuita. Nos fue permitida por nuestros principios, que incluyen la aceptación de la modalidad *de re*. Iniciamos con pequeños diferencias posibles, siempre bajo los principios, y terminamos con una diferencia posible extrema. ¿Dónde tendríamos que parar de concebir diferencias? Partimos de cierta tolerancia: la identidad de los objetos *tolera* ciertas pequeñas diferencias. Luego dijimos que si esa tolerancia se da, entonces el objeto resultante también tolera una mínima diferencia y permanece siendo ese objeto. Como dice Lewis discutiendo esta paradoja, “pequeñas diferencias suman grandes diferencias”[†] (LEWIS 1986, p. 246). O, como lo frasea Quine: “[...] tú puedes cambiar cualquier cosa por cualquier cosa entre escenarios fáciles mediante alguna serie conectada de mundos posibles”[†] (QUINE 1982 p. 127).

La preguntas que surgen ahora son: ¿ese Marx que está en w_n y que tiene todas las propiedades del Darwin actual, sigue de hecho siendo nuestro Marx actual? ¿Ese Darwin que está en w_n y que tiene todas las propiedades del Marx

[†] “Little differences add up to big differences”.

[†] “[...] you can change anything to anything by easy stages through some connecting series of possible worlds.”

actual, en verdad es nuestro Darwin actual mismo?¹³

Únicamente por motivos de exposición, llamemos a ese Marx que está en w_n y que tiene todas las propiedades del Darwin actual, "Marx-de- w_n ". Análogamente, llamemos "Darwin-de- w_n " a ese Darwin que está en w_n y que tiene todas las propiedades del Marx actual. También llamemos a nuestro Marx actual (ese mismo hombre de carne y hueso que murió hace muchos años y escribió *Das Kapital*), "Marx-de-@"; análogamente, para el Darwin actual reservamos el nombre "Darwin-de-@".

Entonces las preguntas son las siguientes: ¿Es Marx-de- w_n = Marx-de-@? ¿Es Darwin-de- w_n = Darwin-de-@? Claramente, responder afirmativamente a una y negativamente a otra sería *ad hoc*. Así que asumiré que la manera de responder a ambas debe ser uniforme. La pregunta entonces es si responderemos afirmativa o negativamente.

Supongamos que respondemos *afirmativamente*: decimos que Marx-de- w_n = Marx-de-@, y así para Darwin. Esta respuesta implica lo siguiente (la 'A' es por "afirmativamente"):

(A1) Rechazar LEIB. Si Marx-de- w_n tiene todas las propiedades de Darwin-de-@, **LEIB** nos obliga a identificarlos. Mejor fraseado, **LEIB** dice que no podemos suponer que hay dos objetos con todas sus propiedades en común: si x tiene todas las propiedades de y , entonces en realidad x e y son uno y el mismo objeto. Así que al aceptar **LEIB** como principio, aceptamos que Marx-de- w_n = Darwin-de-@ (pues tienen todas las propiedades en común) y que Darwin-de- w_n = Marx-de-@ (pues también tienen todas sus propiedades en común.) Por lo tanto, si respondemos afirmativamente a la pregunta de si es Marx-de- w_n = Marx-de-@, esa respuesta implicará la negación de **LEIB**.

Supongamos que respondemos *negativamente*: decimos que Marx-de- w_n \neq Marx-de-@, y así para Darwin. Esto también nos trae alguno de los siguientes problemas, que por simplicidad ejemplificaremos sólo con Marx, aunque surgen los mismos problemas para Darwin (la "N" es por "negativamente"):

(N1) Negamos MOD DE RE. Pues aceptar que Marx podría ser diferente nos llevó a decir que podría ser de la manera en que w_n lo representa. Así que para decir que de hecho Marx no podría ser como w_n lo representa, diremos que Marx no podría haber sido diferente. Negamos **MOD DE RE** y llegamos al superesencialismo: la postura de acuerdo a la cual un objeto no puede cambiar *ninguna* de sus propiedades.

¹³ No sobra recordar que ésta pregunta no es semántica: no es la pregunta de a cuál objeto, si es que a alguno, vamos a llamar "Marx" o "Darwin". La pregunta es: ¿cuál objeto, si es que alguno, es idéntico a ese objeto que actualmente vivió y murió y que se llamaba "Karl Marx"? Es una pregunta *metafísica*, sobre objetos; no una pregunta sobre cómo los nombramos.

(N2) Negamos **LOG=**. Pues para decir que $\text{Marx-de-}w_n \neq \text{Marx-de-}@$, podemos decir que la relación de identidad no es como la describe la lógica clásica. Esto nos da tres opciones: que la identidad o no sea transitiva, o no sea simétrica, o no sea reflexiva. Si no es transitiva, para algún w_i de la serie $\langle @, \dots w_n \rangle$, nuestro Marx actual (ése que renombramos "Marx-de-@") es Marx-de- w_i , y Marx-de- w_i es Marx-de- w_{i+1} , pero Marx-de- w_{i+1} no es Marx-de-@. Si la identidad no es simétrica, podemos decir que Marx-de-@ es Marx-de- w_i , pero que Marx-de- w_i no es Marx-de-@. Si la identidad no es reflexiva, diremos que Marx no es idéntico a sí mismo en w_i . Así que o cambiamos nuestra lógica clásica o rechazamos que haya identidad transmundo (*i.e.*, identidad de un objeto a través de sus posibilidades)

(N3) Negamos **NEC=**. Pues para decir que $\text{Marx-de-}w_n \neq \text{Marx-de-}@$, podemos decir que Marx es *contingentemente idéntico* a sí mismo: en algunos mundos donde existe es idéntico a sí mismo, en algunos mundos donde *también* existe, no lo es. Por ejemplo, Marx existe en el mundo actual y en el mundo w_n , pero en el mundo w_n Marx no es idéntico consigo mismo. O podemos decir que Marx es *contingentemente diferente* a sí mismo: en w_n , Marx es diferente de sí mismo, pero en el mundo actual Marx es idéntico a sí mismo. Así que cambiamos nuestra lógica modal.

(N4) Negamos **CONC-->POS**. Diremos que alguna de las posibilidades de Marx, aunque concebibles, ya no son posibles. Así que investigamos un vínculo diferente (o con más restricciones que las que tenemos ahora) entre lo que es concebido por los agentes racionales y la posibilidad metafísica.

(N5) Negamos alguno de los axiomas de la lógica **S5-UC** que esté funcionando en la derivación de la conclusión. (Por ejemplo, la transitividad de la accesibilidad). Nos quedamos con un sistema más débil que **S5-UC**.

Hemos llegado al problema: que nuestros principios nos han obligado a aceptar que (en nuestro ejemplo) Marx y Darwin podrían haber compartido todas sus propiedades. Sabemos que ambos eran diferentes actualmente, pero nuestras asunciones no hacen nada cuando intentamos responder sobre los hechos de su identidad en otros mundos posibles. Pues si respondemos algo sobre esas identidades, contradecimos algún principio. Pero si las asunciones parecían obviamente verdaderas, y la conclusión es una contradicción con alguna de ellas, tenemos entonces que hemos llegado a una *paradoja*.

Pero podemos intentar responder –ya sea afirmativa o negativamente, pero responder– a las preguntas ¿Es $\text{Marx-de-}w_n = \text{Marx-de-}@$? ¿Es $\text{Darwin-de-}w_n = \text{Darwin-de-}@$?, y hacerlo sin contradecir algún principio. Podemos intentar la siguiente salida:

(Salida opcional) Suplementamos nuestros principios con uno que impida llegar al mundo problemático w_n . Así que tenemos una *teoría* modal

o una lógica más fuerte que **S5-UC**.

El argumento que revisamos concluye en la existencia del mundo w_n . Esta conclusión es la *Paradoja de Chisholm*, y hemos visto que, para bloquearla, necesitamos o rechazar alguno de nuestros principios o suplementarlos.

La opción que investigaré será la **Salida opcional** (suplementar los principios), combinada con la opción **N5**: negar un axioma de la lógica modal. Pero antes de ver cómo podríamos fundamentar esta doble decisión, será conveniente revisar un par de paradojas de indiscernibilidad más.

2.3. La Paradoja de Mackie

Esta paradoja fue presentada por Penelope Mackie en el segundo capítulo de su libro *How Things Might Have Been* (MACKIE 2007, pp. 25 – 30), con el nombre “El Argumento de Indiscernibilidad”. Mackie inicia la presentación de su Paradoja suponiendo que no hay *esencias individuales*.¹⁴ Hasta el momento, tampoco me he comprometido con este tipo de esencias en mi argumentación, así que podemos aceptar esta asunción.

De nuevo, tenemos la intuición de que Marx y Darwin caen bajo **MOD DE RE**. Además, los principios hasta ahora asumidos no marcan límites para las posibilidades de Marx y Darwin, fuera de la consistencia lógica.¹⁵

Sin embargo, si no hay esencias individuales (de las que hablaré en el capítulo 3 de esta tesis), entonces puede darse el siguiente caso: Marx y Darwin tienen propiedades esenciales, pero ambos comparten esas propiedades. Así que la diferencia actual entre ellos es una diferencia accidental. Entonces, si Marx y Darwin comparten todas sus propiedades esenciales, toda propiedad que Marx podría haber tenido es una propiedad que Darwin podría haber tenido, y viceversa.

Digamos que todas las propiedades que Marx tiene en un mundo posible es una *vida de Marx*, y análogamente para Darwin. Si toda propiedad que Marx podría haber tenido es una propiedad que Darwin podría haber tenido y viceversa, entonces toda vida posible de Marx es una vida posible de Darwin (y viceversa), siempre y cuando las especificaciones de esas historias no impliquen lógicamente la identidad con Marx o Darwin. Podrían hablar directamente de otros individuos, como “*x* es el mejor amigo de Friedrich Engels”, siempre que ello no implique la (no-)identidad de *x* con Darwin o con Marx (cf. MACKIE 2007, capítulo 2, nota 15). Con Mackie, llamemos a estas especificaciones de la vida de

14 Hablaremos más de esto en el siguiente capítulo pero, *grosso modo*, una esencia individual es un conjunto de propiedades que identifica a un único objeto a través de todos los mundos posibles.

15 De hecho, **MOD DE RE** sí marca un límite, pero no uno sustantivo: un *C*-objeto es tal que no podría haber sido un *ens-necessarium*: todo *C*-objeto no podría no poder cambiar.

Darwin o Marx (respectivamente), *especificaciones de vida que son "indiferentes al poseedor"* con respecto a Marx o Darwin, respectivamente.

Si toda vida (que pueda especificarse de manera indiferente a su poseedor) que Darwin podría tener es una vida que Marx podría tener, y dado que una de ellas es la vida actual de Darwin (hasta el punto de la indiferencia de poseedor respecto a Marx), tendremos que la vida actual de Darwin es una vida que Marx podría haber tenido. Entonces podemos postular un mundo $w(m)$ en el que Marx lleve la vida actual de Darwin. En $w(m)$, Marx, por ejemplo, habrá escrito *On the origin of species*, habrá sido inglés, habrá viajado en el *Beagle*, etcétera.

Además, dado que la vida actual de Marx es una vida posible de Darwin, también tendremos un mundo $w(d)$ en el que Darwin viva la vida actual de Marx. En $w(d)$, Darwin habrá escrito *Das Kapital*, será amigo de Engels, será alemán, etcétera.

Finalmente, nada nos impide postular que $w(d) = w(m)$. Es decir, nada nos impide postular que Darwin vive la vida actual de Marx *en el mismo mundo* en el que Marx vive la vida actual de Darwin. Así que tendremos un mundo en el que todo sea completamente igual (hasta el punto de la 'indiferencia al poseedor' respecto al par de Marx y Darwin), excepto que en él Darwin y Marx vivirán, cada uno, la vida del otro. Siguiendo a Mackie, llamemos a este mundo $w(s)$, donde la 's' es por *switching*, intercambio. Ha habido un intercambio entre las propiedades de Marx y Darwin.

El mundo $w(s)$ es el mundo w^* que vimos en los preeliminares a este capítulo: un mundo en el que dos objetos actualmente distintos llegan a compartir todas sus propiedades. Este mundo es paradójico: en él, Marx vive una vida (hasta el punto de la indiferencia a su poseedor respecto a Marx) indiscernible de la vida actual de Darwin, y viceversa. Como dice la misma Penelope Mackie (MACKIE 2007, p. 27):

[...] Estamos imaginando que $w(s)$ es *exactamente* como el mundo actual, excepto por el intercambio de roles entre Sócrates y Platón [en nuestro ejemplo, Marx y Darwin, C.R.]. Pero, *inter alia*, esto parece implicar que todo lo que pasó hasta el venir a la existencia de [Darwin] en el mundo actual —esto es, el curso de eventos que actualmente resultó en la existencia de [Darwin]— no era suficiente para que fuera [*Darwin*] quien entonces vino a la existencia, en lugar de alguien más. *Ex hypothesi*, $w(s)$ contiene exactamente ese curso de eventos, y aún así en $w(s)$ es [Marx], no [Darwin], quien es el individuo producido por los padres actuales de [Darwin], en precisamente el mismo lugar, y en exactamente las mismas circunstancias, en las cuales [Darwin] actualmente vino a la existencia. Pero esto parece cercano a lo increíble. Pues ¿qué *más* podría requerirse para que fuera [*Darwin*] quien vino a la existencia desde ese curso de eventos? Esta

versión del argumento de indiscernibilidad parece implicar que el hecho de que fuera [Darwin] quien emergió desde este curso de eventos es un accidente fantástico, pues podría haber sido también otro quien lo hiciera.[†]

Ahora preguntamos, como con la Paradoja de Chisholm y utilizando el mismo aparato notacional: ¿Es Marx-de- $w(s)$ = Marx-de-@? (Y así para Darwin). Como con la paradoja de Chisholm, dependiendo de cómo respondamos esta pregunta, nos comprometeremos con la negación de uno de nuestros principios (de nuevo, la "A" es por "respuesta afirmativa", la "N" es por "respuesta negativa"):

(A1) Negamos **LEIB**. Si Marx-de- $w(s)$ tiene todas las propiedades de Darwin-de-@, **LEIB** dice que son exactamente el mismo. Pero queríamos decir que Marx-de- $w(s)$ es el mismo Marx actual. Así que negamos **LEIB**.

(N1) Negamos **MOD DE RE**. Pues aceptar que Marx podría ser diferente nos llevó a decir que podría haber vivido la vida actual de Darwin. Así que para evitar decir esto, diremos que Marx no podría haber sido diferente. Negamos **MOD DE RE** y llegamos al superesencialismo (por las razones ya dadas en la sección anterior.)

(N2) Negamos **LOG=**. Pues para decir que Marx-de- $w(s)$ \neq Marx-de-@, podemos decir que la relación de identidad no es como la describe la lógica clásica. Esto nos da tres opciones: que la identidad o no sea transitiva, o no sea simétrica, o no sea reflexiva. Como sólo tenemos dos mundos (@ y $w(s)$), no hay "intermediarios" entre Marx-de-@ y Marx-de- $w(s)$, así que la transitividad de la identidad no es relevante. Si la identidad no es simétrica, podemos decir que Marx-de-@ es Marx-de- $w(s)$, pero que Marx-de- $w(s)$ *no* es Marx-de-@. Si la identidad no es reflexiva, diremos que Marx no es idéntico a sí mismo en $w(s)$. Así que cambiamos nuestra lógica clásica.

(N3) Negamos **NEC=**. Pues para decir que Marx-de- $w(s)$ \neq Marx-de-@, podemos decir que Marx es *contingentemente idéntico* a sí mismo: en algunos mundos donde existe es idéntico a sí mismo, en algunos mundos donde *también* existe, no lo es. Por ejemplo, Marx existe en el mundo actual y en el mundo $w(s)$, pero en el mundo $w(s)$ Marx no es idéntico consigo

† "[...] We are imagining that $w(s)$ is *exactly* like the actual world except for the switch of roles between Socrates and Plato. But, *inter alia*, this appears to imply that everything that happened up to the coming into being of Socrates in the actual world—that is, the course of events that actually resulted in the existence of Socrates—was not sufficient for it to be *Socrates* who then came into existence, rather than someone else. *Ex hypothesi*, $w(s)$ contains exactly that course of events, yet in $w(s)$ it is Plato, not Socrates, who is the individual produced by Socrates' actual parents, at precisely the time and place, and in exactly the same circumstances, in which Socrates actually came into existence. But this seems close to incredible. For what *more* could be required for it to be *Socrates* who came into existence out of that course of events? This version of the indiscernibility argument appears to imply that the fact that it was Socrates who emerged from this course of events is a fantastic accident, since it could just as well have been someone else who did so.[...]"

mismo. O podemos decir que Marx es *contingentemente diferente* a sí mismo: en $w(s)$, Marx es diferente de sí mismo, pero en el mundo actual Marx es idéntico a sí mismo. Así que cambiamos nuestra lógica modal.

(N4) Negamos **CONC-->POS**. Diremos que alguna de las posibilidades de Marx, aunque concebibles, no son metafísicamente posibles. Por ejemplo, diremos que, aunque nuestros principios nos permiten concebir la existencia del mundo $w(s)$, éste no es posible. Así que investigamos un vínculo diferente entre lo que puede ser concebido por los agentes racionales y la posibilidad metafísica.

(N5) Negamos alguno de los axiomas de la lógica **S5-UC** que esté funcionando en la derivación de la conclusión. Nos quedamos con un sistema más débil que **S5-UC**.

(Salida alternativa) Suplementamos nuestros principios con uno que impida llegar al mundo problemático $w(s)$. Así que tenemos una *teoría* modal o una lógica más fuerte que **S5-UC**.

A diferencia de la Paradoja de Chisholm, la Paradoja de Mackie no utiliza un razonamiento "por pasos". La estrategia del razonamiento de Mackie está en notar que, sin límites para las posibilidades, nada nos impide fabricar cierto tipo de colecciones de propiedades y luego "mover" de manera "inmediata" (*i.e.*, sin utilizar un razonamiento "por pasos") a los individuos que podrían ser los poseedores de esas colecciones de propiedades.

Ahora revisaremos una tercera paradoja de indiscernibilidad. Notaremos cómo Kripke propone resolverla, e investigaremos, en el siguiente capítulo, una generalización de su estrategia para responder a las otras dos paradojas.

2.4. El argumento de Kripke para la esencialidad de *ser ficticio*

¿Podría el protagonista de una historia de ficción haber sido un objeto real? Por ejemplo, ¿podría Pegaso, esa figura mítica, haber sido un animal de carne y hueso? Kripke piensa que no. Pero, a primera vista, no parece metafísicamente imposible encontrarnos con evidencia paleontológica que muestre que hubo un animal (ahora extinto) con figura de caballo, que fue el resultado de una larga cadena evolutiva que le brindó alas y la capacidad de volar, etcétera. Digamos que cualquier x es *Pegaso-indiscernible* si x tiene todas las propiedades que la mitología le *atribuyó* a Pegaso.¹⁶

¹⁶ Hasta donde la consistencia lo permita. Si el mito le atribuye a Pegaso propiedades contradictorias, por supuesto, esto eliminaría de entrada la posibilidad de encontrarnos con un pegaso de carne y hueso. También deberíamos dejar de lado, *por mor* de la argumentación, propiedades que el mito le atribuya a Pegaso que sean incoherentes con el ser de carne y hueso. Sin embargo, si de hecho el mito le atribuyó a Pegaso propiedades

Si Kripke desea sostener que no existe ningún objeto concreto que sea Pegaso-indiscernible, debe darnos un argumento. Y Kripke sostiene tal tesis y nos da tal argumento. La tesis puede encontrarse en las *addenda* de *El nombrar y la necesidad* (KRIPKE 2005, pp. 154 – 155):

[...] sostengo la tesis metafísica de que, suponiendo que no hay ningún Sherlock Holmes, uno no puede decir de ninguna persona posible que ella *habría sido* Sherlock Holmes si hubiera existido. [...] Así, ya no podría yo escribir, como una vez lo hice, que “Holmes no existe, pero en otras situaciones objetivas habría existido”.

Sustituyendo el ejemplo de Holmes por el de Pegaso, la tesis parece ser la siguiente: si realmente no ha existido Pegaso,¹⁷ entonces es imposible que exista algún animal¹⁸ x tal que: si x hubiera existido, x hubiera sido Pegaso. Asumiendo que existir es ser idéntico con algo, y ' Ax ' se interpreta como x es animal y 'pegaso' denota a Pegaso, tenemos la siguiente fórmula para representar la tesis de Kripke: $[\neg\exists y(Ay \wedge y = \text{pegaso})] \Rightarrow \neg(\diamond\exists x)\diamond(Ax \wedge x = \text{pegaso})$. Es decir, si realmente ningún animal es Pegaso, entonces en ningún mundo posible existe un objeto x tal que, en algún mundo, x es tanto animal como Pegaso.

El argumento –como el mismo Kripke reconoce– es de una brevedad críptica, pero parece ser el siguiente (KRIPKE 2005, pp. 154 – 155):

Muchas personas posibles distintas, e incluso personas reales, tales como Darwin o Jack el Destripador, pudieron haber realizado las hazañas de Sherlock Holmes; pero no hay ninguna persona de la que podamos decir que *habría sido* Holmes si hubiese realizado estas hazañas. Pues, si así fuera, ¿cuál de entre ellas lo sería?

Utilizando nuestro ejemplo, diríamos:

Muchos animales distintos, e incluso animales reales, tales como Giacomo o Monarchos,¹⁹ pudieron haber vivido la historia de Pegaso; pero no hay ningún animal del que podamos decir que *habría sido* Pegaso si hubiese tenido esas propiedades. Pues, si así fuera, ¿cuál de entre ellos sería?

contradictorias, podemos restringir la definición: x es *Pegaso-indiscernible* si x satisface todas las fórmulas abiertas que están en la intersección de todos los conjuntos X de fórmulas abiertas tales que: si P es una propiedad de Pegaso, entonces Px está en X . Dejo de lado las preguntas de si deberíamos incluir *una* de las propiedades contradictorias (y por ello no su negación), y la de si las propiedades contradictorias podrían contar como esenciales para el Pegaso de ficción.

17 Si x existiera actualmente y fuera Pegaso, diríamos que las historias de la “mitología” en realidad no eran mitología. El caso relevante es el caso de las historias completamente ficticias, de todos modos.

18 Asumo que el ser animal implica ser concreto.

19 Giacomo fue el caballo ganador del Derby de Kentucky en 2005, Monarchos lo ganó en 2001. Véase <http://horseracing.about.com/od/history/l/bl Derbywin.htm> (accesado en Noviembre de 2009).

Este argumento puede reconstruirse, me parece, de la siguiente manera: hay más de un objeto posible que es posiblemente tanto animal como Pegaso-indiscernible (por **CONC-->POS** y **MOD DE RE**).²⁰ Dado que no hay otra cosa en el ser Pegaso que ser Pegaso-indiscernible (esto lo sabemos por **LEIB**), tendríamos que más de un objeto posible es posiblemente tanto animal como Pegaso mismo, es decir, tendríamos un mundo w^* en el que dos objetos comparten todas sus propiedades. (Tendremos, entonces una **Paradoja de indiscernibilidad**). Sin embargo, necesariamente, sólo puede haber *una* sola cosa *idéntica* con Pegaso (esto lo sabemos por **LOG=** y **NEC=**). Así que, a menos de negar alguna de las leyes implicadas en este argumento (**MOD DE RE**, **CONC-->POS**, **LEIB**, **LOG=** o **NEC=**) tendremos que negar que haya objetos posibles que son posiblemente tanto animales como Pegaso-indiscernibles; y así llegamos a la conclusión kripkeana.

El lector notará que parece haber algo tramposo: ¿qué nos justifica el asumir que hay *más de un* objeto posible que es posiblemente Pegaso-indiscernible y animal? Alguien que defienda la existencia del Pegaso concreto defenderá justamente lo contrario: hay *exactamente un* objeto posible que posiblemente es animal y Pegaso-indiscernible: el Pegaso concreto mismo. Esto quiere decir que, si el argumento Kripke no pide la cuestión, tendrá que darse una justificación para esta premisa.

Antes de comentar más, veamos cómo representaríamos este argumento formalmente:²¹

1. $(\diamond \exists x) \diamond (Ax \wedge Px)$	Asunción	para
<i>reductio</i>		
2. $[(\diamond \exists x) \diamond (Ax \wedge Px)] \Rightarrow (\diamond \exists y) (\diamond \exists z) \diamond [y \neq z \wedge Ay \wedge Az \wedge Px \wedge Pz]$	Premisa	
3. $(\diamond \exists x) (\diamond \exists y) \diamond [x \neq y \wedge Ax \wedge Ay \wedge Px \wedge Py]$	De 1, 2, por	<i>Modus Ponens</i>
4. $(\Box \forall x) \Box [(\diamond Px) \Rightarrow (\diamond x = \text{pegaso})]$	Por	definición de Pegaso-indiscernibilidad (Px) y LEIB

²⁰Parece que otro fraseo aquí podría ser: hay muchos animales posibles que son posiblemente Pegaso-indiscernibles. La diferencia con el fraseo del texto principal es la siguiente: esta manera de hablar no impide (a menos que asumamos la esencialidad de *ser animal*) que un animal posible sea Pegaso-indiscernible en un mundo en el que no es animal, lo cual sería un caso que no es relevante para la argumentación: la tesis que Kripke quiere negar es que hay animales posibles que son, posiblemente, tanto animales como Pegaso-indiscernibles.

²¹'Ax' se interpreta como 'x es un animal', 'Px' como 'x es Pegaso-indiscernible', 'pegaso' denota a Pegaso y el cuantificador existencial con '!' es el cuantificador existencial unitario: la fórmulas abiertas que liga tal cuantificador son satisfechas por exactamente un objeto: $(\exists ! x)(Fx)$ se traduce como: $(\exists x)(Fx \wedge (\forall y)(Fy \Rightarrow y = x))$.

5. $(\Box \forall x) \Box ((\Diamond x = \text{pegaso}) \Rightarrow (\Box x = \text{pegaso}))$

Por **NEC=**

6. $(\Box \forall x) \Box ((\Diamond Px) \Rightarrow (\Box x = \text{pegaso}))$

De 4, 5, por S.H.

7. $(\Box \forall x) \Box [\Diamond (Ax \wedge Px) \Rightarrow \Box (x = \text{pegaso})]$

De 6, por el teorema

$\Diamond(\alpha \wedge \beta) \Rightarrow \Diamond(\alpha)$ y S.H.

8. $(\Box \forall x)(\Box \forall y) \Box [\Diamond (Ax \wedge Px) \wedge \Diamond (Ay \wedge Py) \Rightarrow \Box (x = \text{pegaso} = y)]$

De 7, por **NEC=** y

LOG=

9. $(\Diamond \exists x) \Box [\Diamond (Ax \wedge Px) \wedge [(\Box \forall y) \Diamond (Ax \wedge Px) \Rightarrow \Box (x = y)]]$

De 1, 8

10. $(\Box \exists! x) [\Diamond (Ax \wedge Px)]$

De 9, por def. de '∃!'

11. $\neg(\Diamond \exists x) \Diamond (Ax \wedge Px)$

De 1-10, por *Reductio*²²

La línea 11 niega que es posible que exista una cosa que sea posiblemente tanto animal como Pegaso-indiscernible.

Claramente, si las premisas son verdaderas tenemos una prueba de la inexistencia de un Pegaso concreto. Pero la premisa 2 es debatible: ¿por qué la posible existencia de algún objeto que es posiblemente tanto animal como Pegaso-indiscernible implica la posible existencia de al menos dos objetos tales? Ésta es la intuición que está haciendo todo el trabajo en el argumento.

Supongamos que negamos ésta intuición. Llamemos a la propiedad modal *ser-un-posible-animal-pegaso-indiscernible*, *A*. Así, diremos que podemos tener la posible existencia de un *A* sin tener la posible existencia de al menos dos *A*. Sólo es posible la existencia de un único posible animal-Pegaso-indiscernible.

Supongamos entonces que decimos que el único *A* posible es Monarchos. Entonces, necesariamente, ningún otro objeto es posiblemente tanto animal como Pegaso-indiscernible: es decir, será esencial a todos los otros objetos el ser $\neg A$.

Así que Monarchos será el único posible *A*, mientras que todos los otros objetos serán esencialmente $\neg A$. ¿Pero cuál es la base de ésta asimetría? En vez de Monarchos podríamos haber seleccionado a Giacomo. Y en vez de Giacomo, a otro caballo concreto. Pues todos los caballos concretos tienen el mismo "derecho" de haber sido Pegaso-indiscernibles. ¿Por qué sólo un objeto actual o meramente posible y no otro? Si aceptamos ésta asimetría, podríamos pensar en dos objetos *cualitativamente indiscernibles* excepto por la *única* propiedad modal *ser posiblemente Pegaso-indiscernible*. ¿Qué justifica esta asimetría?

Alguien podría defender la siguiente teoría: "Lo que nos obliga a aceptar que un *x* posible sea el *único* que es *A* es, justamente, su identidad: como para todos los demás objetos, sólo hay un único *x* que sea *x* mismo (a saber, *x*). Por ello el *x* que sea *A* es el mismo y único *x*." Pero esta teoría pide la cuestión. El punto es que es intuitivamente correcto sostener que muchos animales podrían haber

²² Mientras que la línea 3 dice que posiblemente existen al menos *dos* objetos posiblemente animales y Pegaso-indiscernibles, la línea 10 dice que *necesariamente* existe *un sólo* objeto posiblemente animal y Pegaso-indiscernible. Contradicción.

vivido la vida de Pegaso, pues claramente muchos podrían haber sido el último eslabón de una cadena evolutiva que resulta en caballos-con-alas, etcétera. La teoría en cuestión *no nos dice* qué tiene de especial el supuesto único Pegaso concreto frente a los demás que también podrían haberlo sido, frente a los que tienen el mismo “derecho” para serlo. La teoría en cuestión simplemente es *ad hoc*.

Por lo tanto, encuentro al rechazo de la premisa 2 del argumento un rechazo infundado. Pero si aceptamos la premisa 2, el argumento corre. Recordémoslo: suponer que posiblemente existe al menos un *A* (un posible animal Pegaso-indiscernible) implica suponer que existe más de uno (a riesgo de poner una restricción *ad hoc*). Pero, por la definición de *Pegaso-indiscernibilidad* y nuestros principios, si posiblemente existe al menos un *A*, entonces éste *A* es único. Pero si nuestros principios nos dicen que la posible existencia de al menos un *A* implica que éste es único, y es *ad hoc* negar que si es posible existencia de al menos un *A* entonces también es posible la posible existencia de más *As*, entonces negamos que posiblemente exista al menos un *A*: negamos que posiblemente exista al menos un posible animal Pegaso-indiscernible.

El argumento de aquí a la conclusión de que Pegaso es esencialmente ficticio es el siguiente. Es imposible que haya un animal concreto que sea Pegaso. Pero claramente hay un Pegaso: el Pegaso ficticio, el de la mitología. Como es imposible que sea un animal concreto, debe ser un Pegaso ficticio. Así, Pegaso es esencialmente ficticio, pues es metafísicamente imposible que sea un animal de carne y hueso.

Argumentos análogos pueden ofrecerse para concluir que Pedro Páramo, Sherlock Holmes, Don Quijote... etcétera, son esencialmente ficticios. No podrían haber sido personas de carne y hueso. Así que podemos generalizar fácilmente la tesis: *Si x es ficticio, entonces x es esencialmente ficticio*. Así que *ser ficticio* es una propiedad esencial de cualquier objeto que la posea. Ésto marca un límite en las posibilidades tanto de los objetos ficticios, como de cualquier objeto que sea posiblemente no-ficticio.

Con este límite, Kripke ha evitado la conclusión de que existe un mundo paradójico w^* : un mundo en el que dos objetos (en este caso, dos caballos concretos) comparten todas sus propiedades, a saber, todas las propiedades de Pegaso. Es decir, so pena de llegar a la **Paradoja de indiscernibilidad**, vemos que Kripke concluye la esencialidad de *ser ficticio*. Kripke toma la **Salida opcional**: no niega ningún principio, sino que postula una nueva tesis y con ello evita la paradoja.

Sin embargo, debemos notar que Kripke sólo bloquea una paradoja: aunque ha evitado la conclusión de que: posiblemente dos objetos *concretos* comparten todas sus propiedades, no ha evitado la conclusión de que: posiblemente dos

objetos *ficticios* comparten todas sus propiedades. Pues dos objetos ficticios compartirán todas sus propiedades esenciales (de hecho, la única: *ser ficticio*), por lo que toda diferencia entre ellos será accidental. Esto permite la entrada a un razonamiento al estilo de Mackie (ver la sección anterior), lo cual a su vez lleva a una **Paradoja de indiscernibilidad**. Por lo tanto, a menos que supongamos que los objetos ficticios tienen esencias individuales, la tesis de Kripke no permite bloquear la paradoja para todos los casos.

¿Cómo, más exactamente, se bloquearía la paradoja? Para ver esto, tenemos que formular más exactamente la tesis esencialista en cuestión. A ello se aboca el siguiente capítulo, al que pasaremos después de un recuento.

2.5. Recuento

Hemos revisado tres paradojas que surgen cuando aceptamos principios sobre la identidad y los cambios de los objetos en un espacio de posibilidades. Vimos que estas paradojas nos ponen ante la necesidad de hacer algo con nuestros principios: o negar alguno, o intentar suplementarlos de manera que la paradoja no surja.

Ahora podemos ver la **Paradoja de indiscernibilidad** desde otro punto de vista. Podemos decir –hablando el lenguaje técnico de la teoría de modelos– que los mundos $@$ y w^* bien podrían pertenecer a dos modelos isomórficos hasta la identidad de los objetos en sus respectivos mundos posibles (en adelante, modelos *isomórficos hasta identidad*).²³

La existencia de estos modelos no es, desde el punto de vista puramente lógico, inaceptable: cumplen –desde el punto de vista puramente lógico– la misma función para la que fueron diseñados: ambos son el aparato suficiente para evaluar las propiedades modelo-teóricas de la teoría modal en cuestión (podemos decir si es consistente, completa, evaluar sus relaciones de consecuencia lógica, etcétera). El problema es un problema de *metafísica*: ¿Es legítimo construir dos modelos isomórficos hasta identidad y usarlos en la evaluación de nuestra teoría de metafísica modal? O, preguntando en la línea de David Kaplan (KAPLAN 1975, p. 122), ¿es la identidad entre los miembros de los dominios de cada w_i (*i.e.*, la identidad transmundana) un *artefacto* del modelo, o es un hecho metafísico que deberíamos imponer en la construcción de los modelos? Es decir: ¿deberíamos decir que puede haber dos modelos M y M^* tales que: si M y M^* son isomorfos hasta identidad, M hace verdadera una

²³La idea intuitiva es que dos modelos son isomorfos *sii* tienen la misma 'estructura' (lo cual implica que tienen el mismo número de objetos). Dos modelos pueden ser isomorfos si tienen la misma 'estructura' y no exactamente las mismas 'características': basta encontrar una función (que será el isomorfismo entre ambos modelos) que 'transforme' las 'características' de uno en las 'características' del otro, preservando la 'estructura'.

oración *O* *sii* M^* también? En este caso, la identidad de los particulares (*i.e.*, los particulares de los dominios de cada mundo) de ambos modelos sería irrelevante. Podríamos tener que en un mundo Marx tiene todas las propiedades que Darwin tiene en el mundo actual (y viceversa), o que Monarchos tiene todas las propiedades del Pegaso ficticio, y ello seguiría haciendo verdaderas a las mismas oraciones. Si no es así, si requerimos que la identidad de los particulares no sea un artefacto de los modelos, podría haber una oración *O* tal que: *O* es (actualmente) verdadera de Darwin; pero evaluada en w^* , *O* es verdadera acerca de Marx.

Decir que la identidad es un artefacto de la representación es aceptar **LEIB**: aceptaríamos que, en la evaluación de las oraciones, lo central está en las propiedades de los individuos: si Marx-de- w^* es cualitativamente idéntico a Darwin-de-@, entonces no hay otra cosa que hacer para que *toda* oración verdadera sobre Darwin sea posiblemente verdadera de Marx, dado que w^* es accesible desde @.

Decir que la identidad no es un artefacto es aceptar identidades desnudas, aceptar la postura *hacceitista*. De nuevo, desde el punto puramente lógico, esto no es problemático (podemos rechazar la “dirección” problemática de **LEIB**: si dos objetos son indiscernibles, entonces son idénticos), el problema está en el nivel *conceptual*.

Por otro lado, también comentamos la estrategia kripkeana. En vez de rechazar algún principio, Kripke acepta una teoría esencialista que bloquee ciertas posibilidades de los objetos, que ponga cierto límite. Este límite impedirá llegar a la paradoja para el caso de los objetos concretos. El límite que Kripke postula es la esencialidad de *ser ficticio*.

Es difícil decir, a partir del pequeño texto fuente de esta exégesis, si Kripke utilizó un razonamiento “por pasos” como Chisholm o una inferencia “inmediata” como lo hace Mackie. El punto importante es el siguiente: si podemos bloquear la paradoja al postular una teoría esencialista, podemos investigar esta opción teórica y ver hasta dónde nos lleva. En el próximo capítulo investigaré esta opción: definiciones, pros y contras, y lo que nos espera en nuestra guerra contra la paradoja.

Capítulo 3. Teorías esencialistas

3.1. Preliminares

Como expuse en el capítulo anterior, los principios lógicos y metafísicos que gobiernan a la identidad y a la modalidad de los objetos, nos llevan a enfrentarnos a tres variantes de lo que llamé **Paradoja de indiscernibilidad**. Esas variantes nos llevan a concluir que existe un mundo w^* en el que dos objetos, actualmente diferentes, intercambian todas sus propiedades.

Revisamos una interpretación de una sugerencia de Kripke sobre los objetos ficticios: según Kripke, ser ficticio es una propiedad esencial de cualquier cosa que la tenga actualmente. También revisamos cómo es que esa propuesta falla: aunque evita la conclusión de que podrían haber varios objetos concretos indiscernibles de un objeto ficticio, *no* evita la conclusión de que podrían haber varios objetos ficticios indiscernibles entre sí.

A pesar de que la respuesta kripkeana falla, me parece que el hecho de que evite un escenario de indiscernibilidad, un mundo w^* , indica que es una respuesta que va por buen camino. En este capítulo investigaré cómo podríamos extender este tipo de opción —la esencialista— a casos más generales que el de los objetos ficticios.

La idea de que una teoría esencialista podría evitarnos caer en la paradoja tiene una base intuitiva muy simple. El esencialismo afirma que las cosas tienen algunas propiedades que no pueden o podrían cambiar, a no ser que dejen de ser ellas mismas. Revisaremos esta intuición inmediatamente, y la formularemos sobre la lógica modal estándar. Luego investigaremos, más exactamente, cómo una teoría esencialista puede bloquear el razonamiento (ya sea al estilo de Chisholm o al de Mackie) hacia la **Paradoja de indiscernibilidad**. Expondré el caso de una teoría esencialista que, además de evitar la paradoja, parece ser la más aceptada por los filósofos: el esencialismo del origen. Finalmente, también expondré los problemas que esa teoría tiene, lo cual nos dará pie al último capítulo de esta tesis.

3.2. Formulando el esencialismo

La intuición esencialista es que hay propiedades que un objeto *debe* retener si es que ha de seguir siendo el mismo objeto. Se suelen reconocer dos tipos de esencialismos: el *general* y el de *individuos*.

El primero postula que una cierta propiedad (o conjunto de propiedades) es *necesaria* para la existencia de un individuo, sin prejuzgar si esta propiedad (conjunto de propiedades) también es suficiente para la existencia de ese

individuo en particular. De hecho, esta propiedad (conjunto de propiedades) también puede ser esencial para otros objetos. Por ejemplo, Kripke argumenta que *ser ficticio* es esencial para cualquiera de sus poseedores. Así que tanto Pegaso como Sherlock Holmes y Pedro Páramo son esencialmente ficticios — aunque, claramente, los tres son objetos distintos, y justamente por ésto es que *no es suficiente* el que un x sea ficticio para que x sea idéntico con el mismo Pegaso, o con el mismo Sherlock Holmes o el mismo Pedro Páramo.

El segundo esencialismo postula *esencias individuales*. Este esencialismo afirma que una cierta propiedad (o conjunto de propiedades) es necesaria y *suficiente* para la existencia de un individuo en particular. Por ejemplo, la propiedad *ser Carlos Romero* es necesaria y suficiente para mi existencia: si un x en cualquier mundo ejemplifica esa propiedad, eso es suficiente para que ese x sea yo mismo; además, en cualquier mundo, ningún x puede ser yo mismo si no ejemplifica esa propiedad. Esto resulta en que una esencia individual individúa a un único x a través de todos los mundos posibles: cualquier objeto y en cualquier mundo posible w es x , si y sólo si y ejemplifica en w la esencia individual de x , por lo que es imposible que una esencia individual sea compartida por varios objetos.

Ahora bien, se suele hacer la distinción (tanto para esencias individuales como para esencias generales) entre propiedades esenciales *triviales* y propiedades esenciales *sustantivas* (por ejemplo, MACKIE 2007, sección 2.2; ROBERTSON 2007, sección 3). Las propiedades esenciales triviales serían aquéllas que *cualquier* objeto necesariamente tiene, en virtud de las verdades de la lógica, las matemáticas y las verdades conceptuales. Por ejemplo, *ser idéntico consigo mismo* es una propiedad que cualquier objeto tiene en virtud de la lógica; *ser tal que $5+7=12$* también es una propiedad que cualquier objeto tiene en virtud de las matemáticas; *ser hombre no-casado si se es soltero* también es una propiedad que tienen todos los objetos, en virtud de las verdades conceptuales. Pero *tener el número atómico 79* no es una propiedad que podamos conocer *a priori* por la lógica, las matemáticas o el análisis conceptual, sin embargo, Kripke sugirió (KRIPKE 2005, tercera conferencia) que es una propiedad esencial del oro. Así que la propiedad *tener el número atómico 79* es una propiedad esencial general sustantiva: cualquier muestra de oro necesariamente la tiene si es que existe, pero esto no en virtud de la lógica o las matemáticas o las verdades conceptuales. Y así para las esencias individuales: trivialmente, es necesario y suficiente para cualquier objeto x en cualquier mundo w , el que tenga la propiedad *ser idéntico con Carlos Romero* para que x sea yo mismo en w : esto lo sabemos por pura lógica. Pero no sería trivial decir, por ejemplo, que cualquier x en cualquier w soy yo mismo, si y sólo si x en w tiene el mismo origen que yo tengo actualmente. Así que el origen, si es una esencia individual, es una esencia

individual *sustantiva*.

Lo que ahora debemos hacer es restringir de inicio los esencialismos (general e individual) a propiedades esenciales *sustantivas*, no-triviales. Además, de nuevo debemos recordar que he estado asumiendo que (i) las constantes son designadores rígidos, y que (ii) los objetos tienen las propiedades que tienen independientemente de cómo se los describa.

Los dos tipos de esencialismos suelen formularse en términos de la lógica modal. Para el esencialismo general, la idea que debemos capturar es que hay ciertas propiedades (no-triviales) que *necesariamente* son necesarias para la existencia de los objetos a los cuales son esenciales:

(Propiedad esencial general)

$(\Box \forall x)(\forall F) F$ es una propiedad esencial de $x =_{df} \Box (Existe(x) \Rightarrow Fx)$

Siguiendo la distinción estándar, una propiedad que no sea esencial es una propiedad *accidental*:

(Propiedad accidental)

$(\Box \forall x)(\forall F) F$ es una propiedad accidental de $x =_{df} \Diamond (Existe(x) \wedge \neg Fx)$

Para el esencialismo individual, la idea es que hay ciertas propiedades que son necesarias y suficientes para la existencia de un individuo en particular. (Dejaremos de lado la complicación que supone notar que las esencias individuales podrían ser *conjuntos* de propiedades.) Si notamos que (como se sugiere en SALMON 1987) la existencia puede definirse como la propiedad de ser idéntico con algo (*i.e.*, $\lambda x[(\exists y)x=y]$), entonces podemos dar la siguiente definición de esencia individual:

(Esencia individual)

$(\Box \forall x)(\Box \forall y)(\forall F) F$ es la esencia individual de $x =_{df} \Box (x=y \Leftrightarrow \Diamond Fy)$

Adviértase que en ninguna de las definiciones hemos restringido los cuantificadores a los *C*-objetos (los objetos contingentes.) Las definiciones, entonces, también se aplican a los *ens necessarium*, e implican que los *ens necessarium* tienen todas sus propiedades esencialmente.²⁴

Ahora nos podemos preguntar si una teoría esencialista nos ayudaría a bloquear la **Paradoja de indiscernibilidad**. Hemos visto (en el capítulo anterior) cómo la propuesta de Kripke bloquea la paradoja en algunos casos, pero falla en bloquearla para otros casos. La pregunta ahora es si podemos bloquear la **Paradoja de indiscernibilidad**, ya sea al estilo de Kripke, ya sea al estilo de Mackie, utilizando una teoría esencialista.

²⁴Véase ROBERTSON 2007, sección 4 y MACKIE 2008, capítulo 1, sobre el esencialismo máximo (la postura, como la de Leibniz, que acepta que los objetos tienen *todas* sus propiedades esencialmente). Lo que se sigue de lo que estoy diciendo es solamente un esencialismo máximo para los *ens necessarium*, así que aquí no debería surgir ninguna perplejidad.

3.3. ¿Podemos evitar la Paradoja de indiscernibilidad apelando a una teoría esencialista?

La respuesta a la pregunta del título de esta sección tiene un aire decepcionante. Llanamente, no podemos. Para mostrarlo, analizaremos una teoría esencialista tal que: (a) bloquea la argumentación a la Paradoja de indiscernibilidad; (b) es la teoría esencialista más plausible (o, al menos, la teoría esencialista *disponible* más plausible.)

Ésta teoría tiene dos componentes, cada uno de los cuales hace un trabajo para brindarle las propiedades (a) y (b). La característica (a) de la teoría está dada por un principio esencialista que sirve para derivar la existencia de esencias individuales, las que, a su vez, bloquean la paradoja. La característica (b) está dada por la intuitividad de una versión inexacta del principio esencialista en cuestión. Al hacer una formulación exacta de este principio, necesitamos postular una tesis que llamaré *principio de tolerancia* y que definiré más abajo. Intuitivamente, la idea de este principio es que las esencias de los objetos “toleran” leves cambios; este principio es aceptado para no llegar al superesencialismo. Así, la plausibilidad de la tesis esencialista está importantemente fundada sobre la verdad del principio de tolerancia.

El problema con el principio de tolerancia es que abre la puerta, de nuevo, a la **Paradoja de indiscernibilidad**. Por ello, a final de cuentas, la teoría esencialista en cuestión es tal que: (i) necesita del principio de tolerancia para ser plausible; (ii) la adopción de este principio abre la puerta a la **Paradoja de indiscernibilidad**. Al final de las cuentas, entonces, la teoría esencialista por sí sola no bloquea la paradoja —y a mostrar ésto me dedicaré ahora.

¿Cómo una teoría esencialista bloquearía la paradoja? Para que la teoría en cuestión evite la paradoja, es necesario y suficiente que postule la existencia de propiedades que cumplan con lo que podemos llamar el *requisito de doble unicidad*. Cualquier teoría que postule propiedades que lo satisfagan bloquea la paradoja, pues las propiedades que satisfacen este requisito nunca son intercambiadas. El requisito es el siguiente (notación: “ $F_w x$ ” se interpreta como “ x es F en w ”):

Requisito de doble unicidad (RDU)

El conjunto A de propiedades satisface el Requisito de doble unicidad =_{df}

$(\forall F \in A)(\forall w)$ Si $F_w x$, entonces:

(i) $[\neg(\exists G \in A)(G \neq F \wedge G_w x)]$, equivalentemente: $[(\forall G \in A)(G_w x \Rightarrow G = F)]$, y

(ii) $[\neg(\exists y \neq x)(F_w y)]$, equivalentemente: $[(\forall y)(F_w y \Rightarrow y = x)]$

Es decir, los conjuntos de propiedades A que satisfagan **RDU** son tales que, en cada mundo, máximo *un* objeto ejemplifica máximo *una* A -propiedad. Ahora

veamos cómo las propiedades que satisfagan **RDU** (si existen) bloquean la **Paradoja de indiscernibilidad**, cuando ésta se alcanza mediante un razonamiento chisholmiano.

Dado que el argumento que nos lleva a la paradoja necesita que la serie $\langle w_1 \dots w_n \rangle$ sea tal que: en w_n los objetos involucrados hayan intercambiado todas sus propiedades, si encontramos un mundo w_m (para $m < n$) en el que este intercambio ya no se dé, entonces habremos bloqueado la inferencia y evitaremos, así, el problema. Así que sea $\langle w_1 \dots w_n \rangle$ una cadena de mundos como la del argumento que nos lleva a la paradoja, y sean M y D propiedades que cumplen con **RDU**.

Supongamos que la teoría esencialista en cuestión postula que M es una propiedad esencial de Marx, y D es una propiedad esencial de Darwin. Esto es decir que: en todo mundo w en el que Marx exista, Marx será M en w (y así para Darwin y D). Por **RDU**, en todo mundo w en el que Marx exista, el hecho de que Marx sea M en w implica que no tendrá otra propiedad de la clase A en w —en particular, Marx no tendrá la propiedad D , pues D también está en A . Además, el hecho de que en w Marx sea M , implica que ninguna otra cosa en w es M . Y así para Darwin e D .

Pero como en todo mundo Marx es M , y en cada mundo donde Marx es M Marx no es D , tenemos que en ningún mundo Marx existe y es D . Análogamente, Darwin no es M en ningún mundo. Como la cadena $\langle w_1 \dots w_n \rangle$ necesita que todas las propiedades sean intercambiadas por Marx y Darwin, y hemos mostrado que las propiedades M e D nunca son intercambiadas, se sigue que nunca llegaremos al mundo w_n , en el que todas las propiedades actuales de Marx sean propiedades de Darwin y todas las propiedades actuales de Darwin sean propiedades de Marx. Ahora nada nos obliga a identificar a Marx con Darwin, y bloqueamos así la inferencia a la paradoja de Chisholm.

Por supuesto, hemos bloqueado la paradoja de Chisholm, pero ésta no era la única forma en que surge la **Paradoja de indiscernibilidad**. Todavía tenemos que bloquear la paradoja de Mackie. Hay que notar que la paradoja de Mackie *no* necesita de una serie de mundos $\langle w_1 \dots w_n \rangle$. Simplemente, necesita asumir que toda posibilidad de un x es posibilidad de otro y , y viceversa. Pero para esto era necesario suponer que no hay esencias individuales: si las hubiera, entonces no toda posibilidad de uno es posibilidad de otro. Pues supongamos que X es la esencia individual de x , y la esencia individual de y es Y . Entonces la posibilidad de x de ser X es una posibilidad que le está bloqueada a esa otra cosa y ; y la posibilidad de y de ser Y le está bloqueada a x . Por lo tanto: si hay esencias individuales, entonces no toda posibilidad de uno es una posibilidad de otro. Por lo tanto, si existen propiedades que satisfagan **RDU**, entonces los objetos que ejemplifiquen esas propiedades no estarán sujetos a la paradoja de

Mackie.

Lo que debe quedar claro es que la segunda condición de **RDU** es un requisito de suficiencia que, cuando se le conjunta con la postulación de que las *A*-propiedades son esenciales, nos brinda la postulación de esencias individuales. Pues como en cada mundo *w* donde exista Marx, éste será *M* (por la postulación de que *M* es esencial a Marx), y ninguna otra cosa es *M* en *w*, tendremos que en cada *w* donde exista, sólo Marx será *M* en *w*. Ésto, por definición, nos da que *M* es la esencia individual de Marx: cualquier *x* es Marx si y sólo si *x* es *M*.

Por todo esto, si encontramos una teoría esencialista satisfactoria, cuyas propiedades esenciales postuladas cumplan con **RDU**, la victoria es nuestra. ¿Qué propiedades podríamos postular que fueran plausiblemente esenciales y que cumplieran con **RDU**?

3.3.1. Nuestra mejor teoría esencialista...

Intuitivamente, las propiedades del origen de un objeto cumplen con **RDU**. Pues,, intuitivamente, en cualquier mundo en donde yo exista, sólo yo tengo mi propio origen: en cada mundo, nadie más nace de mis mismos gametos. (Esto es una condición intramundo, *de dicto*: no prejuzga que yo tenga diferentes gametos en diferentes mundos). Así que mi origen cumple con la condición (ii). Pero también con la (i): pues, intuitivamente, si me origino de *X* en *w* (por ejemplo, de la fusión de ciertos gametos), entonces no me origino de ninguna otra fusión de gametos *Y* en *w*: uno no puede ser hijo de los Fox y de los Azcárraga en un mismo mundo.

Así que basta con postular que el origen es esencial a Marx y a Darwin, además de aceptar que el origen cumple con **RDU**, para obtener la tesis de que el origen de cada objeto es su esencia individual (o un constituyente necesario de ésta), y así bloquear la paradoja. Basta con ser, entonces, *esencialistas del origen* para bloquear la paradoja. Kripke, Colin McGinn, John L. Mackie, Nathan Salmon, Graeme Forbes y el equipo de Louis deRosset y Guy Rohrbaugh son ejemplos de filósofos que han defendido el origen-esencialismo (**OE**). (KRIPKE 2005, pp. 109-114, especialmente la nota al pie 56; SALMON 1979, 2005b, especialmente el capítulo 7; FORBES 1985 capítulo 3, 1986, 2001; ROHRBAUGH & DEROSSET 2004, 2006; MCGINN 1976; MACKIE 1974).

Kripke dió ejemplos intuitivos y "algo como una prueba" (como él mismo dice en *op. Cit.*, nota 56) para mostrar que el origen-esencialismo para los artefactos y seres biológicos es muy plausible; Mackie, McGinn y Forbes lo defendieron para los entes biológicos; Salmon y deRosset-Rohrbaugh para artefactos. Muchos de éstos filósofos han afirmado que la tesis parece poder generalizarse a casos más generales. Robert C. Coburn nos expone esta idea y nos da algunas definiciones que nos serán útiles (COBURN 1986, p. 170):

Probablemente todas las cosas persistentes emergen de una manera que es análoga o a la construcción intencional o al desarrollo biológico. Si es así, entonces todas las cosas persistentes tienen un “origen”, donde x es el origen de $y =_{df}$ (a) x es un conjunto de uno o más objetos, individual y conjuntamente distintos de y , de los cuales y se desarrolló *vía* algún proceso biológico natural, o algún proceso natural análogo a un proceso de desarrollo biológico; o (b) x es algún pedazo de materia que es distinto de y , o una colección de uno o más objetos que son individual y conjuntamente distintos de y , desde los cuales y fue intencionalmente hecho de acuerdo a algún plan, o de los cuales y vino a ser *vía* algún proceso o azar no-intencional. En cualquier caso, muchas de las cosas en nuestro alrededor con las que tenemos un comercio diario tienen orígenes en el sentido indicado.[†]

La idea, entonces, es que muchos objetos parecen estar sujetos a la esencialidad de su origen, aunque todavía nos falta saber exactamente qué tipo de objetos. Graeme Forbes (1985, capítulo 3), menciona que los miembros de un conjunto podrían ser los “originadores” del conjunto en cuestión; ROHRBAUGH & DE ROSSET (2004, p. 710, nota 10) mencionan que su argumentación podría cubrir casos como las sinfonías). Dejemos de lado la pregunta sobre el alcance de la tesis, y pasemos a intentar formularla de manera más exacta.

Podemos primero notar las intuiciones que están detrás. El *locus classicus* está en en Kripke (*op. Cit*, p. 112-113), donde leemos:

¿Cómo podría ser *esta mismísima mujer* [la Reina de Inglaterra, Isabel II] una persona que se hubiese originado a partir de otros progenitores, de un espermatozoide y un óvulo completamente diferentes? Podemos imaginar, *dada* esta mujer, que varias cosas en su vida hubiesen cambiado: que se hubiese convertido en una mendiga, que su sangre real hubiese permanecido ignorada, y así sucesivamente. [...] Pero lo que es más difícil imaginar es que hubiese nacido de padres diferentes. Me parece que cualquier cosa proveniente de un origen diferente no sería ese objeto.

En el caso de esta mesa [...], podemos no saber de qué pedazo de madera

† “Probably all persisting things emerge in a way that is analogous to either intentional construction or biological development. If so, then all persisting things have an “origin”, were x is the origin of $y =_{df}$ (a) x is a set of one or more objects, severally and jointly distinct from y , from which y developed via some natural biological process or some natural process analogous to a process of biological development, or (b) x is some hunk of stuff that is distinct from y , or a collection of one or more objects that are severally and jointly distinct from y , out of which y was made intentionally and according to some plan, or from which y came to be via some nonintentional process or happenstance. In any case, most of the things in our surroundings with which we have daily commerce have origins in the sense indicated. [...]”

procede la mesa. Ahora bien, ¿podría haber sido hecha *esta mesa* de un pedazo de madera completamente *diferente* o, incluso, de agua hábilmente endurecida hasta hacerla hielo —de agua tomada del río Támesis?

Advirtamos que, intuitivamente, la historia de los *C*-objetos es contingente. Es decir, posiblemente un *C*-objeto dado *x* tiene una historia diferente. Es decir: hay mundos posibles en los que *x* es diferente a como es actualmente. Pero (por **MOD DE RE**), ése *x* en otros mundos debe ser nuestro mismo actual *x*, (y es por esto que Kripke dice: “Podemos imaginar, *dada* esta mujer, que varias cosas en su vida hubiesen cambiado”). Pero si la historia de los *C*-objetos es contingente, entonces, si es que hemos de buscar su esencialidad en algo, esa esencialidad debe estar en su venir a la existencia —en su *origen*.

No es trivial ver qué propiedades, objetos u eventos, presentes en o constituyentes de mi origen, estarían en una relación esencial conmigo. Pero parece haber una intuición compartida: que la materia de la que actualmente se viene a la existencia, configurada como actualmente estuvo configurada, es lo que es esencial del origen. El espermatozoide y el óvulo de la Reina, la materia de la mesa, parecen ser las cosas relevantes para la esencialidad. Pero como con la materia de la mesa podríamos haber hecho una silla, y como el espermatozoide y el óvulo podrían haberse modificado por ingeniería biológica para resultar en otro objeto, también la configuración de la materia (o el plan con que se diseña, en el caso de los artefactos) parece esencial.

Generalicemos estos casos. Siguiendo a Forbes (*op. Cit.*), podríamos hablar de “propágulas”, como los originadores de los organismos. Más generalmente, intentando cubrir el caso de los organismos, de los artefactos y de todos los otros tipos de objetos que caigan bajo el esencialismo del origen, podemos hablar de “originadores” e incluir en este término tanto la materia original como la configuración de ésta.

La tesis de que *muchos* objetos tienen esencialmente a su origen parece, pues, apoyada por la intuición. Sin prejuzgar qué tipos de objetos (artefactos, seres biológicos, conjuntos...) satisfacen esta tesis, llamemos *O* a la clase de objetos que caen bajo ella. Entonces podemos poner de manera formal a la tesis de la siguiente manera (“*Orig* $\langle x, y \rangle$ ”, se interpreta como “*x* se origina de *y*”):

Esencialismo del origen (EO)

$$(\forall x \in O)(\Box \forall y) \Box [Orig \langle x, y \rangle \Rightarrow \Box (Existe(x) \Rightarrow (Existe(y) \wedge Orig \langle x, y \rangle))]$$

Es decir: necesariamente, si actualmente *x* se origina de *y*, entonces en cualquier mundo *w* en el que *x* exista, *y* también existirá y *x* se originará de *y* en *w*.

Ahora bien, también parece que el origen es “flexible”, en el siguiente sentido. Mi originador podría haber sido *un poco* diferente. Suponiendo que mi originador se constituye completamente por el espermatozoide y el óvulo de los

que de hecho surgió el cuerpo que me trajo a la existencia, podemos decir que: el espermatozoide o el óvulo (o ambos) podrían haber tenido un par de átomos menos, o haber tenido una configuración mínimamente diferente —y, aún así, ellos podrían haber llegado a constituir mi originador. Así que los orígenes también aceptan una variación, aún si ésta es mínima. Cerrar estas posibilidades nos lleva a postular algo contraintuitivo: que el pedazo de plástico de mi silla no podría haber tenido *ni una* partícula diferente. Cerrar estas posibilidades, para decirlo brevemente, nos lleva al superesencialismo. Por otro lado, también es difícil decir cuál es el número diferente de partículas que el plástico en cuestión podría haber tenido, y aún así tener abierta la posibilidad de ser el originador de mi silla.

Esta flexibilidad, que llamaremos **Tolerancia**, constituye una intuición que da plausibilidad a **EO**. **Tolerancia** es, como su nombre nos dice, la intuición de que los originadores *toleran* ciertas diferencias posibles, de que los originadores también están sujetos a **MOD DE RE**. Una formulación más exacta (que, por supuesto, no elimina toda la vaguedad, pero a ello volveremos más abajo) sería así ("*MinDif* $\langle x, y \rangle$ " se interpreta como "x es mínimamente diferente de y", y suponemos que ésta es una relación que puede satisfacerse incluso por objetos en mundos posibles diferentes):

(Tolerancia)

$$(\forall x \in O)(\Box \forall y)(\Box \forall z)[(Orig \langle x, y \rangle \wedge MinDif \langle y, z \rangle) \Rightarrow \Diamond Orig \langle x, z \rangle]$$

Tolerancia y **EO** nos dan la siguiente imagen: es esencial a los objetos en O el haberse originado de sus originadores actuales, y estos originadores podrían haber sido mínimamente diferentes y aún así ser los originadores esenciales de los objetos en O . Pero, aunque podrían haber sido mínimamente diferentes, los originadores actuales no podrían haber sido completamente diferentes y aún así ser los originadores esenciales de los objetos en O .

Ya tenemos la formulación de la teoría origen-esencialista. Como vimos, se compone por dos tesis características. Como también vimos, esta teoría postularía propiedades que satisfacen **RDU** —lo cual, a su vez, resultaría en que esta teoría bloquea el argumento a la **Paradoja de indiscernibilidad**. Así que si aceptar el esencialismo del origen nos libra de la paradoja. ¿Por qué no aceptarlo?

3.3.2. ...Y sus problemas

El esencialismo del origen es por sí mismo un esencialismo problemático. Pues aunque parece intuitivamente verdadero, enfrenta dos problemas (enunciados explícitamente por Penelope Mackie (2007, p. 98)):

(i) El problema de justificar por qué es el *origen* esencial, y no otra propiedad de los objetos involucrados.

(ii) El problema de justificar qué *propiedades relativas al origen* son las esenciales de tales objetos, y no otras.

Llamaré al primer problema el *problema de la fundamentación del esencialismo del origen*; al segundo, lo llamaré el *problema de la demarcación del esencialismo del origen*.

El problema de la demarcación es difícil, pero no diré mucho sobre él. Parece más importante resolver el problema de la fundamentación: antes de decir qué propiedades del origen son las esenciales, deberíamos decir si *de hecho* hay propiedades del origen esenciales.

Para resolver el problema de fundamentación tenemos abiertas dos opciones.

La primera sería un argumento a la mejor explicación: si decimos que el esencialismo del origen bloquea la paradoja, entonces podremos afirmar que la mejor explicación de por qué nuestra metafísica modal no es paradójica es que el esencialismo del origen es verdadero. Sin embargo, surgirá la pregunta ¿por qué éste esencialismo y no otro? Después de todo, *cualquier* otra teoría esencialista que postule la existencia de propiedades que satisfacen **RDU** también bloquea la paradoja. Entonces, para responder, podemos dejarlo todo a la intuición, y decir que el esencialismo del origen es la teoría que bloquea la paradoja y, además, es la más intuitiva (apelaremos a la contingencia del desarrollo o hechos por el estilo). O podemos intentar dar un argumento que pruebe, *específicamente*, la verdad del esencialismo del origen.

Los autores arriba mencionados han intentado dar un argumento que pruebe específicamente la verdad del esencialismo del origen, y ésta sería la segunda estrategia. Me enfocaré en la estrategia seguida por Kripke, Salmon, Forbes y (argüiblemente), Rohrbaugh-deRosset.²⁵

Los argumentos de Salmon, Forbes y Rohrbaugh-deRosset parecen compartir una misma estrategia:

- (1) afirmar que dos hechos son *conjuntamente* posibles en un mismo mundo (*i.e.*, que dos hechos son *composibles*);
- (2) afirmar el hecho trivial de que un objeto sólo tiene un originador en un mismo mundo;
- (3) afirmar que un originador sólo origina una cosa en un mismo mundo; y luego, basándose en **NEC=** aplicado a las dos mesas,
- (4) concluir el esencialismo del origen.

Cómo se logra concluir (4) se ve con el siguiente ejemplo: si yo podría

²⁵ Argüiblemente, módulo un criterio de individuación para argumentos, Forbes, Salmon y Rohrbaugh con deRosset estarían utilizando el mismo argumento presentado de diferentes maneras. Véase CAMERON & ROCA 2006 y ROBERTSON & FORBES 2006 para la defensa de que, esencialmente, los tres argumentos mencionados proceden con la misma estrategia. El argumento de J.L. Mackie (1974) ha sido criticado por McGinn (1979), y el argumento de éste ha sido criticado por Forbes (1985, capítulo 3).

originarme de mis gametos actuales G_c y los gametos G_o de los Obama podrían dar origen a otra persona cualquiera, entonces en un mundo w pasan conjuntamente las dos cosas: yo me origino de mis G_c , mientras que los gametos G_o originan a una persona que podemos llamar "Carl" sin prejuizar sobre su identidad. Sabemos que Carl es diferente de mí en w , notando que ninguna cosa se origina de dos diferentes originadores en un mismo mundo. Además, la tercera premisa nos permite concluir que, en cualquier mundo w , los gametos G_o sólo originan a Carl en w , y mis gametos G_c sólo me originan a mí en w . Así que en cualquier mundo donde esos gametos de los Obama den origen a algo, ese algo será Carl mismo; pero como Carl y yo podríamos nacer en un mismo mundo, entonces no seremos los mismos objetos. Por **NEC=**, Carl y yo somos necesariamente diferentes. Como no hay nada especial en Carl y los gametos de los Obama, generalizamos y obtenemos: para todo mundo w , cualquier cosa que no se origine de mis gametos actuales en w , no es yo mismo. Por equivalencia lógica, obtenemos: en cualquier mundo donde exista, me origino de mis gametos actuales. Es fácil ver que podemos correr argumentos análogos para otros objetos además de mí —y obtener así **EO** en cada caso.

Más esquemáticamente, y con requisitos técnicos que más adelante serán más claros, la estrategia parece ser la siguiente (véase ROBERTSON 2007, suplemento y HAWTHORNE & SZABÓ-GENDLER 2000):

$(\Box \forall x)(\Box \forall y)(\Box \forall z)(\Box \forall u):$

Composibilidad (COMP). Si un objeto x se origina de y con la configuración P (se P -origina), y es posible que *cualquier* objeto u ($u \neq x$) sea *el único* que se P -origine de z ($z \neq y$), entonces es posible que x se P -origine de y , y u sea *el único* que se P -origine de z .

Unicidad del origen (UO). Es imposible que: x se P -origine de y y x *también* se P -origine de z (en el mismo mundo).

Suficiencia (SUF). Si es posible que x sea *el único* que se P -origine de y , entonces necesariamente: todo x' que sea *el único* que se P -origine de y , es el mismo x .

Por **NEC=** aplicado a x y u , derivamos una versión de **EO**:

Esencialismo del origen* (EO*). Si x se P -origina de y , entonces necesariamente: x es el único que se P -origina de y .

Comentarios: las cláusulas sobre la configuración (o el P -origen) son necesarias para evitar los casos en que, por ejemplo, concebimos un pedazo de madera M tal que: M actualmente origina a una mesa m , pero podría haber originado a una silla t . Claramente $m \neq t$ y, a menos que tengamos cláusulas sobre configuración, rechazamos **SUF**. Con la cláusula sobre la configuración, este caso se elimina de inicio: simplemente ejemplificamos la configuración P sobre "haberse diseñado como mesa de acuerdo a su diseño actual".

Las cláusulas de unicidad ("el único P -originado") son para evitar la

posibilidad de casos del tipo Barca-de-Teseo. Como lo permite **MOD DE RE** y **CONC-->POS**, supongamos: x se P -origina de y en w_i y en t_i . Ahora supongamos que a x le cambiamos una parte en t_{i+1} , parte que integramos en otro z . En t_{i+2} , le cambiamos todavía otra parte a x para integrarla a z —y así hasta t_{i+n} , cuando le hemos cambiado toda las partes al x original y las hemos integrado a z . Ahora toda la materia original de x conforma a z , y **SUF** (a menos que tenga cláusulas de unicidad), identifica incorrectamente a x con z , y por ello la rechazamos de inicio. Con las cláusulas de unicidad, impedimos estos casos.

También podemos notar que: (i) **UO** es un ejemplo de la primera condición de **RDU** (es decir, de la condición $[(\forall G \in A)(G_w x \Rightarrow G = F)]$), suplementada con las cláusulas de configuración y unicidad, y (ii) **SUF** se parece mucho a la segunda condición de **RDU** ($[(\forall y)(F_w y \Rightarrow y = x)]$), aunque difieren en que **SUF** postula necesidad, mientras que **RDU(ii)** es una condición intra-mundo.

Si el argumento es correcto, entonces el origen es esencial. Si además aceptamos que el origen también cumple con las dos condiciones de **RDU** (de hecho, la primera se *supone* en el argumento), tendremos que el origen es la esencia individual (o un constituyente esencial de esa esencia) de todos aquellos objetos en O . Con la esencia individual, bloqueamos el argumento que nos llevaba a la **Paradoja de indiscernibilidad**, al menos para todos aquéllos objetos de los que podamos decir, con alguna plausibilidad, que tienen esencialmente su origen (*i.e.*, al menos para todos los objetos en O .)

Entonces: ¿es exitoso el argumento para **EO**?

La premisa más controversial es **SUF**. **COMP** y **UO** parecen triviales (**UO** parece ser una consecuencia de **LEIB**, mientras que **COMP** parece ser un hecho obvio: sin principios metafísicos que lo impidan, si dos hechos son posibles, entonces ambos son compositibles). Así que ahora investiguemos qué pasa con **SUF**. Tenemos dos problemas.

Primer problema. De nuevo pensemos en el caso Barca-de-Teseo. Si z podría originarse de y después de una serie de cambios en x (el que originalmente se origina de y), entonces parece que z podría originarse de y sin haber existido x ; y por ello z se originaría de y y sería diferente a x . Contraejemplo a **SUF** (que, recordemos, decía que cualquier cosa originada de y es el mismo x .)

De hecho, este contraejemplo es la misma **Paradoja de indiscernibilidad**. A través de una serie de cambios, x y z llegan a compartir todas sus propiedades. Así que una premisa esencial para concluir **EO** está sujeta a la misma **Paradoja de indiscernibilidad**. Así que no podemos establecer **EO** sin antes haber bloqueado la paradoja —y la estrategia de bloquearla con el mismo **EO** falla de inicio.

Podemos intentar bloquearla de la siguiente manera: aceptamos el *esencialismo del orden*. Decimos que es esencial a z el haber sido precedido por

x : necesariamente, si x no existe antes que z , entonces z no existe. Aplicamos entonces las cláusulas de unicidad y evitamos la paradoja (véase FORBES 2001; SALMON 2005b, p. 273). Pero permanece la pregunta de si el esencialismo del orden es suficientemente intuitivo como para postularlo en apoyo de **EO**, o si es un esencialismo tan sustantivo como para necesitar de un argumento. Como dice Robertson (2007):

El hecho de que el partidario de esta línea general de argumentación para la esencialidad del origen puede defenderse contra [este problema] al adoptar la esencialidad del orden [...] no debería llevar a uno a subestimar la importancia de [este problema]. Si la mejor defensa contra él es adoptar una marca de esencialismo que tiene menos apoyo de la intuición que del que el esencialismo del origen ya gozaba, entonces el argumento para el esencialismo del origen parece ofrecer menos apoyo para la tesis que el que ya le daba la intuición.[†]

Segundo problema. Supongamos que de hecho llegamos a la conclusión y tenemos **EO**. Ahora, por **Tolerancia** ($(\forall x \in O)(\Box \forall y)(\Box \forall z)[(Orig \langle x, y \rangle \wedge MinDif \langle y, z \rangle) \Rightarrow \Diamond Orig \langle x, z \rangle]$) si es esencial a x el haberse originado de y (más las cláusulas de configuración y unicidad), entonces también es esencial a x el haberse originado de y' , donde y' es mínimamente diferente de y . Pero la noción de *mínima diferencia* da paso a vaguedad, de la siguiente manera. Si y es mínimamente diferente de y' , y y' a su vez es mínimamente diferente de y'' , entonces y es mínimamente diferente de y'' . Y así para y y y''' , donde y''' es un objeto *muy* diferente de y . Si no queremos llegar a esta conclusión ¿en dónde hemos de poner el punto tal que un objeto ya es *muy* diferente de otro? Hacerlo parece *ad hoc*. Por otro lado, simplemente negar **Tolerancia** le quita plausibilidad a **OE**.

Así que **OE** (suplementado con **Tolerancia**, que le da plausibilidad) tiene al menos dos graves problemas: vaguedad y una nueva **Paradoja de indiscernibilidad**. Claramente, antes de que los partidarios de **OE** resuelvan estos problemas, no pueden clamar haber bloqueado la paradoja.

En el siguiente capítulo revisaremos otra estrategia para resolver la **Paradoja** de una buena vez —la estrategia que llamaré $\neg 4$.

† “The fact that the advocate of this general line of argumentation for the essentiality of origin can defend against the recycling problem by adopting the essentiality of order [...] should not lead one to underestimate the significance of the recycling problem. If the best defense against it is to adopt a brand of essentialism that has less support from intuition than origin essentialism itself already enjoyed, then the argument for origin essentialism seems to offer little more support for the claim than the intuition did in the first place”.

Capítulo 4. La estrategia $\neg 4$

Logic is sometimes conceived as metaphysically neutral, so that nothing controversial in metaphysics is logically valid. That conception devastates logic. Just about every putative principle of logic has been contested on metaphysical grounds.[†]

—Timothy WILLIAMSON (2006)

4.1. Preliminares

¿Hay una manera de evitar la paradoja? Hemos revisado la que parecía la mejor opción: suplementar nuestros principios con una nueva tesis. La opción que nos resta es rechazar alguno de los principios involucrados, y esa opción es la que revisaremos ahora.

Nuestros principios son **LOG=**, **NEC=**, **LEIB**, **MOD DE RE** y **CONC-->POS**, además de la lógica **S5-UC**. Vimos que ellos nos llevan a la **Paradoja de indiscernibilidad** de diferentes maneras. Argumenté que esta paradoja es una paradoja, un problema que nos debe hacer cuestionar aquellos principios. Ahora tomaremos la:

Estrategia restrictiva

Para evitar la **Paradoja de indiscernibilidad**, debemos rechazar al menos uno de los principios involucrados en la deducción de tal paradoja.

Argumentaré que una **Estrategia restrictiva**, la que niega el axioma 4, puede evitar la **Paradoja de indiscernibilidad**.

Notando que el razonamiento que nos lleva a la paradoja de Chisholm hace un uso esencial de la transitividad en la relación de accesibilidad definida sobre una clase de mundos posibles, han habido filósofos que intentan evitar esta paradoja rechazando la relación de accesibilidad transitiva. Ésta relación caracteriza modelo-teóricamente al axioma 4 que, a su vez, caracteriza a la lógica **S4**. Hugh Chandler parece haber sido el primero en proponerlo (Chandler 1976), y Nathan Salmon ha sido quien ha defendido este acercamiento con más detalle (Salmon 1989, 2005b Apéndice IV).

Salmon argumenta que algunas instancias del axioma modal alético 4 son al menos no-verdaderas, mientras que otras son definitivamente falsas.²⁶ Por lo

[†] "La lógica a veces ha sido concebida como metafísicamente neutral, de manera tal que nada controversial en metafísica es lógicamente válido. Esa concepción devasta a la lógica. Casi cada principio de la lógica ha sido contestado sobre bases metafísicas."

²⁶ En "Modal paradox: parts and counterparts, points and counterpoints" (SALMON 2005b,

tanto, la lógica **S4** no puede ser la lógica modal alética correcta.²⁷

El axioma en cuestión afirma que si una proposición es necesaria, entonces su ser necesaria es algo necesario: $\Box\phi \Rightarrow \Box\Box\phi$. Equivalentemente, afirma que si una proposición es posiblemente posible, entonces es posible: $\Diamond\Diamond\phi \Rightarrow \Diamond\phi$. Al dar una interpretación formal del axioma, resulta que éste se valida en todos los modelos cuya relación de accesibilidad entre mundos sea transitiva. Es decir, todos los modelos con la siguiente condición en su relación de accesibilidad validan el axioma: para todo índice i, j, k : si $w_i R w_j$ y $w_j R w_k$, entonces $w_i R w_k$. Es decir: si, dado un mundo, otro es posible relativo a éste y existe otro mundo posible relativo a este último, entonces este tercero es posible relativo al primero. Este axioma hace que un mundo sea posible relativo a otro, sin importar cuántos mundos “intermedios” existan entre ellos, siempre y cuando ambos estén conectados por una cadena de mundos conectados entre sí. Esto es ilustrado en el diagrama 3.

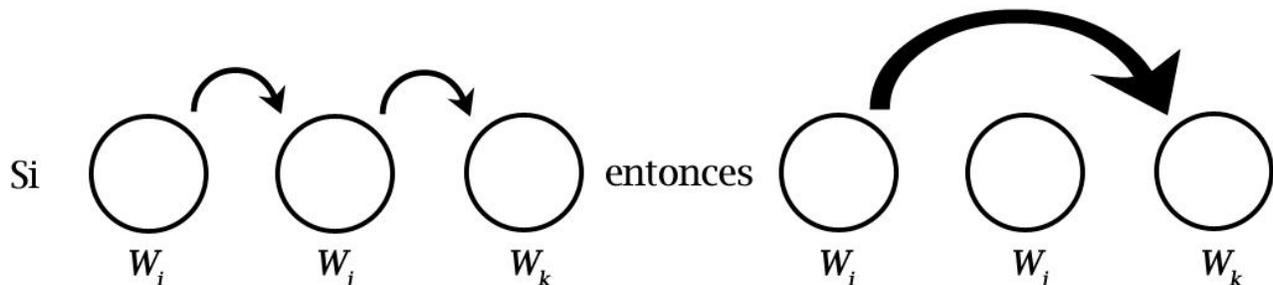


Diagrama 3. Relación transitiva de accesibilidad entre mundos.

Salmon defiende que esto no es así: existen mundos w para los cuales otro mundo w' es posible relativo a ellos, y a su vez este w' tiene acceso a otro w'' ; sin embargo, w'' no es posible relativo a w . La motivación de Salmon para el abandono de la relación transitiva de accesibilidad (y con ello del axioma 4) surge de su aceptación de una intuición esencialista: la intuición de que **EO** es una tesis verdadera.²⁸ Más abajo veremos que la *verdad* de **EO** ni siquiera es requerida para negar 4.

Supongamos que **EO** es verdadero. Supongamos que una mesa $T_{@}$ se origina actualmente de un pedazo de madera $M_{@}$. Por **EO**, la propiedad *originarse de*

Apéndice IV), Salmon defiende que la relación de accesibilidad admite grados, al menos en un sentido epistémico: para ciertos mundos w_i y w_j , el hecho de si $w_i R w_j$ es un hecho que no sabemos determinadamente si se obtiene.

27 En adelante, hablaré de axiomas y sistemas modales omitiendo decir que me refiero a la modalidad *alética*, la modalidad metafísica.

28 Salmon también argumenta que el axioma *B* (que dice que si algo es el caso entonces es necesariamente posible) aunque parece ser necesario, no parece ser *lógicamente* necesario. En el texto no me ocuparé de este axioma, sino que me enfocaré en 4, dado que es este el axioma relevante para la discusión de las paradojas.

$M_{@}$ será una propiedad que $T_{@}$ tenga en cada mundo en el que exista. Apliquemos entonces **Tolerancia**. Diremos que $T_{@}$ podría haberse originado de una madera M_2 mínimamente diferente de M_2 : diferente, digamos, por un sólo átomo. Por **REP MUN**, diremos que hay un mundo w_2 , accesible al nuestro, en el que $T_{@}$ se origina de M_2 . Y por más aplicaciones de tolerancia, llegaremos a decir que hay un mundo w_n , accesible a w_{n-1} , en el que $T_{@}$ se origina de M_n , donde M_n tiene *todos* los átomos de diferencia con $M_{@}$. Pero, por una aplicación del axioma 4, si el mundo w_2 es accesible desde nuestro mundo actual $@$, y a su vez w_2 tiene acceso a w_3 , y éste a w_2 y..., y w_{n-1} tiene acceso a w_n , entonces nuestro mundo tiene acceso a w_n . Así que será posible que $T_{@}$ se origine de M_n , un pedazo de materia completamente diferente al actual originador de $T_{@}$, $M_{@}$. Por lo tanto, **EO** es falso si aceptamos **Tolerancia** y nuestros principios.

Por lo tanto, o rechazamos algún paso del razonamiento, o tenemos incoherencia entre estas dos tesis. La solución de Salmon es rechazar un paso del razonamiento: el paso en el que se utiliza la transitividad. En la cadena de mundos habrá un mundo w_i en el que el pedazo de madera en cuestión, M_i , no es un posible originador de $T_{@}$; sin embargo, existirá un pedazo M_{i-1} que sí es un posible originador de $T_{@}$. Entonces, si M_{i-1} fuera el actual originador de $T_{@}$, M_i sería otro de sus posibles originadores (pues M_{i-1} y M_i comparten suficiente materia). Pero M_i es imposible un originador de $T_{@}$, y M_{i-1} es un originador posible de $T_{@}$.

Esto nos muestra un caso de algo que es imposible, pero que sería posible si otra posibilidad fuera actual: algo *posiblemente posible* que es, de hecho, *imposible*. El argumento fue que *necesariamente* $T_{@}$ no se origina de M_{i-1} ; sin embargo, *no es necesario que necesariamente* $M_{@}$ no se origine de tal pedazo, pues (como es posible) si se hubiera originado de M_{i-1} , entonces sería posible que se hubiera originado de M_i . Este caso implica que no todo lo que es *posiblemente posible* es *de hecho posible* —equivalentemente, no todo lo que es necesario es *necesariamente necesario*. Esto invalida el axioma 4.

Si el axioma 4 no es lógicamente verdadero (pues tiene contraejemplos surgidos en la metafísica), entonces no podemos apelar a él al construir experimentos mentales que lo involucren esencialmente. Como no podemos llegar a la Paradoja de Chisholm sin un razonamiento que involucre al axioma 4, si rechazamos este axioma de entrada podemos bloquear la inferencia.

Claramente, muchos tenemos intuiciones esencialistas. Pero ello no es suficiente, pues o suplementamos esas intuiciones con principios de tolerancia análogos a **Tolerancia** o no lo hacemos. Si no lo hacemos, nuestras intuiciones corren el riesgo de parecer demasiado estrictas (diremos, por ejemplo, que una mesa sólo podría haberse originado de un pedazo de materia exactamente traslapado con su originador actual: no cabe diferencia siquiera de un átomo), y

si lo hacemos, damos entrada a razonamientos que nos llevarán a paradojas chisholmianas.

Si rechazamos el axioma 4, entonces dada la cadena de mundos $\langle @, \dots w_n \rangle$, hay un mundo w_i en la cadena que ya no es accesible desde @, es decir, que no es una posibilidad para @. Así rompemos la cadena y nunca llegamos, en la argumentación, al mundo paradójico w^* .

Así que el panorama dialéctico parece ser el siguiente: tenemos los principios, y nos llevan a una paradoja. Bloqueamos esta paradoja con una teoría esencialista, que suplementamos con un principio de tolerancia que la hace plausible, pero éste, a su vez, lleva de nuevo a la paradoja. Como la solución esencialista (junto con el principio de tolerancia) era la más plausible, la fuente de la paradoja debe estar en otro lado, y eso nos lleva a los principios de nuevo — y el candidato más plausible para ser la fuente de la paradoja es el axioma 4. Uno podría preguntarse dos cosas. ¿De verdad es el más plausible? Y: Si bloqueamos la paradoja rechazando el axioma 4, entonces ¿Para qué invocar a la intuición esencialista?

4.2. Rechazando el axioma 4

¿De verdad 4 es el candidato más plausible para ser rechazado? Lewis dice que no. Leemos en *On the plurality of worlds* (LEWIS 1986, p. 246):

Se supone que hay una defensa [contra la paradoja de Chisholm] diferente [a la propuesta por el teórico de las contrapartes]. En lugar de basarse en la intransitividad de la relación de contraparte, para bloquear el movimiento fatal —de hecho, sin asumir teoría de las contrapartes en absoluto— en lugar de ello podríamos confiar en la inaccesibilidad de los mundos. [...] Así que desde el punto de vista de [$@$], [w_n] en realidad no cuenta —¡Sin preocupaciones! No es así, *si* ignoramos mundos inaccesibles desde el nuestro, que tenemos un caso de diferencia haecceitística. [29] [...] Digo yo: esto no es defensa, esto es capitulación. En estas cuestiones de haecceitismo y esencia, ¿Con qué derecho ignoramos mundos que son condenados a ser inaccesibles? Accesibles o no, ellos aún son mundos. Todavía creemos en ellos. ¿Por qué no cuentan?†

29 Cuando Lewis habla de “diferencia haecceitística”, se refiere a lo que en esta tesis he venido nombrando como ‘diferencia desnuda’: una diferencia entre dos objetos que no se basa en absoluto en las propiedades de ellos: x y y comparten todas sus propiedades, pero aún así son diferentes.

† “There is supposed to be a different defence available. Instead of relying on intransitivity of the counterpart relation to block the fatal move -indeed, without assuming counterpart theory at all - we could instead rely on inaccessibility of worlds. So from the standpoint of W_1 , W_3 doesn't really count - no worries! It is not so, *if* we ignore worlds inaccessible from ours, that we have a case of haecceitistic difference. [...] Say I: this is no defence, this

Lewis afirma que el rechazo de 4 no es una defensa del esencialismo, sino una capitulación ante el anti-esencialismo: pues aún si afirmáramos que los mundos paradójicos en los que existen diferencias desnudas son inaccesibles desde nuestro mundo, estos mundos *existen*. Y si existen, y en ellos se da la diferencia haecceitística (desnuda), entonces (por **REPMUN**) esta diferencia es posible, en contradicción con la teoría esencialista (y con los otros principios). Conclusión: inaccesibles o no, los mundos paradójicos son mundos posibles que existen, y ellos satisfacen la contradicción al esencialismo. El rechazo de 4 no ha logrado una defensa de la teoría.

La argumentación de Salmon para responder a Lewis apela en este punto a una concepción global sobre las posibilidades, o algo mucho más abarcador que las posibilidades. Salmon considera que tenemos *dos* nociones que suelen ser conflagradas: una noción genérica y una noción particularmente modal. La noción genérica es la noción de una *manera de ser de las cosas*, ("a way for things to be") la noción modal es la noción de una *manera en que las cosas podrían haber sido* ("a way things *might have been*") (SALMON 1989, p. 11). Ambas son nociones que involucran entidades abstractas máximas, entidades *de acuerdo a las cuales* los hechos o estados de cosas se obtienen o no. La diferencia es que la noción modal es la noción de un tipo de entidades de acuerdo a las cuales ciertos hechos o estados de cosas *podrían* haberse dado; mientras que la noción genérica incluye a las entidades de acuerdo a las cuales el mundo está en cierto estado (ciertos hechos o estados de cosas se obtienen), incluso si el mundo *no pudiera* haber estado en tal estado. En otras palabras, la noción genérica incluye tanto mundos posibles como *mundos imposibles*.

Escribe Salmon (1989, p. 12):

Estrictamente hablando, un mundo posible no es una manera de ser de las cosas que podrían haber existido; es una manera de ser de las cosas tal que las cosas podrían haber sido de esa manera. [...] Así que la palabra "posible" contribuye algún significado especial a la frase, y más significado del que es acomodado por la noción genérica de una total manera-de-ser-de-las-cosas-incluso-si-las-cosas-no-podrían-haber-sido-de-esa-manera. Estrictamente hablando, un mundo posible no es ninguna vieja manera de ser de las cosas, sino una clase modalmente especial de una manera total de ser de las cosas, a saber, una manera total de ser que las cosas *podrían haber sido*. Un mundo posible es una manera total de ser de las cosas que se conforma a los constreñimientos metafísicos concernientes a lo que podría haber sido. La noción genérica de una manera total de ser de las cosas es una noción sin un

is capitulation. In these questions of haecceitism and essence, by what right do we ignore worlds that are deemed inaccessible? Accessible or not, they're still worlds. We still believe in them. Why don't they count?"

término propio.[†]

Así que tenemos maneras de ser de las cosas de dos tipos: una tal que las cosas *podrían* haber sido de esa manera, y otra tal que las cosas *no podrían* haber sido de esa manera. Llamemos a la noción genérica 'mundos'. Entonces los mundos incluyen tanto a los mundos posibles como a los mundos imposibles. Si esto es correcto, entonces, según Salmon "la idea de que la necesidad metafísica se corresponde con la verdad en todo mundo cualquiera está francamente errada" (p. 15).[†] Pues, si tomamos en cuenta a los mundos imposibles, entonces "ninguna proposición es verdadera de acuerdo a todo mundo y toda proposición es verdadera de acuerdo a algún mundo" (*ibid.*).^{††} Incluso las falsedades lógicas serán verdaderas en algún mundo: en un mundo lógicamente imposible, puede darse el caso de que tanto p como $\neg p$. Incluso las verdades lógicas serán falsas en algún mundo: en un mundo lógicamente imposible, una proposición no se implica a sí misma. Si nunca consideramos estos casos, de acuerdo con Salmon, es porque siempre nos restringimos a la necesidad lógica o a la necesidad metafísica.

¿Por qué creer que existen los mundos imposibles? Bueno, según Salmon, casi cada argumento a favor de la existencia de mundos posibles tiene un análogo para la existencia de mundos imposibles, incluso lógicamente imposibles. Y cada argumento a favor de sólo la existencia de mundos posibles, confunde la noción de *mundo* con la de *mundo posible*, o confunde la noción de *mundo imposible* con la de *mundo que no puede existir*, o ambas. Pues

[...] el hecho de que $[T@]$ no se puede originar de $[M_n]$ implica (*entails*) que originarse de $[M_n]$ es una manera de ser tal que $[T@]$ no puede ser de esa manera. Se sigue de esto último que $[T@]$ originándose de $[M_n]$ y Sócrates siendo sabio y ... (donde "todas" las cuestiones de hecho son fijadas), es una manera máxima en la que las cosas (en general) no pueden ser. Se sigue de el hecho de que $[T@]$ no se puede originar de $[M_n]$, entonces, que hay una

† "Strictly speaking, a possible world is not a way for things to be that might have existed; it is a way for things to be such that things might have been that way. Similarly, a possible history or possible state for an individual is not a history or state that might have existed, but a history or state that the individual might have had or might have been in. Thus the word "possible" contributes some special meaning to the phrase, and more meaning than is accommodated by the generic notion of a total way-for-things-to-be-even-if-things-could-not-have-been-that-way. Strictly speaking, a possible world is not any old total way for things to be, but a modally special kind of total way for things to be, namely a total way that things might have been. A possible world is a total way for things to be that conforms to metaphysical constraints concerning what might have been."

† "[...] the idea that metaphysical necessity corresponds to truth in every world whatsoever is flatly mistaken".

†† "[...] no proposition is true according to every world and every proposition is true according to some world."

manera máxima en que las cosas no pueden ser. (Salmon 1989, p. 16)^{†††}

La única diferencia, entonces, entre mundos posibles e imposibles no es que un tipo de mundos exista y el otro no: ambos tipos de mundos existen, sólo que uno podría haberse realizado y otro no. Todo esto, por supuesto, supone que las nociones modales de necesidad y posibilidad no se definen o analizan en términos del aparato de mundos posibles, sino viceversa: uno tiene que comprender primero las nociones modales para poder comprender la noción de mundo posible.

Con toda esta teoría Salmon está en la posición de poder definir la relación de accesibilidad entre mundos, *i.e.*, la noción de *ser posible relativo a otra posibilidad*: w tiene acceso a w' si y sólo si todo hecho de w' es una posibilidad en w , si y sólo si todo hecho necesario en w' se obtiene en w . De esta manera, “[...] la relación de accesibilidad de un mundo con otros mundos es interna al mundo, *vía* las posibilidades en este mundo” (*ibid.*, p. 19).^{††††}

Uno puede entonces responder a Lewis de la siguiente manera: los mundos inaccesibles no cuentan porque no son posibles. Son mundos imposibles. En ellos, suceden cosas que son prohibidas por los principios metafísicos. En la serie de mundos $\langle @, \dots, w_n \rangle$ postulada por el argumento a la paradoja de Chisholm, por ejemplo, debe haber un mundo que sea imposible, un mundo al que @ ya no pueda tener acceso. Y ésto es asegurado por el esencialismo, lo cual nos lleva a una pregunta que dejamos pendiente.

4.3. Del esencialismo a $\neg 4$

Recordemos lo que dije hace algunas páginas:

Así que el panorama dialéctico parece ser el siguiente: tenemos los principios, y nos llevan a una paradoja. Bloqueamos esta paradoja con una teoría esencialista, que suplementamos con un principio de tolerancia que la hace plausible, pero éste a su vez lleva de nuevo a la paradoja. Como la solución esencialista (junto con el principio de tolerancia) era la más plausible, la fuente de la paradoja debe estar en otro lado, y eso nos lleva a los principios de nuevo —y el candidato más plausible para ser la fuente de la paradoja es el axioma 4. Uno podría preguntarse dos cosas. ¿De verdad es el más plausible? Y: Si bloqueamos la paradoja rechazando el axioma 4,

††† “The fact that Woody cannot originate from m entails that originating from m is a way Woody cannot be. It follows from the latter that Woody originating from m and Socrates being wise and... (where “all” questions of fact are fixed), is a maximal way that things (in general) cannot be. It follows from the fact that Woody cannot originate from m , therefore, that there is a maximal way things cannot be.”

†††† “[...] a world's accessibility relations to other worlds is internal to the world, via the possibilities at that world”

entonces ¿Para qué invocar a la intuición esencialista?

He intentado responder, junto con Salmon, a la primera pregunta en la sección anterior. En esta sección intentaré responder, también siguiendo a Salmon, a la segunda pregunta.

Hay que notar que la *verdad* de **OE** suplementado con **Tolerancia** *no* se requiere para rechazar el axioma 4. Lo único que se requiere es que la conjunción de ambos sea *coherente*, que sea *lógicamente posible*. Pues ambas tesis resultan, dice Salmon, en “[..] un *modelo* o *interpretación* que respeta tanto la interpretación pretendida de las constantes lógicas, incluyendo “necesariamente” [...] e invalida S4” (1989, p. 31).[†]

Claramente estas intuiciones son lógicamente posibles: por todo lo que la lógica pura nos dice, un objeto podría tener ciertas propiedades que ejemplifica en su origen (llamémoslas *propiedades-de-origen*) de manera esencial. (O no). Por todo lo que la lógica pura nos dice, las propiedades-de-origen esenciales de la clase relevante de objetos, podrían cumplir con sus respectivos principios de tolerancia. (O no). Entonces las intuiciones son lógicamente posibles (aún más, son lógicamente contingentes).

El argumento para rechazar 4 desde las intuiciones parece ser el siguiente: por definición, un axioma lógico es lógicamente necesario. Y si es lógicamente posible que un pretendido axioma falle, entonces no es lógicamente necesario, y no puede contar como axioma lógico, ni la lógica que lo supone puede ser un sistema correcto. Así que rechazamos **S4**.

Dos comentarios sobre esto. Primero, uno *no* necesita ser esencialista del origen para rechazar 4. Uno puede adoptar cualquier esencialismo *acerca de individuos*³⁰ —con la condición de que tal esencialismo sea lógicamente posible. Cualquier intuición, que no sea inconsistente, sobre cualquier propiedad esencial de los individuos tendrá igual fuerza argumental contra 4. Así que si uno no tiene la intuición esencialista del origen pero sí la intuición sobre, digamos, *destrucción-esencialismo* (por definición: las propiedades de la destrucción de un objeto les son esenciales), uno puede correr argumentos análogos al de Salmon, llegar a la paradoja chisholmiana, y negar 4.

Así que una intuición sobre cualquier tipo de esencialismo de individuos es lógicamente incompatible con 4. Y si esto es así, entonces parece que el esencialismo del origen no tiene ningún privilegio en el plano lógico: está al mismo nivel que los otros esencialismos. Quizá el esencialismo del origen tenga más apoyo intuitivo, pero, por todo lo que dice la argumentación de Salmon, el aspecto relevante no es el intuitivo —sino la mera coherencia, la posibilidad

† “[...] a model or interpretation that both respects the intended interpretation of the logical constants, including “necessarily” [...], and invalidates S4.”

³⁰ *Versus* esencialista de clases (naturales o no), conjuntos, universales, etcétera.

lógica. Por lo tanto, todo esencialista sobre los individuos, e incluso todo aquél que no sea esencialista pero crea que el esencialismo es lógicamente posible, rechazará el axioma 4.

Segundo, a menos que tengamos un argumento para tal o cual teoría esencialista, uno también podría decir que la paradoja chisholmiana es una reducción al absurdo de la conjunción de las intuiciones esencialistas y sobre la tolerancia con **MOD DE RE**. El principio de tolerancia tiene un fuerte apoyo intuitivo, pero parece quedar abierta la posibilidad teórica de negar este principio y aceptar tanto el esencialismo como el axioma 4: nos volveríamos *superesencialistas* respecto a los originadores.

Por mi parte, acepto la argumentación de Salmon, dado que tengo la intuición de que el principio de tolerancia es verdadero y estoy convencido de la posibilidad lógica del esencialismo. Por lo tanto rechazo 4, y acepto que este rechazo bloquea la paradoja de Chisholm —con lo que también bloquea la de Mackie.

Este bloqueo puede hacerse de dos maneras. Primera: sin el axioma 4, **EO** queda libre de problemas, y tenemos abierta la opción de postular que las propiedades relevantes del origen también satisfacen **RDU**, con lo cual tendremos que el origen es la esencia individual de muchos objetos. Con esencias individuales, bloqueamos de inicio la motivación para la paradoja de Mackie. Aunque, incluso si no postulamos que las propiedades del origen satisfacen **RDU** (y con ello, tampoco postulamos que el origen de x es su esencia individual), podemos bloquear la paradoja de Mackie si bloqueamos la de Chisholm.

La segunda manera de bloquear la paradoja de Mackie es la siguiente. Notemos que el razonamiento a esta paradoja no opera por pequeños cambios. Simplemente apela a las definiciones y principios: si no prohíben una posibilidad, entonces no hay razón para decir que ésta no se obtiene —no una razón, al menos, pre-teórica. Y nuestros principios no prohíben que toda posibilidad de x sea una posibilidad de y —*excepto por LEIB*. Pues, por lo que **LEIB** dice, la posibilidad de x de que tenga todas las propiedades de y no existe —éste hecho sólo se da en mundos imposibles: si y tiene todas las propiedades de x , entonces y es el mismo x . Sin embargo, ésta posibilidad está permitida por **MOD DE RE** —o así parecía. **MOD DE RE** nos permite mover a los objetos en las posibilidades, *siempre que esas posibilidades estén dentro de los límites de la consistencia*. Pero si aceptamos a **LEIB** como principio, entonces la posibilidad de los mundos *switching* (los mundos donde se da el intercambio de propiedades) está excluida de inicio.

De hecho, fuera de la negación de **LEIB**, no hay otra motivación para llegar a la paradoja, más que la intuición de los “pequeños cambios” que funcionaba en

la derivación de la paradoja de Chisholm. Como dice Lewis al hacer su presentación de la paradoja de Chisholm (1986, p. 249, énfasis mío):

Por favor concédanme que las esencias toleran suficiente revisión de tal manera que, necesariamente, un ocupante del rol de Adán podría haber ocupado el rol de Noém [donde "Noém" es el "intermedio" entre Adán y Noé] y un ocupante del rol de Noém podría haber ocupado el rol de Noé. No hay punto en negarlo –*si lo hacen, sólo incrementaré el número y decreceré la diferencia de roles intermedios hasta que la negación se vuelva absurda*, así complicando el ejemplo sin ningún buen propósito.[†]

Así que neguemos que las esencias son así de tolerantes: **LEIB** lo impide. Entonces, tendríamos que poner más "roles intermedios" (*i.e.*, hacer más larga la serie de mundos), a través de los cuales los cambios sean cada vez más pequeños, lo cual iría más acorde con la intuición y nos permitiría llegar a la paradoja de Chisholm –*si tuviéramos el axioma 4*. Sin éste, no tenemos nada que nos asegure la inferencia, y hemos bloqueado *por fin* la **Paradoja de indiscernibilidad**, tanto en su forma chisholmiana como en su forma mackieana.

† "Please grant me that essences tolerate enough revision so that, necessarily, an occupant of the Adam role could have occupied the Noam role, and an occupant of the Noam role could have occupied the Noah role. No point denying it - if you do, I will just increase the number and decrease the difference of intermediate roles until the denial becomes absurd, thereby cluttering the example to no good purpose. "

Conclusiones

Como dije en la introducción, los filósofos se interesan en sistematizar las intuiciones que tenemos sobre hechos cotidianos. En esta tesis me enfoqué en la sistematización de las intuiciones sobre la modalidad *de re*, sobre los modos en que las cosas pueden (y deben) ser, conjugadas con las intuiciones y leyes lógico-metafísicas sobre la identidad y la diferencia. Hice notar cómo es que esa sistematización llevaba a un escenario paradójico de diferentes maneras, que llamé **Paradoja de indiscernibilidad**.

Como también dije en la introducción, hay varias salidas para esta paradoja. De inicio, uno puede rechazar la lógica dentro de la cual esta paradoja y varias soluciones ofrecidas a ella se inscriben. Por ejemplo, uno puede quedarse con la lógica de primer orden clásica y construir una teoría sobre ella, sin suplementarla con axiomas modales y constantes lógicas modales. Este acercamiento es el que toman los teóricos contrapartistas. Pero se han ofrecido muchas opciones y varias maneras de definir los conceptos relevantes. Las nociones de *necesidad* (y *posibilidad*), de *mundo posible*, de *esencialidad*, han estado, cada una, sujetas a un vigoroso debate.

La extensión misma de ese debate, y la extensión de mi tesis, me impidieron explorar otras opciones. Para intentar bloquear la **Paradoja de indiscernibilidad**, tomé la salida que casaba mejor con mis intuiciones: la idea de que la lógica modal estándar es la lógica de lo necesario y lo posible correcta, mientras que los objetos pueden cambiar algunas de sus propiedades, siempre que este cambio no sea demasiado radical. Así que acepté seguir el camino del esencialismo para intentar bloquear el problema.

Como vimos en la tesis, el esencialismo mismo daba pie a la misma paradoja. De nuevo, aquí hay varias opciones. Guiado todavía por mis intuiciones, intenté ser lo más conservador posible respecto a la base lógica con la que formalizaríamos nuestra teoría. Esto implicaba intentar retener a toda costa la lógica modal estándar —lo cual, irónicamente, nos llevó a “mutilarla” y rechazar uno de sus axiomas. Este rechazo, si está bien fundado, falsea la creencia de muchos filósofos de que la lógica correcta para la posibilidad metafísica es la lógica **S5**: como el axioma 4 no es válido, la lógica **S5** —que incluye esencialmente al axioma 4— no puede ser la lógica correcta para la metafísica. La argumentación de Salmon me convenció de que la lógica modal que tiene los axiomas “transparentemente lógicamente verdaderos” —como el mismo Salmon dijo— es la lógica **T**, cuyo axioma característico afirma que todo lo necesariamente verdadero es verdadero *simpliciter*. Axiomas más arriesgados parecen correr el riesgo de estar sujetos a la paradoja.

Aquí notamos también cómo el esencialismo y la estrategia $\neg 4$ de Salmon se

llevaban de la mano. El esencialismo necesita de la negación del axioma para poder ser verdad, mientras que la negación del axioma necesita de la *coherencia* del esencialismo para estar bien fundamentada. Negamos el axioma y la paradoja no surge, con lo cual podemos regresar a intentar dar una versión más intuitiva de la teoría esencialista.

En cierto sentido, volvimos al inicio. Vimos que uno de los principios que habíamos supuesto —que todo lo necesario es necesariamente necesario, equivalentemente, que todo lo posiblemente posible es posible— debe negarse, si es que intentamos hacer metafísica modal esencialista. Con la posibilidad de hacer metafísica modal esencialista, tenemos el comienzo de una explicación sobre los modos en que las cosas podrían ser: al menos ahora sabemos que la postura que afirma que las cosas sí podrían “tolerar” ciertas diferencias y aún así permanecer ellas mismas (siempre que esas diferencias no sean radicales), *ésta* postura es una postura coherente y plausible. El único costo que, hasta ahora, parece requerir, es el rechazo de un axioma modal. Qué suceda con las otras posturas que no analicé aquí, será tema de investigaciones posteriores, donde también me preguntaré sobre su coherencia, plausibilidad y costos.

Finalmente, me gustaría notar mi motivación inicial para investigar la modalidad *de re*. Los filósofos interesados en la aplicación práctica de la filosofía podrían no ver inmediatamente cómo la investigación sobre los modos en que las cosas pueden ser, su potencialidad y su esencia, podría ser aplicado fuera del minado campo de la metafísica analítica. Sin embargo, al tomar cursos de historia de la filosofía noté que la discusión en torno a la esencia de las cosas, así como la discusión en torno a “casos límite” de aplicación de ciertos conceptos, eran discusiones centrales en varias épocas y en varias ramas. Seguramente la historia de la filosofía está plagada con concepciones rivales de lo que sea ser esencial (y hoy en día hay un debate sobre ello), pero intentar comprender las implicaciones de *una* de ellas me pareció una tarea digna de una tesis. Esto, pensé, podría darme la oportunidad de tener una visión con la cual entrar a debates no sólo metafísicos (como los debates en torno al cambio, a la identidad, a la existencia y a la posibilidad), sino también de otras ramas filosóficas. Y aunque puede ser que al preguntarse cuál sea la esencia de la mente, de lo bueno o de la belleza, uno no avance mucho apelando a las discusiones en torno al esencialismo, quizá el apelar a ellas sí que evite alguno que otro retroceso.

De cualquier manera, ha sido emocionante el sumergirme en este debate y poder notar que hay un surtido catálogo de problemas, profundos y estimulantes, esperando respuesta. Me gustaría, entonces, concluir parafraseando a Kripke (2005, p. 75): Hay que fijarse cuánto puede alzarse el oleaje en los mares del ser. Hasta en las partes menos profundas.

Referencias

ADAMS, Robert.

- 1979. Primitive thisness and primitive identity. *The Journal of Philosophy*, **76**, 1: pp. 5-26.

BLACK, Max.

- 1952. The identity of indiscernibles. *Mind* **61**, #242, pp. 153 – 164. Reimpreso en Jaegwon Kim & Ernest Sosa (eds.): *Metaphysics: An Anthology*, pp. 66 – 71. Blackwell.

CAMERON, ROSS & ROCA, Sonia.

- 2006. Rohrbaugh and deRosset on the necessity of origin. *Mind* **115**, pp. 361 – 366.

CHANDLER, Hugh.

- 1976. Plantinga and the contingently possible. *Analysis* **36**, pp. 106 – 109. Reimpreso en TOOLEY, Michael (1999): *Necessity and possibility. The metaphysics of modality*, pp. 310 – 314.

CHANG, C.C. & KEISLER, Jerome H..

- 1990. *Model theory*. Tercera edición. North Holland: Studies in logic and the foundations of mathematics.

CHISHOLM, Roderick.

- 1967. Identity through possible worlds: Some questions. *Noûs* **1**, #1, pp. 1–8. Reeditado en CHISHOLM 1989, capítulo 3.
- 1979. *Person and object*. Open Court Publishing Company.
- 1989. *On metaphysics*. University of Minnesota Press.

COBURN, Robert C..

- 1986. Individual essences and possible worlds. *Midwest Studies on Philosophy*, **11**: *Essentialism*.

COCHIARELLA, Nino B. & FREUND, Max A.

- 2008. *Modal logic. An introduction to its syntax and semantics*. Oxford University Press.

FINE, Kit.

- 1978. Model theory for modal logic – Part I: The *De Re/De Dicto* distinction. *Journal of Philosophical Logic* **7**, pp. 125 – 156.

FITTING, Melvin.

- 1999. Barcan both ways. *Journal of Applied Non-Classical Logics* **9**, pp. 329–344. (Disponible en <http://comet.lehman.cuny.edu/fitting/bookpapers/pdf/papers/Barcan.pdf>, accesado en Noviembre 2009.)

FORBES, Graeme.

- 1984. Two solutions to Chisholm's Paradox. *Philosophical Studies* **46**, pp. 171 – 187.
- 1985. *The metaphysics of modality*. Oxford University Press.
- 1986. In defense of absolute essentialism. *Midwest Studies in Philosophy* **11**, pp. 3 – 29.
- 1994. A new riddle of existence. *Philosophical perspectives* **8: Logic and language**, pp. 415 – 430.
- 2001. Origins and identities. Editado en Bottani & Giaretta & Carrara (eds.): *Individuals, essence and identity. Themes of analytic metaphysics*, 2001. Reidel.
- 2008. Identity and the facts of the matter. *Draft de Septiembre 2008*, disponible en http://spot.colorado.edu/~forbesg/pdf_files/Id_and_FoMs.pdf (accesado en Octubre 2009)

GIBBARD, Allan.

- 1975. Contingent identity. *Journal of philosophical logic* **4** #2, pp. 187 – 221.

HAWTHORNE, John & SZABÓ-GENDLER, Tamar.

- 2000. Origin essentialism: The arguments reconsidered. *Mind* **109**, #434, pp. 285 – 298.
- 2006. *Conceivability and Possibility*. Oxford University Press.

HALLECK, John.

- 2008. Logic systems: System S5–Universal closure. Disponible en: <http://home.utah.edu/~nahaj/logic/structures/systems/s5-uc.html>, accesado última vez en Septiembre de 2009.

HAYAKI, Reina.

- 2006. Contingent objects and the Barcan Formula. *Erkenntnis* **64**, pp. 75 – 83.

HAZEN, Allen.

- 1979. Counterpart-theoretic semantics for modal logic. *The Journal of Philosophy* **76** #6, pp. 319 – 338.

HOBBS, Thomas.

- 1655. *De Corpore*.

HUGHES, G.E. & CRESSWELL, M.J.

- 1968. *A new introduction to modal logic*. Routledge.

KAPLAN, David.

- 1975. How to Russell a Frege-Church. *The Journal of Philosophy* **72**: #19, pp. 716 – 729.

KRIPKE, Saul.

- 1963. Semantical considerations on modal logic. *Acta Philosophica Fennica* **16**: *Modal and many valued logic*, pp. 83 – 94.
- 2005. *El nombrar y la necesidad*. Traducción de Margarita M. Valdés de 1981: *Naming and necessity*, Blackwell. UNAM, Instituto de Investigaciones Filosóficas.

LEIBNIZ, Gottfried Wilhelm von.

- 1686. *Discours de métaphysique*. Traducción de Jonathan Bennet, *Discourse on metaphysics*. Disponible en <http://www.earlymoderntexts.com/pdf/leibdm.pdf> (accesado en Noviembre 2009).

LEWIS, David.

- 1968. Counterpart theory and quantified modal logic. *The Journal of Philosophy* **65**, #5, pp. 113 – 126. Reimpreso en Lewis 1983a, pp. 26 – 38.
- 1983a. *Philosophical Papers, Volume I*. Oxford University Press.
- 1986. *On the plurality of worlds*. Blackwell.

MACKIE, J.L..

- 1974. *De what re is de re modality?*. *The Journal of Philosophy* **71**, #16, pp. 551 – 561.

MACKIE, Penelope.

- 2007. *How things might have been*. Oxford University Press.

MARCUS, Ruth Barcan.

- 1967. Essentialism in modal logic. *Noûs* **1**, pp. 90 – 96.

McGINN, Colin.

- 1976. On the necessity of origin. *The Journal of Philosophy* **73**, #5, pp. 127 – 135.

MENZEL, Christopher.

- 2009. Actualism. *The Stanford Encyclopedia of Philosophy* (Edición de invierno de 2008). Edward N. Zalta (editor). Disponible en <http://plato.stanford.edu/archives/win2008/entries/actualism/> (accesado en Noviembre de 2009)

PLANTINGA, Alvin.

- 1974. *The nature of necessity*. Oxford University Press.
- 1976. Actualism and possible worlds. Reimpreso en PLANTINGA 2003, pp. 103 – 121.
- 1983. On existentialism. Reimpreso en PLANTINGA 2003, pp. 158 – 175.
- 2003. *Essays in the metaphysics of modality*. Oxford University Press.

QUINE, Willard Van Orman.

- 1961. Reference and modality. Editado en *From a Logical Point of View*, pp. 139 – 159. Harvard University Press.
- 1982. Worlds away. Reimpreso en *Theories and Things*, ensayo 14, pp. 124 – 128.

ROBERTSON, Teresa.

- 1998. Possibilities and the arguments for origin essentialism. *Mind* **107**, #428, pp. 729 – 749.
- 2000. Essentialism: Origin and order. *Mind* **108**, #434, 299 – 307.
- 2007. Essential vs. accidental properties. *Stanford Encyclopedia of Philosophy (Fall 2008 Edition)*, Edward N. Zalta, (ed). Disponible en:

<http://plato.stanford.edu/archives/fall2008/entries/essential-accidental>
(accesado en Octubre 2009)

ROBERTSON, Teresa & FORBES, Graeme.

- 2006. Does the new route reach its destination?. *Mind* **115**, #458, pp. 367 – 373.

ROHRBAUGH, Guy & DEROSSET, Louis.

- 2004. A new route to the necessity of origin. *Mind* **113**, #452, pp. 705 – 725.
- 2006. Prevention, independence, and origin. *Mind* **115**, #458, pp. 376 – 385.

RUSSELL, Bertrand.

- 2005. On denoting. *Mind* **14**, #56, pp. 479 – 493.

SALMON, Nathan.

- 1979. How *not* to derive essentialism from the theory of reference. *The Journal of Philosophy* **76**, #12, pp. 703 – 725.
- 1987. Existence. *Philosophical Perspectives* **1**, pp. 49 – 108. Reeditado en SALMON 2005, pp. 9 – 49.
- 1989. The logic of what might have been. *The philosophical review* **98**, #1, pp. 3 – 34. Reeditado en SALMON 2005, pp. 129 – 149.
- 1996. Transworld identification and stipulation. *Philosophical Studies* **84**: 203 – 223. Reeditado en SALMON 2005b, como Apéndice V: Crossworld identification and stipulation.
- 2005. *Metaphysics, mathematics and meaning: Philosophical papers, Volume I*. Clarendon Press, Oxford.
- 2005b. *Reference and essence*. 2a edición. Prometheus Books.

STALNAKER, Robert.

- 1979. Anti-essentialism. Reeditado en STALNAKER 2003, pp. 71 – 83.
- 2003. *Ways a World Might Be*. Oxford.

VAN INWAGEN, Peter.

- Por aparecer. McGinn on existence. Por aparecer en *The Philosophical Quarterly*.
- 1985. Plantinga on trans-world identity. Reeditado en VAN INWAGEN 2001, pp. 186 – 205.

- 2001. *Ontology, identity and modality. Essays in metaphysics.* Cambridge University Press.

WILLIAMSON, Timothy.

- 2006. Stalnaker on the interaction of modality with quantification and identity. En THOMPSON & BYRNE (editores) *Content and modality. Themes from the philosophy of Robert Stalnaker.* Oxford University Press.

YABLO, Stephen.

- 1987. Identity, essence and indiscernibility. *Journal of Philosophy* **84** #6, pp. 293 – 314.