UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO

FACULTAD DE FILOSOFÍA Y LETRAS

Sobre los mundos que pudieron haber sido: una defensa del modelo semántico Stalnaker-Lewis para condicionales contrafácticos

TESIS

Que para obtener el grado de

Licenciado en Filosofía

PRESENTA

Luis Rodolfo Vázquez Cabrera.

DIRECTOR DE TESIS

Mtro. Gabriel Ramos García.

Ciudad Universitaria, Ciudad de México, 2018.

Universidad Nacional Autónoma de México
UNAM
Biblioteca Central
Dirección General de Bibliotecas de la UNAM

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.

ÍNDICE

Introducción

Capítulo I: El problema de las condiciones de verdad de los condicionales contrafácticos

1.1 Las afirmaciones condicionales del lenguaje natural y las condiciones de verdad del condicional material

1.2 Las condiciones de verdad de los condicionales contrafácticos

1.3 Alternativas para un análisis de los enunciados condicionales

Capítulo II: La semántica Stalnaker-Lewis para condicionales contrafácticos

2.1 La lógica condicional de Robert Stalnaker

2.2 La Teoría de contrafácticos de David Lewis

2.3 Problemas en Stalnaker y las ventajas de la reformulación de Lewis

2.4 El problema de la invalidez de la transitividad contrafáctica

Capítulo III: Análisis de una alternativa a los mundos posibles

3.1 Cuestionando el modelo Stalnaker-Lewis

3.2 Las objeciones de Michael Woods a la semántica de mundos posibles para condicionales contrafácticos

3.3 La teoría de Michael Woods de condicionales contrafácticos

3.4 Crítica a las tres tesis de Michael Woods

Conclusiones

Bibliografía

Introducción

Una de las cosas que uno debe tener en cuenta cuando se introduce en el estudio de la lógica formal es que se trata de una herramienta con alcances limitados. Muchas características del lenguaje natural[1] se pierden en la traducción al lenguaje lógico, por lo que hay muchos elementos de los razonamientos y procesos inferenciales que comúnmente hacemos en contextos reales que no son capturados por la maquinaria de la lógica clásica. Pero eso no necesariamente se trata de un defecto como muchos podrían pensar sino más bien de una virtud. El hecho de que sea un sistema que abarque tan poco es lo que le permite ser también un sistema eficiente y funcional, la abstracción de esos elementos del lenguaje que no son estructurales es lo que le permite alcanzar un entendimiento claro de nociones tales como la implicación y la validez. Por lo que uno como filósofo debe aprender a vivir con la idea de que habrá cosas que queden fuera del alcance de la lógica clásica. Sin embargo, en respuesta a limitaciones como esas, también es posible expandir el sistema de la lógica clásica para poder abarcar más fenómenos del lenguaje y del razonamiento, en otras palabras, la insuficiencia de la lógica clásica formal para abarcar toda la riqueza del lenguaje es la que motiva a buscar lógicas alternativas que expliquen otros fenómenos. Esto es

[1] En el transcurso de esta tesis usaré la frase "lenguaje natural" para referirme al lenguaje nativo en el que un hablante se comunica (como el inglés o el español por ejemplo). Cuando use la frase generalmente la usaré para resaltar su carácter preteórico, por ejemplo si hago mención de un condicional del lenguaje natural me estoy refiriendo a una afirmación condicional que un hablante competente de su lengua enunciaría de manera cotidiana independiente del análisis que un teórico de la lógica pudiera hacer acerca de ese condicional, incluso si está inserto en una estructura inferencial. Así mismo cuando hable de contextos de habla o enunciación normales me refiero a los contextos en los que ese mismo hablante haría una afirmación, es decir a los contextos de uso cotidiano del hablante que emplea la lengua para comunicarse con otros hablantes. Esta idea es opuesta a la noción de "lenguaje formal" por la cual yo entiendo el lenguaje de las teorías lógicas que (entre otras cosas) abstraen elementos del lenguaje natural y los representan mediante un aparato técnico construido artificialmente, por lo que en ese sentido considero también a los lenguajes formales "lenguajes artificiales".

precisamente lo que sucede en el caso de los condicionales contrafácticos. Los condicionales contrafácticos son afirmaciones condicionales en las cuales implícitamente se presupone la falsedad de la cláusula condicional, como es entendida en lógica clásica. Esta suposición implícita de falsedad choca con las propiedades semánticas que la lógica clásica le adjudica a las afirmaciones condicionales; esto sucede porque en el análisis de la lógica clásica todo condicional con antecedente falso es verdadero, lo cual implica que todo condicional contrafáctico es verdadero bajo ese análisis, entrando claramente en conflicto con el modo en que se usan estos condicionales en contextos de habla normales. Dada esta inadecuación del análisis clásico con las propiedades semánticas de los contrafácticos es que se vuelve necesario un análisis alternativo que explique las circunstancias en las que estos son verdaderos, es decir que explique sus condiciones de verdad. En esta tesis mi propósito central es justamente responder a la pregunta de cuáles son las condiciones de verdad de los condicionales contrfácticos, concretamente lo que me propongo hacer una exposición y una evaluación del modelo semántico Stalnkaer-Lewis (que se compone de las lógicas condicionales de los filósofos Robert Stalnaker y David Lewis) el cual intenta establecer un planteamiento teórico para dar cuenta precisamente de las condiciones de verdad de los condicionales contrafácticos. En otras palabras, lo que pretendo es revisar las teorías de Stalnaker y de Lewis, señalar sus dificultades y cómo pueden ser salvadas; así como también examinarlas a la luz de una teoría alternativa que cuestione sus supuestos para finalmente evaluar también la teoría alternativa y dar un veredicto de si el modelo Stalnaker-Lewis se sostiene aún después de ese examen. Este modelo es el núcleo del trabajo porque su planteamiento basado en la semántica de mundos posibles es una de las explicaciones más mencionadas en el campo de las lógicas condicionales. Después de revisar la los textos de la bibliografía encontré que los textos constantemente se referían al modelo

Stalnaker-Lewis como el punto de partida de la manera en que se entienden los condicionales contrafácticos en la actualidad por lo que me pareció que valía la pena hacer una investigación y un examen de esa teoría. En las páginas que siguen haré una exposición del modelo semántico en cuestión para luego ponerlo a prueba señalando sus problemas y evaluando un modelo alternativo, para así finalmente dictaminar si las condiciones de verdad propuestas en esta teoría dan cuenta satisfactoriamente de la semántica de los condicionales contrafácticos. El trabajo tendrá entonces una estructura de tres momentos: i) planteamiento del problema, ii) exposición del modelo semántico Stalnaker-Lewis y iii) examinación crítica del modelo. Cada uno de estos momentos comprenderá un capítulo de la tesis. En el Capítulo I voy a delimitar el objeto de estudio dando una breve taxonomía de los enunciados condicionales, posteriormente voy a exponer de manera más detallada el problema de las condiciones de verdad de los condicionales en general y luego de los condicionales contrafácticos, ya que el problema de los condicionales contrafácticos es un caso particular del problema de los condicionales, todo esto en el contexto de la inadecuación del análisis clásico de condicionales para explicar las propiedades semánticas de estos enunciados. En el Capítulo II expondré las teorías de Stalnaker y de Lewis, señalando algunos de los problemas que estas presentan, tales como el problema del supuesto de unicidad, el problema del principio de tercero excluso condicional y el problema de la invalidez de la transitividad contrafáctica. Finalmente, en el Capítulo III exploraré una teoría alternativa al modelo Stalnaker-Lewis que intenta también dar una explicación de las condiciones de verdad de los condicionales contrafácticos. La teoría en cuestión es la del filósofo Michael Woods. Elegí este filósofo porque su teoría rechaza de manera frontal el modelo de mundos posibles propuesto por Stalnaker y Lewis, lo cual me permitirá someter a examen este modelo. Por lo que posteriormente pondré a prueba el modelo Stalnaker-Lewis

siguiendo a Woods y su cuestionamiento de la viabilidad de la semántica de mundos posibles para entender los contrafácticos, así mismo evaluaré las objeciones de Woods para determinar si efectivamente tiran o no el modelo Stalnaker-Lewis; consideraré también las ventajas de la teoría de Woods en contraste con las ventajas de las teorías basadas en mundos posibles que revisaremos para finalmente dar un veredicto acerca de cuál de ellas es una mejor teoría de los condicionales contrafácticos. Por último en las conclusiones trataré de dar una perspectiva general de los hallazgos de esta investigación y haré una evaluación de si he logrado responder a mi pregunta central.

CAPÍTULO I

EL PROBLEMA DE LAS CONDICIONES DE VERDAD DE LOS CONDICIONALES CONTRAFÁCTICOS

1.1 Las afirmaciones condicionales del lenguaje natural y las condiciones de verdad del condicional material

Una relación fundamental en lógica clásica es la relación que llamamos condicional. Siendo una relación binaria, podemos decir que dos oraciones P y Q se encuentran en una relación condicional cuando la verdad de P es condición para la verdad de Q o viceversa. En el lenguaje natural esta relación se expresa generalmente mediante oraciones con estructura 'si… entonces' por ejemplo la oración 'si entrenamos duro entonces ganamos la olimpiada', en consecuencia, a estas oraciones les llamamos oraciones condicionales. De entre los tipos de oraciones condicionales del lenguaje natural, las oraciones que a la lógica le competen son las oraciones condicionales asertóricas, pues dado que son oraciones declarativas que afirman la verdad de un hecho, ellas son susceptibles de ser evaluadas como verdaderas o falsas, es decir son portadoras de verdad. En contraste, quedan fuera (al menos de la lógica clásica) oraciones tales como las oraciones condicionales interrogativas ('si no encuentro las llaves, entonces ¿cómo entraré a la casa?') o las oraciones condicionales imperativas ('si te reúnes con Javier, salúdalo de mi parte'), las cuales no son portadoras de verdad.

Por su parte los enunciados condicionales asertóricos pueden ser divididos en dos tipos distintos: los condicionales indicativos y los condicionales subjuntivos o contrafácticos. De acuerdo con Michael Woods[2], los condicionales indicativos son

[2] Woods, Michael. Conditionals. Oxford University Press. Oxford. 1997. P.10

aquellos que involucran la aplicación de la estructura ‘si… entonces’ a un par cualquiera de enunciados asertóricos y cuyo significado puede ser explicado dando cuenta de la interacción entre el significado de dicha estructura y el significado de las dos oraciones que la completan. Por ejemplo, el significado del condicional indicativo ‘si esta nevando, entonces es invierno’ se compone del significado de la estructura condicional y de los dos enunciados que ocupan los lugares vacios de dicha estructura, es decir, el enunciado ‘esta nevando’ y el enunciado ‘es invierno’. Es necesario para que el enunciado condicional sea significativo que los enunciados que ocupen la estructura sean ellos mismos enunciados declarativos, de manera que enuncien un hecho sobre el mundo y que puedan ser susceptibles de verdad o falsedad, en otras palabras, que los enunciados involucrados en el condicional sean también portadores de verdad, para que el enunciado condicional sea también susceptible de ser evaluado como verdadero o falso. De ahí que elementos del discurso tales como las preguntas o las ordenes al ser insertadas en una estructura condicional del tipo “si… entonces” no producen un enunciado susceptible de verdad o falsedad pues ni las preguntas ni las órdenes son portadoras de verdad dado que no afirman ni niegan hechos.

Los condicionales contrafácticos por otro lado son aquellos que además de consistir en un par de enunciados insertos en una estructura del tipo “si… entonces” se distinguen desde el punto de vista estructural por involucrar expresiones tales como ‘habría’ y ‘hubiera’, por lo que, siguiendo a Woods[3], su significado estará dado por (i) las oraciones que lo componen, (ii) la estructura condicional ‘si… entonces’ como se presenta en los condicionales indicativos y (iii) por las propiedades semánticas de sus expresiones características. Un enunciado

[3] Woods, Michael. Conditionals. Oxford University Press. Oxford. 1997. P.10

contrafactual típico es el enunciado ‘si Mario hubiera cruzado la calle con cuidado entonces no lo habrían atropellado’, en un enunciado como éste (a diferencia del condicional indicativo) las oraciones que intervienen en la estructura condicional no son enunciados que puedan ser evaluados como verdaderos o falsos de manera independiente a esa esctructura condicional, es decir que su significado está ligado a la estructura semántica del condicional contrafactual completo que, como dijimos, además del ‘si… entonces’ entraña el uso de ‘hubiera’ y ‘habría’. Adicionalmente, un rasgo prominente atribuido al condicional contrafáctico es la presuposición de que el antecedente refiere a un hecho que no sucedió o no es el caso en el momento de la locución del condicional, por lo que el condicional contrafactual pareciera estar mentando algo que es verdad dada la verdad hipotética de una afirmación que es de hecho falsa.

Veamos ahora el tratamiento que se le dan a las oraciones condicionales en la lógica clásica. En lógica clásica el símbolo ‘→’ nos sirve para expresar la relación condicional entre dos proposiciones, así como para traducir oraciones con la estructura ‘si... entonces’ del lenguaje natural al lenguaje formal, de manera que un enunciado como ‘Si hoy cobro mi salario entonces lo gastaré hoy mismo en ropa’ se expresaría en el lenguaje lógico con la fórmula ‘P→Q’ donde ‘P’ simboliza ‘hoy cobro mi salario’ y ‘Q’ simboliza ‘gastaré mi salario hoy mismo en ropa’. En esta formulación a la oración que ocupa el lugar de ‘P’ se le llama antecedente y a la oración que ocupa el lugar de ‘Q’ se le llama consecuente.

Tanto en el lenguaje formal como en el lenguaje natural, la manera en que sabemos cuando una afirmación es verdadera o falsa es mediante sus condiciones de verdad, es decir las circunstancias que deben darse para que una afirmación sea verdadera o falsa. En lógica clásica el análisis semántico de las condiciones de verdad

de toda construcción condicional es un análisis veritativo funcional, esto significa que un enunciado condicional es verdadero o falso dependiendo del valor de verdad de los enunciados simples que lo componen. Dado que la lógica clásica es una lógica bivalente hay cuatro posibles combinaciones de valores de verdad para los enunciados moleculares, cada una de las cuales arroja un valor de verdad para el enunciado completo 'P→Q'.

Una manera sencilla de expresar las condiciones de verdad es mediante una tabla de verdad que muestre las 4 posibles asignaciones de verdad de una condicional: i) antecedente y consecuente verdaderos, ii) antecedente verdadero y consecuente falso, iii) antecedente falso y consecuente verdadero y iv) antecedente y consecuente falsos.

P	→	Q
V	V	V
V	F	F
F	V	V
F	V	F

Obsérvese que en el análisis tradicional, el único caso de falsedad es cuando el antecedente es verdadero y el consecuente falso, todos los otros casos arrojan un valor verdadero para el compuesto. Este análisis expresa las condiciones de verdad de lo que se conoce en lenguaje técnico como el *condicional material*. Como veremos a continuación el condicional material no siempre describe adecuadamente las

propiedades semánticas de las oraciones condicionales del lenguaje natural, es decir que las condiciones de verdad de las oraciones condicionales del lenguaje natural no siempre coinciden con las condiciones que la lógica clásica le atribuye a enunciados de este tipo. Las condiciones de verdad de la fórmula ‘P→Q’ garantizan que una afirmación de tipo condicional en el lenguaje de la lógica clásica es equivalente a una disyunción de la forma ‘¬P ∨ Q’ (donde el símbolo ‘∨’ expresa una relación de disyunción entre dos afirmaciones) que se leería como “o bien es verdadera ¬P o bien es verdadera Q o son verdaderas ambas” y por ende es verdadera en caso de que ‘P’ sea falsa y/o ‘Q’ sea verdadera. Sin embargo no podríamos afirmar con seguridad que las oraciones del lenguaje natural que expresan relaciones condicionales expresan una disyunción entre la falsedad del antecedente y la verdad del consecuente, sino que, en cambio, afirmaríamos que expresan la verdad del consecuente dada la verdad (hipotética o factual) del antecedente. Retomando la oración condicional "Si hoy cobro mi salario entonces lo gastaré hoy mismo en ropa" desde el punto de vista de la lógica clásica un posible caso de verdad del enunciado sería que el antecedente fuera falso y el consecuente verdadero de manera que aún cuando es falso que cobro hoy mi salario de todos modos lo gasto hoy mismo en ropa, pero ¿cómo puedo gastarlo si no lo he cobrado todavía?

Quisiera aventurar la hipótesis de que la intención detrás del funcionamiento de esta conectiva es modelar la noción de inferencia deductiva, es decir tener un constructo el cual exprese que la verdad del consecuente es garantizada por la verdad del antecedente. Es por ello que cuando en un condicional material el consecuente es falso, siendo el antecedente verdadero, el compuesto resulta falso también. Esto sucede porque es en este caso y sólo en éste en el cual no se cumple la garantía de que el antecedente herede la verdad al consecuente, lo cual es requerimiento obligado en una inferencia deductiva. Sin embargo al tratar de adecuar este constructo

semántico al uso de otras expresiones condicionales del lenguaje natural puede resultar contraintuitivo pensar que un condicional como "si estudio paso el examen" es verdadero no sólo cuando estudio y paso el examen sino también cuando ni estudio ni paso el examen e incluso cuando no estudio y sí paso el examen. Como veremos también más adelante, cuando se afirma un condicional en un contexto de habla normal usualmente existe la creencia hipotética en el antecedente además de que se presupone algún tipo de conexión entre antecedente y consecuente, ya sea causal, temporal o de otra índole. Es por eso que resulta tan extraño que un condicional material sea verdadero cuando la condición expresada en el antecedente sea falsa.

Considero yo que la lógica (entendiéndola en un sentido amplio como campo del conocimiento, no solamente reduciéndola a la lógica clásica) es una disciplina que tiene como una de sus tareas modelar el razonamiento correcto y que debe podernos dar cuenta de los razonamientos y estructuras inferenciales válidas que hacemos cotidianamente los seres humanos sin conocimiento teórico de la lógica, así como también de otros razonamientos de los que no hacemos cotidianamente (como los que hacen los matemáticos o las maquinas). En el caso del condicional material de la lógica clásica, es evidente que no captura el uso natural de todo enunciado condicional (aunque de algunos enunciados condicionales sí) y eso nos puede resultar problemático cuando intentamos aplicar el modelo formal de la lógica clásica para el uso que consideramos correcto de muchos enunciados y razonamientos del lenguaje natural; parece en cambio, que el condicional material únicamente garantiza que no suceda que el antecedente sea verdadero al mismo tiempo que el consecuente falso, sus condiciones de verdad son pues condiciones restrictivas: la estructura veritativo funcional condicional de la lógica clásica evita que un condicional sea verdadero cuando es el caso que el antecedente sea verdadero y el consecuente falso. De manera que el condicional material funciona más bien como una paráfrasis de "en este

momento no sucede al mismo tiempo que el antecedente sea verdadero y el consecuente falso" por lo que no puede usarse para entender muchos de los condicionales del lenguaje natural que no comparten esas condiciones de verdad como los que involucran contenido implícito como temporalidad, causalidad, etc. A diferencia del lenguaje de la lógica formal clásica, en el lenguaje natural el uso típico de una oración condicional entraña una asunción acerca de la verdad del antecedente para afirmar, en consecuencia de esta asunción, la verdad del consecuente, por lo que el condicional material usado en el lenguaje de la lógica formal clásica no alcanza a capturar el tipo de relación que muchos enunciados condicionales suelen expresar en contextos de habla no formales. Que una fórmula condicional pueda ser verdadera cuando el antecedente es falso va en contra de la manera en que funcionan muchas de las oraciones condicionales en contextos no formales. De los anterior concluyo que no basta con la lógica clásica para lograr los objetivos que al inicio de este parrafo planteo como objetivos de la lógica como disciplina. Para ello se requieren otras herramientas adicionales a la lógica clásica.

1.2 Las condiciones de verdad de los condicionales contrafácticos

En el caso de las afirmaciones condicionales contrafácticas es aún más claro que no pueden ser evaluadas como son evaluadas las fórmulas del lenguaje lógico formal, y que la creencia hipotética en la verdad del antecedente es lo único que permite que estas afirmaciones puedan ser juzgadas como verdaderas o falsas, pues ellas involucran antecedentes que son de hecho falsos. Consideremos el enunciado "Si hubiera salido temprano de mi casa, no habría llegado tan tarde al trabajo”, en esta oración pareciera que esta implícita la suposición de que de hecho no salí temprano de mi casa por lo que para evaluar su verdad o falsedad es necesario asumir lo

contrario, a saber, que sí salí temprano. A partir de esa asunción es como puedo considerar la verdad o falsedad del consecuente de este tipo enunciado condicional. Veremos a continuación, de manera más detallada, cómo el análisis tradicional para condicionales de la lógica clásica es insuficiente para determinar las condiciones de verdad de un enunciado contrafactual.

Los condicionales contrafácticos son enunciados condicionales del lenguaje natural que cuando son correctamente empleados presuponen la falsedad del antecedente. Cuando digo "si hubiera nacido en Francia me sabría la marsellesa de memoria" quiero decir que aun cuando nací en México y no en Francia, si los hechos fueran diferentes y hubiera nacido en Francia entonces sería verdad que me sabría la marsellesa de memoria. Lo mismo cuando afirmo que “si México hubiera ganado el partido contra Holanda en el mundial pasado, sería campeón del mundo” o también “si hubiera estudiado astronomía mi tesis sería acerca de los anillos de Saturno” estoy afirmando que si algo que de hecho no sucedió hubiera sido verdadero entonces podría afirmar la verdad de otro hecho diferente. En un condicional contrafáctico se parte de una hipótesis alternativa a lo que de hecho es el caso y se hace una declaración acerca de cómo sería el mundo si la hipótesis fuera verdadera. Un condicional contra fáctico es pues un condicional contra los hechos.

Dadas las condiciones de verdad del condicional material de la lógica clásica hay dos problemas evidentes que se presentan cuando se intenta representar con estos las oraciones condicionales contrafácticas, el primero es que, como habíamos señalado, al proferir un condicional contrafactual existe implícita la presuposición de la falsedad del antecedente, si asumimos como dada esta presuposición de que todos los contrafácticos tienen antecedente falso, al adecuarlos a la estructura veritativo funcional del condicional material nos veríamos obligados a afirmar que toda

construcción condicional contra factual es verdadera, lo cual es evidentemente falso si comparamos los siguientes dos enunciados condicionales, el primero un condicional indicativo regular y el segundo un contra fáctico:

i) Si Leonardo no pintó la mona lisa entonces alguien más lo hizo

ii) Si Leonardo no hubiera pintado la mona lisa entonces alguien más lo habría hecho.

Es claro que, dado que alguien de hecho pinto la mona lisa, el primer condicional es verdadero mientras que el segundo es muy difícil que pudiera resultar verdadero también.

El segundo problema tiene que ver con el modo en el que son expresados los condicionales contra factuales, que es el modo subjuntivo, que es evidente por la inclusión de la partícula "hubiera", en contraste con el modo indicativo del condicional material. Esto es, el condicional material hace una declaración de un estado de cosas actual, está afirmando algo que en este momento de hecho es o no el caso, el contra fáctico por el contrario afirma el antecedente de manera hipotética, se podría decir que hasta imaginaria, contenido semántico que queda fuera del análisis tradicional de condicionales de la lógica clásica. Es por esto que las condiciones de verdad de los condicionales contra fácticos no pueden ser las mismas del condicional material.

Pareciera entonces que, en el caso de los condicionales contrafácticos es aún más patente la inadecuación del análisis de la lógica clásica a las características semánticas de los condicionales en el lenguaje natural. Y es aún más problemático puesto que parece que una característica central de los condicionales contrafácticos es la falsedad del antecedente, lo que haría que bajo el análisis de la lógica clásica todo

condicional contrafáctico fuera verdadero, lo cual es muy difícil de aceptar. Ante este escenario es necesario entonces proponer un análisis semántico alternativo que pueda dar cuenta de mejor manera de las condiciones de verdad de los condicionales contrafácticos. Es por esto que la pregunta central de esta tesis será precisamente la de en qué circunstancias un condicional contrafáctico es verdadero o falso. Es decir, la pregunta de cuáles son las condiciones de verdad de los condicionales contrafácticos. Para responder esa pregunta, analizaré en el capítulo II de esta tesis, dos teorías alternativas al análisis de la lógica clásica que intentan formular las condiciones de verdad de este tipo de enunciados en otros términos, para después en el capítulo III someterla a prueba y evaluar su viabilidad. La observación acerca del antecedente de un contrafáctico como hipótesis imaginaria es lo que dará pie a estas teorías que serán el núcleo de la tesis: las teorías semánticas de Robert Stalnaker y David Lewis de enunciados contra fácticos, las cuales van a usar el aparato semántico de mundos posibles, para capturar la idea de hechos que no son el caso pero que podrían serlo o pudieron haberlo sido, llamaré a estas dos teorías en conjunto como el modelo semántico Stalnaker-Lewis para condicionales contrafácticos.

1.3 Alternativas para un análisis de los enunciados condicionales

A manera de preámbulo a la discusión del modelo Stalnaker-Lewis revisaré brevemente la discusión acerca de las condiciones de verdad de los enunciados condicionales en general, para luego centrarme exclusivamente en los condicionales contrafácticos. Conectaré este último apartado del capítulo I con la siguiente sección, usando como puente la idea del test de Ramsey para los condicionales, el cual juega un rol central en el planteamiento del modelo Stalnaker-Lewis.

En lo que concierne a los condicionales del lenguaje natural hay varias posibles alternativas de análisis para sus condiciones de verdad. Una primera alternativa es que los condicionales siempre tienen valor de verdad y que además son veritativo funcionales, lo cual quiere decir que su valor depende del valor de verdad de las oraciones que lo componen sin necesidad de que estas estén relacionadas semánticamente, como en el caso del enunciado 'si la luna es de queso la capital de Francia es Paris' cuya verdad no depende del contexto de enunciación ni de la relación entre las dos oraciones simples, sino sólo del resultado que arroje la función proposicional "si… entonces" al ser ocupada por las oraciones simples. De modo que como son veritativo funcionales y siempre tienen valor de verdad son reductibles a condicionales materiales como los de la lógica proposicional, y por ende son verdaderos para cualquier combinación de valores de verdad excepto en el caso en el que el antecedente es verdadero y el consecuente falso. Lo anterior tiene como corolario que un condicional con el antecedente falso será siempre verdadero, lo que parece ir en contra de nuestras intuiciones acerca de cuándo un enunciado de este tipo es significativo, por ejemplo en el caso de un enunciado como "Si 3+5=7 entonces lloverá mañana" . Aún cuando este es verdadero dado que su antecedente es falso, pareciera que no podría ser declarado a menos que fuera en circunstancias muy especiales, pues parece no haber conexión semántica o de ningún otro tipo entre los enunciados que componen la afirmación condicional. Sin embargo, en el caso en que se redujeran los condicionales simples a condicionales materiales, la dificultades que esto presenta podrían ser salvadas -sostiene Michael Woods[4]- si además de sus condiciones de verdad distinguiéramos sus condiciones de asertibilidad (es decir las circunstancias en las cuales un enunciado puede ser declarado) y estableciéramos que un condicional simple puede ser verdadero y sin embargo no ser asertible. Por

[4] Woods, Michael. Conditionals. Oxford University Press. Oxford. 1997. P.12

ejemplo, el enunciado “si 3+7=11 entonces la tierra gira alrededor del sol” es verdadero en virtud de que es una función de verdad y en virtud tanto de la verdad del consecuente como de la falsedad del antecedente, y sin embargo no es asertible en el lenguaje natural puesto que carece de una conexión semántica entre las oraciones simples y no establece relaciones de causalidad ni temporalidad que son las que más comúnmente expresan los condicionales en el lenguaje natural. De este modo podríamos entender los condicionales del lenguaje natural como condicionales materiales con la salvedad de que si bien un condicional puede ser formalmente verdadero, puede no ser asertible en el lenguaje natural. Según lo anterior, un análisis de las condiciones de verdad de los condicionales deberá también incluir un análisis de sus condiciones de asertibilidad, además de una explicación de la discordancia entre las condiciones de asertibilidad de un enunciado en el lenguaje natural y su correspondiente condicional material.

Por otro lado, una alternativa diferente es que los condicionales simples no siempre tienen valor de verdad, es decir que pueden ser o bien verdaderos, o bien falsos, o bien carecer de valor de verdad. En este caso un condicional simple sería verdadero si tanto el antecedente como el consecuente son verdaderos, falso si el antecedente es verdadero y el consecuente falso o bien carente de valor de verdad en caso de que el antecedente fuera falso (sin importar el valor de verdad del consecuente). El análisis anterior previene la posibilidad de que un enunciado como “Si 3+5=7 entonces lloverá mañana” sea verdadero únicamente en virtud de la falsedad de su antecedente, como sí sucede en la lógica proposicional.

Una de las explicaciones más intuitivas de la semántica de las oraciones condicionales es la que ofrece F. P. Ramsey en su ensayo “Proposiciones generales y

causalidad”[5]. Ramsey sostiene que al hacer la aserción “Si P entonces Q” se afirma que Q es verdadera bajo el supuesto de que P también lo es. De modo que si dos personas discuten “¿si fuera el caso que P, sucedería Q?” lo que hacen es añadir hipotéticamente P a su cuerpo de creencias y discutir si sobre esa base Q resultaría verdadera. Lo que tiene como consecuencia que si P fuera falsa la discusión cesaría de tener significado. Así entonces, un condicional es verdadero según Ramsey si añadiendo el antecedente a mi cuerpo de creencias el consecuente resulta verdadero a consecuencia de esa adición. Esto se conoce como la prueba de Ramsey para la aceptabilidad de los condicionales. Un caso en que la prueba de Ramsey se encuentra con dificultades y que es relevante para la discusión sobre los condicionales contrafactuales es el que se presenta cuando se sabe de cierto (o al menos se presume) la falsedad del antecedente. Si procediera a añadir hipotéticamente la verdad del antecedente a mi conjunto de creencias entre las cuales ya se encuentra la creencia de que la oración en el antecedente es falsa ello entrañaría hacer una revisión no solo de mi creencia de que esa oración es falsa, sino de mis demás creencias también, pues en la prueba de Ramsey la aceptabilidad del consecuente se da como consecuencia no solo del antecedente sino de su incorporación a mis otras creencias. En este caso la aceptabilidad de la afirmación condicional se da sólo si la revisión a mis creencias es una revisión *mínima*, es decir si no conlleva a una revisión radical de todo mi cuerpo de creencias. Este procedimiento de revisión de creencias puede ser visto como la consideración de un mundo alterno posible al que está representado por mi cuerpo de creencias. Esta línea de pensamiento conduce a las teorías de mundos posibles para condicionales, que revisaré en el siguiente capítulo.

[5] F.P. Ramsey, “General propositions and Causality” en: Ramsey, Foundations of Mathematics and other Logical Essays. New York, 1950. Pp. 237-257.

CAPÍTULO II

LA SEMÁNTICA STALNAKER-LEWIS PARA CONDICIONALES CONTRAFÁCTICOS

2.1 La lógica condicional de Robert Stalnaker

En "A Theory of Conditionals"[6] Robert C. Stalnaker construye una lógica condicional con miras a establecer las condiciones de verdad de toda construcción condicional del lenguaje natural. Stalnaker sostiene que los enunciados condicionales expresan una proposición que es una función de otras dos proposiciones, es decir, que un enunciado condicional es un enunciado molecular que está compuesto por dos enunciados atómicos o moleculares, insertados ambos en una estructura del tipo 'si… entonces'. Sin embargo no todo condicional del lenguaje natural es una función de verdad, lo que significa que no todo condicional cumple que su valor de verdad esté determinado por el valor de verdad de los enunciados que lo componen. Por ejemplo, es posible que yo sepa el valor de verdad de 'Armando estudió derecho' y de 'Armando se recibió de abogado' y no obstante no saber el valor de verdad de 'si Armando hubiera estudiado derecho se habría recibido de abogado'. En estos casos en que el valor de verdad de un condicional no depende del valor de verdad de los enunciados que lo componen es necesaria una teoría que explique cuáles son entonces sus condiciones de verdad.

Stalnaker intenta formular una teoría sobre los condicionales que incluye a los condicionales no veritativo funcionales atacando lo que él llama el *problema lógico de los condicionales* que consiste en describir cuáles son las propiedades formales de la función condicional "si,… entonces" lo cual implica determinar las condiciones de

[6] Robert C. Stalnaker, 'A Theory of Conditionals', en: Nicholas Rescher (ed.), Studies in Logical Theory, Oxford 1968.

verdad de los enunciados condicionales. Sin embargo Stalnaker aborda el problema preguntándose primero cómo de hecho evaluamos una afirmación condicional, es decir cómo uno decide la verdad o falsedad de una oración condicional cuando se enfrenta a ella. Tomemos por ejemplo la oración 'si le regalo una flor a mi novia, ella se casará conmigo' ¿cómo uno decide si considerar verdadera o falsa una afirmación como esa? Stalnaker ofrece tres respuestas a esa pregunta de las cuales la última evita algunas de las dificultades de las otras dos.

La primera respuesta es lo que se conoce como el análisis veritativo funcional. Según este análisis, de las cuatro posibles asignaciones de verdad para una afirmación de la forma "si,… entonces" sólo una de ellas (antecedente verdadero y consecuente falso) da como resultado la falsedad del enunciado condicional completo, de manera que si el antecedente es falso o el consecuente es verdadero, el enunciado condicional será necesariamente verdadero. Pero esta postura es inaceptable dado que la falsedad del antecedente no es razón suficiente para la afirmación del condicional. Una segunda respuesta tiene que ver con las limitaciones de la primera. El análisis veritativo funcional de los condicionales no es satisfactorio porque no incluye la idea de *conexión* entre las proposiciones que componen el condicional y que está implícita en algunas afirmaciones de este tipo. Si se tomara en cuenta esta idea de conexión, podríamos entender las afirmaciones condicionales como declaraciones que establecen una relación que se da entre el antecedente y el consecuente (lógica, causal, temporal, etcétera), y por ende su evaluación dependería de si de hecho hay tal conexión o no. Lo anterior implicaría por supuesto precisar lo que se entiende por esta idea de conexión, sin embargo esta estrategia de evaluación de condicionales también falla debido a que un condicional puede ser verdadero aun cuando no exista conexión alguna entre antecedente y consecuente. Consideremos el ejemplo propuesto anteriormente y supongamos que el consecuente es verdadero

independientemente de su relación con el antecedente, es decir, supongamos que mi novia se casará conmigo le regale una flor o no, podemos suponer que ella siempre ha estado muy enamorada de mí o que quiere casarse conmigo porque soy millonario o cualquier otra cosa que haga verdadero el consecuente, de manera que el hecho de que yo le regale una flor no hará realmente una diferencia. En ese caso es claro que nuestra afirmación condicional inicial es verdadera aun cuando no haya relación entre antecedente y consecuente.

La tercera respuesta está relacionada con el test de Ramsey que habíamos visto previamente. Y es esta alternativa la que Stalnaker usa para construir su lógica condicional. La lógica condicional de Stalnaker toma como punto de partida la observación hecha por F.P. Ramsey[7] acerca de cómo, en condiciones de habla normales, se evalúa un enunciado condicional. Si nos preguntamos por la verdad o falsedad de un enunciado condicional, señalaba Ramsey, lo que hacemos en esencia es suponer la verdad del antecedente añadiéndolo hipotéticamente al conjunto de todas nuestras otras creencias acerca del mundo y a partir de ahí juzgar si el consecuente es verdadero o falso como resultado de esta adición. El condicional completo deberá tener el mismo valor de verdad que el consecuente. Aquí la idea de conexión es a veces relevante a la evaluación del condicional y a veces no. Si se considera que hay una conexión entre antecedente y consecuente, entonces junto con el antecedente se debe también añadir el consecuente, pues éste es consecuencia de aquel. Y si no existe dicha conexión y el consecuente es verdadero independientemente de su relación con el antecedente (y además creo que lo es) entonces al añadir el antecedente a mi conjunto de creencias, el consecuente seguirá perteneciendo a ese conjunto. En los dos casos se debe afirmar la verdad del

[7] F.P. Ramsey, "General propositions and Causality" en: Ramsey, Foundations of Mathematics and other Logical Essays. New York, 1950.

condicional. De modo que bajo este análisis la idea de conexión no es indispensable para la evaluación de un condicional. Ramsey asume que uno no tiene conocimiento alguno sobre la verdad o falsedad del antecedente cuando hacemos tal suposición, no obstante se puede aplicar tanto para el caso en que se sabe que el antecedente es verdadero como para el caso en que se sabe que es falso. Si el antecedente es verdadero, el valor de verdad del condicional será simplemente el mismo que el del consecuente. El otro caso sin embargo requiere más atención, pues si al evaluar un condicional sé que el antecedente es falso, incluirlo como verdadero a mi conjunto de creencias entrañaría una contradicción (dado que si el antecedente es de hecho falso en mi conjunto de creencias y yo agrego a ese conjunto su negación, entonces en mi conjunto tendría una proposición y su negación lo cual haría a ese conjunto contradictorio). Por lo que es necesario hacer ajustes a mi cuerpo de conocimiento eliminando o cambiando las creencias que entran en conflicto con el antecedente, siempre intentando que esos ajustes sean los mínimos posibles. De esta manera podemos evaluar la verdad o falsedad de una oración condicional mediante tres procedimientos: (i) la adición del antecedente a nuestro conjunto de creencias, (ii) el ajuste necesario de ese conjunto para mantener la consistencia sin modificar la creencia hipotética en el antecedente y (iii) dictaminar si el consecuente es o no verdadero bajo la hipótesis anterior.

Ahora bien, ya que dimos una primera aproximación a las condiciones de evaluación de los condicionales falta establecer de manera más estricta cuáles son sus condiciones de verdad. Para ello Stalnaker se vale de la noción técnica de *mundo posible,* la cual es un análogo ontológico del cuerpo de creencias hipotético que describe Ramsey, y del aparato semántico para lógica modal desarrollado por

Kripke[8]. La intuición detrás de esta aproximación es que cuando uno hace una afirmación condicional está afirmando algo acerca de un mundo posible en el que el antecedente es verdadero. La lógica modal, explica Stalnaker, provee una manera de hablar acerca de lo que es verdad en el mundo actual, en todos los mundos posibles, en al menos un mundo no especificado, por lo que construir una semántica condicional a partir de ella nos permitirá también hablar acerca de lo que es verdad en situaciones posibles que de hecho no son el caso, es decir, situaciones contrafactuales.

Así pues, puestas en términos de mundos posibles las condiciones de verdad de un enunciado condicional se pueden formular de la siguiente forma: dado el condicional 'si A entonces B' para saber si este es verdadero considérese un mundo posible en el cual A es verdad y que difiere mínimamente del mundo que de hecho es el caso. Si B es verdad en ese mundo posible, el condicional será verdadero en el mundo actual (o sea el mundo real). Y si B es falso en el mismo mundo posible, también lo será el condicional en el mundo que es el caso. Aquí el conjunto de todas mis creencias acerca del mundo que proponía Ramsey está representado en la formulación de Stalnaker por el mundo que de hecho es el caso, e igualmente, la idea de que el mundo posible que considera a A como verdadero debe ser mínimamente diferente al mundo actual proviene del requisito que ponía Ramsey acerca de hacer ajustes a mi conjunto de creencias una vez que supongo la verdad de un antecedente. De manera que el mundo a partir del cual se evalúa la verdad del consecuente debe ser diferente del mundo real sólo en la medida en que la inclusión de la verdad del antecedente lo requiera.

[8] S. Kripke, "Semantical Analysis of Modal Logics, I" Zeitschrift für mathematische Logik und Grundlagen der Matematik, vol. 9. 1963. Pp.67-96.

Este primer planteamiento va a volverse a formular usando como modelo el aparato formal de la lógica modal. La teoría semántica de Stalnaker usa el símbolo '>' para representar todo tipo de construcción condicional, por lo que (A>B) se lee como "si A entonces B". También establece como parte de su aparato formal una relación R de accesibilidad entre mundos de manera que siendo α y β mundos posibles la relación (α, R, β) se lee como "β es posible respecto de α". Además establece una función binaria de selección *f* que toma como argumentos una proposición (el antecedente de un condicional) y un mundo posible base, y cuyo valor es un mundo posible en el que la proposición argumento es verdadera. Lo cual quiere decir que si partimos de un condicional A>B acerca del mundo base α, la función selecciona el antecedente A del condicional y el mundo base, y arroja el mundo posible β en el cual A es verdadero, por lo que $f(A, \alpha) = \beta$. La función de selección pretende capturar la idea de la adición hipotética del antecedente al conjunto de mis creencias acerca del mundo por lo que el mundo posible β que es resultado de la función debe ser el mundo posible más parecido a α de entre todos los mundos posibles.

Con estos elementos Stalnaker enuncia de manera más estricta las condiciones de verdad de los enunciados condicionales:

A>B es verdadero en α si B es verdadero en $f(A, \alpha)$;

A>B es falso en α si B es falso en $f(A, \alpha)$.

La teoría incluye también un mundo absurdo λ en el que las contradicciones y todas sus consecuencias son verdaderas. El mundo λ está aislado con respecto a los demás mundos posibles, lo que quiere decir que no es accesible desde ningún otro mundo y además que desde él no es accesible ningún otro mundo tampoco. Su propósito es dar

cuenta del caso en que el condicional analizado tenga un antecedente que sea falso en todo mundo posible accesible desde el mundo base (o sea un antecedente imposible respecto al mundo base).

Los mundos que la función de selección arroja como resultado están ordenados uno respecto de otro en la medida en que son similares al mundo base de modo que el mundo base los precede a todos ellos y son seleccionados justamente por su similitud[9]con él. Existen además algunas restricciones acerca de los mundos seleccionados en casos específicos:

1) Para cualquier antecedente A y mundo base α, A debe ser verdadero en f(A, α). Es decir que en el mundo seleccionado por la función, el antecedente que es argumento de ella debe ser verdadero. Así, al evaluar un condicional, el hipotético mundo en el que el consecuente es verdadero o falso es un mundo en el que el antecedente es verdadero.

2) Para cualquier antecedente A y mundo base α, f(A, α) = λ sólo si no hay un mundo posible respecto de α en el cual A sea verdadero. Esta restricción señala lo que antes había apuntado, que si el mundo seleccionado por la función es el mundo absurdo, el antecedente A es falso en todos los mundos accesibles desde el mundo base α.

3) Para cualquier antecedente A y mundo base α, si A es verdadero en α, entonces f(A, α) = α. Esto significa que si el antecedente es verdadero en el

[9] En los textos en inglés acerca del modelo Stalnaker-Lewis el término usado para describir esta relación de los mundos posibles con el mundo base es *similarity*. Aunque algunos autores como Gladys Palau lo traducen como *similaridad*, de acuerdo con la Academia Mexicana de la Lengua la palabra *similaridad* no existe en el idioma español, por lo que para este trabajo usaré el término *similitud*.

mundo base[10] entonces el mundo seleccionado por la función es el mundo base mismo, ya que por un lado la función selecciona el mundo máximamente similar al mundo base en el que el antecedente es verdadero y por otro si el antecedente es ya verdadero en él no hay mundo más parecido que la función pueda seleccionar que él mismo.

4) Para todo mundo base α y todo antecedente A y A', si A es verdadero en f(A', α) y A' es verdadero en f(A, α), entonces f(A', α) = f(A, α). Lo cual quiere decir que si dos antecedentes son verdaderos en un mundo, no puede haber otro mundo diferente en el que los ambos antecedentes sean verdaderos, debe ser el mismo.

2.2 La Teoría de contrafácticos de David Lewis

La teoría semántica de contrafácticos de David Lewis se distingue de la teoría de Stalnaker porque mientras que Stalnaker busca explicitar las condiciones de verdad de toda construcción condicional Lewis limita su dominio de explicación a los condicionales contra fácticos. El análisis de Lewis parte del tratamiento que Nelson Goodman[11] previamente había dado a los condicionales contra fácticos reduciéndolos a implicaciones estrictas también llamadas necesitantes. La implicación estricta es

[10] Usualmente al emplear la construcción contrafáctica en el lenguaje natural se presupone la falsedad del antecedente, en el caso en que el antecedente de un condicional contrafáctico formulado en el lenguaje natural sea verdadero, lo más probable es que el hablante tenga una creencia equivocada con respecto a la verdad del antecedente. Este caso elimina la necesidad de apelar a un mundo posible diferente del mundo base pues la cualidad contrafáctica del condicional desaparece. En otras palabras, un condicional contrafáctico con el antecedente verdadero no es en realidad un contrafáctico.

[11] Goodman N. "The Problem of Counterfactual Conditionals", The Journal of Philosophy, Vol. 44, No. 5.

también un tipo de enunciado condicional que se caracteriza por que su verdad es necesaria, por lo que en términos de la lógica modal se le representa de esta forma □ (P → Q). En el lenguaje de la lógica modal lo que este enunciado expresa es que el condicional material (P → Q) es verdadero en todo mundo accesible desde el mundo en que □ (P → Q) es enunciado. Lewis va a mostrar con un contraejemplo que las características semánticas de los condicionales contrafácticos no son adecuadamente representadas por el condicional estricto, como se podrá observar las condiciones de verdad de un condicional estricto son demasiado permisivas y nos obligan a aceptar la verdad de contra fácticos que son patentemente falsos bajo contextos de habla normales. David Lewis se sirve de la falacia del refuerzo del antecedente para ilustrar esta dificultad, esta falacia se expresa en el siguiente esquema: ((P→Q) → ((P^R) → Q)), este esquema es un esquema válido de la lógica proposicional pues no hay ninguna instancia en la combinatoria de asignación de valores de verdad para las letras proposicionales en que el antecedente resulte verdadero y la conclusión falsa, sin embargo si este esquema se observa sustituyendo las variables proposicionales por enunciados proposicionales del lenguaje natural puede darse el caso en que la verdad del enunciado implicado de este mismo esquema sea difícil de aceptar aun dada la verdad del enunciado implicante, precisamente por el significado de los enunciados simples, por ejemplo tomando el diccionario P:Pedro duerme bien abrigado, Q: Pedro pasará frío durante la noche, R: Pedro duerme en una tina con hielo, considérese la siguiente fórmula que es una instancia de la falacia ((P→ ¬Q) → ((P^R) → ¬ Q)), puesto el esquema proposicional en forma de razonamiento del lenguaje natural este resulta sumamente contra intuitivo:

Si es verdad que si Pedro duerme bien abrigado no pasará frío durante la noche entonces también es verdad que si Pedro duerme bien abrigado y en una tina con hielo entonces no pasará frío durante la noche.

La falacia del refuerzo del antecedente recibe su nombre como se puede notar del hecho de que al añadir proposiciones al antecedente de un condicional (al "reforzarlo") el valor de verdad del condicional completo no varía, por lo que si una proposición implica a otra no importa cuantas proposiciones más se pongan en conjunción a la proposición implicante, la proposición implicada aún se sigue de esta conjunción puesto que ya se seguía de la primera proposición implicante cuando estaba sola, la proposición implicante original era pues suficiente para asegurar la verdad de la proposición implicada. David Lewis va a hacer pues uso de esta discordancia entre la estructura y el uso del lenguaje lógico para mostrar que el análisis de los contrafácticos como implicaciones estrictas es un análisis fallido.

Supóngase que el siguiente enunciado contrafáctico tiene las mismas condiciones de verdad de un condicional estricto: i) Si hubiera comprado un coche nuevo ayer, hubiera quedado en la ruina. Ahora, asumiendo la verdad de este condicional considérese la siguiente cadena de enunciados contra fácticos: ii) Si hubiera comprado un coche nuevo ayer y después me hubiera sacado la lotería me habría quedado en la ruina; iii) Si hubiera comprado un coche nuevo ayer y después me hubiera sacado la lotería e inmediatamente hubiera gastado todo en mejoras apara mi coche me habría quedado en la ruina; iv) Si hubiera comprado un coche nuevo ayer y después me hubiera sacado la lotería e inmediatamente hubiera gastado todo en mejoras apara mi coche y luego hubiera vendido ese coche por el doble del monto que le invertí me habría quedado en la ruina. Al observar esta secuencia de contrafácticos podemos notar que, si bien dada la verdad de i) el condicional iii) parece razonablemente verdadero mientras que ii) y iv) son claramente falsos. Pero si el condicional contrafáctico tiene las mismas condiciones de verdad del condicional estricto entonces la verdad de i) garantiza que no importa cuantas más proposiciones añadamos al antecedente todo condicional de la secuencia va a ser verdadero. No

obstante, si observamos los contrafácticos ii) y iv) es claro que no podríamos aceptar su verdad bajo un contexto de habla razonable aun aceptando la verdad de i). Es por esto que Lewis no puede aceptar que los condicionales sean equivalentes a los condicionales necesitantes.

En contraste, Lewis va a introducir la idea de que el condicional contrafáctico es un condicional variablemente estricto, esto es a diferencia de un condicional necesitante para el cual los mundos accesibles a él son todos los mundos posibles, para un contrafáctico los mundos accesibles a él no son todos los mundos posibles sino los mundos que cumplan las condiciones semánticas relevantes. Así pues mientras que un condicional estricto $\Box (P \rightarrow Q)$ necesita que el condicional $(P \rightarrow Q)$ sea verdadero para todos los mundos que son accesibles a él (es decir: todos los mundos posibles) para un contrafáctico hay una asignación de mundos accesibles a él que son los mundos que más se asemejan al mundo real excepto en los aspectos solicitados por el enunciado, de manera que en el ejemplo de la secuencia anterior, cada vez que el antecedente se refuerza el conjunto de mundos accesibles es diferente para cada contráfactico de la secuencia lo cual permite que cada uno de los enunciados sea evaluado en condiciones semánticas diferentes a las del enunciado inicial.

La teoría de Lewis está emparentada con la de Stalnaker en tanto que ambas recurren a la semántica de la lógica modal y a la noción de similitud comparativa. La herramienta central que la teoría de Lewis usa para representar la idea de que un condicional contrafáctico es verdadero cuando el consecuente es verdadero en el mundo es la de *sistema de esferas* de accesibilidad, un sistema de esferas de accesibilidad de un mundo posible cualquiera es un conjunto de conjuntos de mundos posibles, donde cada subconjunto del sistema es una esfera. Una esfera de

accesibilidad es un conjunto de mundos posibles accesibles a un mundo base que cumplen con cierto grado de similitud, todos los mundos de una esfera son igualmente similares al mundo base. De manera que dado un mundo base *i* Lewis asigna un sistema de esferas de mundos accesibles a él, en donde cada esfera representa un grado de similitud respecto al mundo base, siendo las esferas más cercanas a *i* los conjuntos de mundos más similares. De tal forma Lewis expresa las condiciones de verdad de los condicionales contrafácticos de la siguiente manera: un enunciado contrafáctico con antecedente P y consecuente Q es verdadero en un mundo base *i*, en relación a un sistema de esferas si y solo si, o bien ningún mundo en que el antecedente es verdadero pertenece a ninguna esfera del sistema (el caso de verdad por vacuidad), o bien alguna esfera en el sistema contiene al menos un mundo en el que el antecedente es verdadero y el condicional material ($P \rightarrow Q$) es verdadero en todo mundo de esa esfera.

En seguida en la sección 2.3, regresaremos a la teoría de Stalnaker para apuntar algunos problemas que tiene su planteamiento para después exponer como es que la reformulación de Lewis de la misma idea resuelve estos problemas

2.3 Problemas en Stalnaker y las ventajas de la reformulación de Lewis

Recapitulando, como vimos en 2.1, en la semántica condicional de Robert Stalnaker, para evaluar la verdad o falsedad de un enunciado condicional contrafáctico de la forma A>B en un mundo α (y de hecho cualquier otro tipo de condicional) se presupone un mundo hipotético β donde el antecedente A es verdadero, mundo que expresamos con la función $f(A, \alpha)$. En este mundo se evalúa la verdad o falsedad de B, de manera que si B es verdadero en β entonces A>B es verdadero en α y en el caso

contrario si B es falso en β entonces A>B es falso en α. Teniendo en cuenta que, siguiendo el paradigma de Ramsey para evaluación de condicionales, el mundo β en la teoría de Stalnaker es el mundo más parecido a α donde A es verdadero.

De este modo tenemos que en el mundo β o bien el consecuente B es verdadero o bien es falso, por lo que en el mundo α o bien A>B es verdadero o bien (A>~B) es verdadero ya que el mundo β seleccionado es el mismo para ambos condicionales, a saber el mundo más parecido a α en que A es verdadero. Por lo que si en el mundo β B es falso entonces A>B es falso y por ende (A>~B) es verdadero, de forma que (A>~B) es equivalente a ~(A>B). Lo anterior significa dos cosas, la primera es que la negación de un condicional contrafáctico en la teoría de Stalnaker es equivalente a ese mismo condicional con el consecuente negado, o sea que:

$$\sim(A>B) \equiv (A>\sim B)$$

Esto tiene la ventaja de reflejar el uso común de los condicionales en el lenguaje natural y además captura las intuiciones que generalmente se tienen acerca de su negación. Por ejemplo si tengo el condicional “si te hubiera conocido en la secundaria habríamos sido pareja entonces” es natural pensar que su negación es “si te hubiera conocido en la secundaria no habríamos sido pareja entonces” o también si tengo el condicional “si fuera millonario gastaría todo mi dinero en alcohol” pareciera natural que su negación sea “si fuera millonario no gastaría todo mi dinero en alcohol”

Lo segundo que la observación hecha en el párrafo anterior significa es que en la teoría de Stalnaker el principio de tercero excluso puede ser expresado de este modo:

$$(A>B) \vee (A>\sim B)$$

gracias a que, cómo habíamos señalado el A-mundo que es más similar al mundo acerca del cual ambos disyuntos se refieren es el mismo para los dos. De manera que en la teoría de Stalnaker subyace el supuesto de que dado un condicional hay un conjunto de mundos accesibles desde él ordenados de modo que hay únicamente un mundo que es el más parecido al mundo base. Llamaremos a este supuesto el supuesto de unicidad.

David Lewis considera este supuesto como uno de los elementos más problemáicos de la teoría de Stalnaker lo cual lo lleva a rechazarlo junto con la formulación de Stalnaker de la negación del contrafáctico y la formulación del principio de tercero excluso. Lewis construye un ejemplo que muestra que es posible encontrar un caso en donde hay más de un mundo máximamente similar a un mundo base. Considérese un condicional contrafáctico de la forma A>B en el que el mundo antecedente fuera uno en que A represente el hecho de que Pelé y Maradona sean compatriotas. En un caso como este se observa que el mundo antecedente en el que ambos son brasileños es igualmente similar al mundo antecedente en el que ambos son argentinos, de modo que elegir mediante la función de selección de Stalnaker cualquiera de los dos sería una elección arbitraria. No existe pues en este caso manera de determinar que cualquiera de los dos mundos es el mundo máximamente similar al mundo base en el orden de mundos supuesto por Stalnaker.

Como vimos en la sección anterior, David Lewis construye una semántica para los condicionales contrafácticos basada en la idea de un sistema de esferas de accesibilidad, las cuales no son otra cosa que conjuntos de conjuntos de mundos accesibles con respecto del mundo base en donde cada conjunto o esfera de mundos es un grado de similitud con el mundo base. A diferencia pues de la teoría de Stalnaker que supone un orden secuencial de mundos en el que cada mundo

representa un grado de similitud más próximo al mundo base, Lewis construye su teoría semántica haciendo que cada grado de similitud esté compuesto no por un solo mundo sino por un conjunto de ellos igualmente similares entre sí al mundo base, y por ende en esos términos establece las condiciones de verdad para los enunciados condicionales contrafácticos.

Como ya se dijo, para Lewis un condicional contrafáctico es verdadero en un mundo α cuando dado un sistema de esferas de mundos con acceso a α, o bien no hay ningún mundo perteneciente a una esfera del sistema donde el antecedente sea verdadero (lo que constituye el caso vacuo) o bien cuando hay al menos un mundo en alguna esfera donde A sea verdadero y el condicional material correspondiente es verdadero en todos los mundos de esa esfera. Como se puede observar, hablar en los términos de Lewis, en términos de conjuntos de conjuntos de mundos posibles requiere que a su vez pensemos en términos de cuantificación sobre mundos para formular las condiciones de verdad de enunciados contrafácticos, es decir que un condicional contrafáctico es verdadero o falso dependiendo de lo que ocurra en una totalidad de mundos posibles de cierto tipo accesibles a él.

Para manejar las nociones cuantificacionales de su teoría Lewis formula los condicionales contrafácticos usando los símbolos '□' y '◊' de la lógica modal, para crear un par de conectivos lógicos usando también el símbolo '→' de la lógica clásica, lo cual le va a permitir referirse con '□→' a totalidades de mundos y con '◊→' a al menos un mundo dentro de los sistemas de esferas. De esta manera un condicional del tipo "si el hecho A hubiera ocurrido, el hecho B sería el caso" puede abreviarse en la teoría de Lewis como:

$$A \Box\!\!\rightarrow B$$

Siendo sus condiciones de verdad las ya mencionadas. De forma paralela se puede abreviar un condicional del tipo “si el hecho A hubiera ocurrido, el hecho B podría ser el caso” de la siguiente manera:

$$A\Diamond\rightarrow B$$

Siendo este enunciado verdadero si y sólo si hay al menos un mundo en alguna esfera donde A sea verdadero y el condicional material correspondiente sea verdadero en al menos algún mundo de esa esfera. Ambas expresiones son interdefinibles como se muestra a continuación:

$$(A\Box\rightarrow B) \equiv df \ \sim (A\Diamond\rightarrow\sim B)$$

$$(A\Diamond\rightarrow B) \equiv df \ \sim (A\Box\rightarrow\sim B)$$

Esta formulación permite a Lewis replantear el problema de la existencia de más de un mundo máximamente similar al mundo base y con ello también logra reformular la negación de un enunciado condicional contrafáctico y por ende del principio de tercero excluso. Tómese por ejemplo los condicionales:

(1) No es el caso que si Pelé y Maradona fueran compatriotas Pelé sería argentino.

y

(2) No es el caso que si Pelé y Maradona fueran compatriotas Pelé no sería argentino.

Podemos simbolizar ambos condicionales en la notación de Stalnaker usando el diccionario

A: Pelé y Maradona son compatriotas

B: Pelé es argentino

De forma que

(1’) ~(A>B)

(2’) ~(A>~B)

En la lógica de Stalnaker las fórmulas son una la negación de la otra, y si se ponen en disyunción forman una instancia del principio de tercero excluso. Expresado así este ejemplo en la Lógica de Stalnaker no hay manera de determinar cuál de los dos condicionales no sería el caso en el mundo en que Pelé y Maradona son compatriotas, esto es, no hay manera de establecer que es falso que Pelé sería o no argentino dado el único mundo en que es compatriota de Maradona y ciertamente aseverar la verdad de ambos al mismo tiempo sería contradictorio, sería de hecho una violación del principio de tercero excluso. Sin embargo si expresamos los mismos condicionales usando la notación de

Lewis tendríamos:

(i) ~(A □→B)

(ii) ~(A □→ ~B)

Lo cual por definición es equivalente a:

(i)’ (A ◊→ ~B)

(ii)’ (A ◊→ B)

De manera que podemos volver a leer los mismos condicionales en el lenguaje natural de la siguiente forma:

(i)‘‘ Si Pelé y Maradona fueran compatriotas Pelé podría no ser argentino

y

(ii)'' Si Pelé y Maradona fueran compatriotas Pelé podría ser argentino

Y dada esta nueva lectura podemos ver que aseverar la verdad de ambos enunciados no resulta necesariamente contradictorio, pues en un mundo en donde Pelé y Maradona fueran compatriotas existe la posibilidad de que Pelé sea argentino, tanto como no lo sea. De modo que en casos como este la verdad o falsedad de un contrafáctico en donde el antecedente exprese el hecho de que Pelé y Maradona son compatriotas no depende de la elección del mundo máximamente similar sino que se reduce a la elección de la esfera más cercana. Si en la esfera ambos mundos son igualmente similares podemos afirmar ambos contrafácticos de posibilidad sin caer en contradicción y también podemos negar ambos contrafácticos de necesidad sin incurrir en contradicción, en ambos casos sin violar el principio del tercero excluso.

Nótese que el condicional (i) es la negación de un condicional contrafáctico por lo que por la equivalencia que hicimos en (i)' tenemos una nueva formulación de la negación de un contrafáctico. O sea que mientras en Stalnaker la negación de un condicional contrafáctico se expresaba en los mismos términos, en Lewis debido a las nociones cuantificacionales que implican la adopción de un sistema de esferas como modelo semántico, la negación de un condicional de caja se fórmula en términos de un condicional de diamante y viceversa. De manera que podemos también ahora reformular el principio de tercero excluso de la siguiente forma:

$$(A\Box\rightarrow B) \vee (A\Diamond\rightarrow\sim B)$$

En conclusión dado que hemos visto cómo es implausible la adopción del supuesto de unicidad que subyace a la teoría de Stalnaker, por las dificultades que señalamos, es preferible por lo tanto la adopción de la formulación de Lewis usando

nociones modales (que a su vez responde al enfoque semántico de Lewis basado en conjuntos de conjuntos como grados de similitud), ya que ello permite dar cuenta de casos en que es evidente que hay más de un mundo máximamente similar al mundo base, como en el ejemplo de Pelé y Maradona en el que el mundo más similar al mundo base puede ser aquel en el que ambos son brasileños o evidentemente también aquel en el que ambos son argentinos. La formulación de Lewis permite igualmente formular la negación de un contrafáctico y el principio de tercero excluso sin los problemas de la formulación de Stalnaker. Habiendo dicho lo anterior, pasaré ahora a discutir en la última sección de este capítulo un problema que presentan ambas teorías así como también una posible solución al mismo apoyándome en el trabajo de la filósofa Gladys Palau.

2.4 El problema de la invalidez de la transitividad contrafáctica

Tanto en la teoría condicional de Stalnaker como en la de Lewis hay ciertos esquemas inferenciales que resultan inválidos, aún cuando en la lógica clásica estos mismos esquemas son válidos. Uno de estos esquemas inferenciales es la *transitividad*, el cual es uno de los tipos de inferencias más intuitivos y más usados en el razonamiento en el lenguaje natural, lo cual representa un problema para el modelo Stalnaker-Lewis porque justamente intenta representar de mejor manera el uso de los condicionales en el lenguaje natural, por lo que a continuación discutiré la invalidez de la transitividad en la semántica de mundos posibles propuesta por Stalnaker y Lewis. Sin embargo es necesario para la exposición de este problema volver a discutir brevemente el refuerzo del antecedente que ya habíamos tocado previamente en la sección 2.2. La inferencia conocida como refuerzo del antecedente es también un esquema válido para la lógica clásica e inválido para el modelo Stalnaker-Lewis. Como habíamos

visto ya, esta inferencia se expresa de la siguiente forma en el lenguaje de la lógica clásica:

(1) $A \rightarrow B \,/\, (A \wedge C) \rightarrow B$

Este esquema inferencial es válido en lógica clásica dado que la verdad de B en la premisa está garantizada de antemano siempre que A sea verdadera, en otras palabras, la verdad de A es suficiente para afirmar la verdad de B en la premisa, así que no importa qué otra proposición aparezca en conjunción con A en la conclusión, B va a seguir siendo verdadera. O puesto de otro modo, si B se sigue de A sola entonces se sigue de A en conjunción con cualquier otra proposición puesto que la verdad de A es condición suficiente para la verdad de B.

Sin embargo, cuando ejemplificamos esta estructura en el lenguaje natural, particularmente tratándose de un condicional contrafáctico, es evidente que el ejemplo no puede ser evaluado con las mismas condiciones de verdad, a menos que estemos dispuestos a ir contra nuestras intuiciones acerca de este tipo de razonamientos. Tomemos como muestra la siguiente instancia del refuerzo del antecedente: "Si hubiera tenido una hija a los 20 años de vida entonces cuando yo tenía 30 ella tendría 10. Por lo tanto, si hubiera tenido una hija a los 20 años y ella hubiera fallecido después de 1 año de nacida, cuando yo tenía 30 años de vida ella tendría 10". Simbolizaremos el argumento anterior de la siguiente manera usando lógica proposicional:

1. $P \rightarrow Q \,/\, (P \wedge R) \rightarrow Q$

Donde:

P: Tuve una hija a los 20 años de vida.

Q: Cuando yo tenía 30 años de vida mi hija tenía 10 años de vida.

R: Mi hija falleció después de un año de nacida.

Proposicionalmente este es un argumento válido pues es imposible que la premisa sea verdadera siendo la conclusión falsa dado que para la falsedad de la conclusión es necesario que tanto ‘P’ como ‘R’ sean verdaderas al mismo tiempo que ‘Q’ sea falsa, lo cual obliga a que la premisa sea falsa también. No obstante, cualquier hablante competente del español tendría problemas para aceptar la conclusión a partir únicamente de la verdad de las premisas, además de que el argumento resulta inválido bajo el modelo semántico que siguen las teorías de Lewis y Stalnaker, es decir el modelo que parte de la noción primitiva presente en el test de Ramsey y que deriva en un análisis de los condicionales contrafácticos en términos de estados de cosas hipotéticos, estados de cosas que formalmente se traducen en los mundos posibles de la lógica modal. A continuación procederé a mostrar que el esquema inferencial de refuerzo conocido como el *refuerzo del antecedente* es efectivamente inválido en estos términos. En este caso en particular usaré el aparato semántico de Stalnaker, por razones de economía de la explicación, analizaré el ejemplo en términos de mundos posibles únicamente, pues al ser un ejemplo sencillo no necesito el aparato de Lewis que usa conjuntos de conjuntos de mundos posibles. Analizado pues este argumento bajo las condiciones de verdad de este modelo se puede mostrar fácilmente que es posible hallar un caso en que la premisa sea verdadera y la conclusión falsa. Simbolicemos el argumento usando notación de Stalnaker y el diccionario de la simbolización previa:

1. $P > Q / (P \wedge R) > Q$

La construcción del caso de invalidez involucra que se seleccionen mundos antecedentes distintos para cada condicional. Es decir que en el caso de la premisa el mundo antecedente hipotético que se selecciona es el más parecido al real (al cual

llamaremos α) en el que es verdadero que tuve una hija a los 20 años de vida y en el que además (para propósitos de la demostración) es falso que ella falleció después de 1 año de nacida[12]. En ese mundo además es verdadero que cuando yo tenía 30 años de vida ella tenía 10. A ese mundo hipotético lo llamaremos β. Dado el estado de cosas anteriormente descrito, en el mundo β la premisa del argumento es verdadera. De modo paralelo, para hacer falsa la conclusión se selecciona el mundo antecedente hipotético más parecido al real en el que tuve una hija a los 20 años pero en el que ella falleció después de 1 año (mundo al que llamaremos γ), evidentemente bajo esta hipótesis el consecuente de la conclusión es falso ya que en el mundo en el que mi hija falleció al año de edad no es posible que a mis 30 ella tenga 10 años de vida, lo que vuelve falsa en γ a la conclusión entera también. De modo que bajo este paradigma semántico (a diferencia del paradigma semántico de la lógica clásica) es posible construir un caso para el refuerzo del antecedente que obedezca las condiciones de verdad de los enunciados condicionales (usando mundos antecedentes distintos para cada enunciado) en el que se muestre que es posible que la premisa sea verdadera y la conclusión falsa, por lo que este esquema inferencial en particular es inválido.

De forma similar podemos ver que la transitividad es inválida bajo el mismo modelo semántico. La transitividad se expresa de la siguiente manera en el lenguaje de la lógica clásica:

(i) $A \rightarrow B$

[12] Si bien el mundo más parecido al real donde se cumple el antecedente puede ser un mundo donde sea verdadero que mi hija falleció al año de nacida, para probar la invalidez del argumento estamos obligados a suponer la verdad del condicional en la premisa, eso obliga a su vez a que el consecuente de la premisa condicional sea verdadero en el mundo antecedente por lo que en ese mundo es obligatorio que mi hipotética hija no haya fallecido después de un año para que pueda ser verdad en ese mundo que cuando yo tenía 30 años de vida ella tenía a su vez 10 años de vida.

(ii) B→C / (A →C)

Y también de forma análoga se puede observar que la validez de este esquema radica en que dado que B es condición suficiente para la verdad de C, la verdad de A garantiza la verdad de C también, ya que A es a su vez condición suficiente para la verdad de B[13]. Ahora, trasladando esta misma estructura al reino de los condicionales contrafácticos podemos formular ejemplos que si bien pueden o no ser evidentemente inválidos a primera vista, sí que se pueden probar inválidos bajo el modelo Stalnaker-Lewis. Considérese el siguiente argumento formulado por Stalnaker[14]: "Si J. Edgar Hoover hubiera nacido en Rusia entonces habría sido comunista. Si Hoover hubiera sido comunista, habría sido un traidor. Por lo tanto, si Hoover hubiera nacido en Rusia habría sido un traidor". Este argumento es válido proposicionalmente pues tiene la estructura proposicional de la transitividad. Sin embargo la sola conclusión parece problemática a simple vista pues si Hoover hubiera nacido en Rusia ya no podría considerársele traidor en los Estados Unidos[15], no obstante su verdad está garantizada por la verdad de las premisas. En cualquier caso podemos construir en los términos del modelo Stalnaker-Lewis una prueba que muestre que es posible que las premisas sean verdaderas y la conclusión falsa. En esta prueba seleccionaremos un mundo antecedente distinto para cada condicional del esquema inferencial. Para hacer verdadera la primera premisa seleccionamos como mundo antecedente el más cercano al mundo real en el que Hoover haya nacido en Rusia, en ese mundo es verdadero también que Hoover es comunista. Para verificar la segunda premisa seleccionamos

[13] Con esta explicación trato de que el lector entienda de manera general el funcionamiento del esquema inferencial, sin embargo tanto en este caso como en el caso del refuerzo del antecedente basta con hacer la tabla de verdad de ambos esquemas para disipar cualquier duda que el lector pudiera tener acerca de la validez de los mismos en la lógica clásica.

[14] Stalnaker, Robert. *A Theory of conditionals*. En Harper, William L. (et al). Ifs. Springer, 1980. P.48

[15] Se entiende en el contexto del ejemplo que Hoover habría sido un traidor a los E.U.

un mundo en el que Hoover es tanto comunista como un traidor (para los Estados Unidos) por lo que en este mundo Hoover es de manera obligatoria de nacionalidad estadounidense y no rusa. Y finalmente no es difícil ver que para falsear la conclusión seleccionamos el mundo en el que es verdadero que Hoover es ruso pero es falso que es un traidor a los Estados Unidos, el cual es un mundo perfectamente plausible. En este escenario podemos constatar que hay una manera de verificar las premisas y falsear la conclusión evaluando cada enunciado condicional en un mundo antecedente distinto. Por lo que la transitividad es inválida para la semántica Stalnaker-Lewis.

Hasta aquí parece ser que la teoría semántica Stalnaker-Lewis de enunciados condicionales nos permite construir un caso que muestra que es posible que las premisas en una cadena inferencial respecto a la cual (en los términos ortodoxos y tradicionales de la lógica clásica) podría predicarse transitividad, sean verdaderas y sin embargo la conclusión falsa. Esto sucede debido a que la teoría permite asignar distintos mundos antecedentes para cada afirmación dentro del esquema inferencial. No obstante, al observar el uso de razonamientos de transitividad contrafáctica en el lenguaje natural podemos percatarnos de que existen muchos ejemplos que difícilmente podríamos considerar inválidos y que sugieren que en contextos de habla normales la transitividad contrafáctica sí es una estructura inferencial válida, aun cuando en la teoría de Stalnaker-Lewis resultan razonamientos inválidos.

Consideremos el siguiente argumento: “Si hubiera terminado mi tesis el año pasado habría entrado a la maestría en este ciclo. Si hubiera entrado a la maestría en este ciclo, estaría recibiendo la beca del Conacyt. Luego, si hubiera terminado mi tesis el año pasado estaría recibiendo la beca del Conacyt”. Pareciera que en este ejemplo suponer la verdad de las premisas garantiza la verdad de la conclusión, como sucede en un argumento válido, dicho de otro modo, al asumir que las primeras dos

afirmaciones condicionales son lo que de hecho es el caso parece inevitable que el tercer condicional también lo sea. Sin embargo, es importante tener presente que en la teoría semántica Stalnaker-Lewis que un condicional contrafáctico sea de hecho el caso, es decir, sea verdadero, depende de que lo que afirman el antecedente y el consecuente sea el caso en un mundo hipotético; considerando estas características de la teroría, es posible suponer mundos hipotéticos para cada condicional del argumento de manera que la conclusión sea falsa aun cuando las premisas sean verdaderas. Intentaré pues asignar un mundo antecedente para cada condicional del argumento con la intención de mostrar su invalidez. Simbolizaré el argumento de la siguiente forma: Sea T: terminé mi tesis el año pasado, M: entré a la maestría maestría en este ciclo, y B: recibo la beca del Conacyt. Sea también α el mundo real. El argumento se expresa de la siguiente forma:

1. T > M
2. M > B/T > B

Asignaré para la premisa 1 el mundo al que llamaré β, que es el mundo más parecido a α en el que T es verdadera, y como necesito para mi demostración que la premisa 1 sea verdadera en α, entonces en el mundo β es verdadera también la proposición M, por lo que en efecto la premisa se hace verdadera en el mundo α; en este mundo no nos interesa el valor de verdad que tenga B. Para la premisa 2 asignare el mundo γ, el cual es el mundo más parecido a α en el que M es verdadera y como la premisa 2 debe ser verdadera en el mundo α, B es también verdadera en el mundo γ; en γ no nos interesa el valor de verdad de T por lo que hasta donde hemos visto γ podría ser o no idéntico a β. Finalmente para la conclusión asignaré el mundo δ el cual es el mundo más parecido a α en el que T es verdadera y B es falsa pues eso hace que la conclusión sea falsa en el mundo α, como para evaluar el valor de verdad de la

conclusión en α no nos interesa el valor de verdad de M en δ, los mundos β y δ bien podrían ser idénticos pero los mundos γ y δ son necesariamente diferentes pues en γ la proposición B es verdadera mientras que en δ la proposición B es falsa.

De esta manera he encontrado una asignación de mundos antecedente para cada enunciado contrafáctico que me permite mostrar que es posible que la conclusión del argumento del ejemplo sea falsa, aun en el caso en que las premisas sean verdaderas, por lo que hemos demostrado su invalidez en la teoría Stalnaker-Lewis. Poniéndolo en términos del contenido semántico del ejemplo, a pesar de que las premisas son verdaderas en el mundo real después de la asignación de mundos hipotéticos donde éstas se verifican, en el mundo real es falso que "si hubiera terminado mi tesis el año pasado estaría recibiendo la beca del Conacyt" si suponemos un mundo donde sí terminé mi tesis el año pasado pero no logré entrar a la maestría. Al lector perspicaz esta asignación le pude parecer ya sospechosa, en el sentido en que parece no cumplir con los requerimientos que (sensatamente) alguien que enunciara este argumento impondría de manera implícita a los enunciados del mismo. Pareciera injusto que, si al asumir la verdad de las premisas suponemos mundos en donde sus antecedentes respectivos son verdaderos, para la conclusión asignemos un mundo donde estas suposiciones no tienen efecto. La teoría Stalnaker-Lewis permite seleccionar mundos diferentes para cada contrafáctico de un esquema inferencial, pero al enunciar el mismo argumento en un contexto de habla normal, el hablante considera implícitamente que los mundos son relevantes al contexto en el que se está argumentando. En el caso de nuestro ejemplo no nos interesa saber si la conclusión puede ser falsa en un mundo donde no entré a la maestría, por el contrario si habíamos asumido previamente que si hubiera entrado a la maestría en este ciclo, estaría recibiendo la beca del Conacyt, los mundos relevantes para evaluar la conclusión son los mundos en los que también es verdad que entré a la maestría. Lo

mismo sucede con el ejemplo de J. Edgar Hoover, en ese ejemplo para evaluar la verdad de "Si Hoover hubiera sido comunista, habría sido un traidor" de manera implícita no consideramos los mundos en que Hoover es estadounidense sino los mundos en que es ruso, los cuales habían sido considerados al asumir la verdad de la premisa previa. De manera que dada la consideración implícita del mundo en que Hoover es estadounidense es imposible hacer esa premisa verdadera, por lo que es imposible también, para ese ejemplo, el caso de premisas verdaderas y conclusión falsa.

Gladys Palau[16] sostiene que las condiciones de verdad de los enunciados contrafácticos cuando aparecen en contextos inferenciales no son las mismas que cuando aparecen de manera aislada, es decir, en un contexto inferencial los condicionales contrafacticos no deben ser evaluados de manera independiente, pues al enunciar una inferencia contrafáctica el hablante supone un contexto al cual los enunciados están ligados, es decir, se presupone que en el contexto argumentativo se establece una relación entre ellos y no se les considera de manera aislada. Palau llama a este contexto el contexto contrafáctico y caracteriza los requerimientos de evaluación de verdad de los enunciados que estén en dicho contexto de la siguiente forma:

Todas las inferencias por transitividad contrafáctica tienen la siguiente forma: si A → B y B → C, luego A → C. Todo lo que un hablante pide cuando enuncia una inferencia de este tipo es que una vez seleccionado el mundo antecedente A más próximo a i [el mundo real] en el que se verifica B, o sea el mundo en el que la primera premisa es verdadera, en el mundo antecedente seleccionado para verificar

[16] Palau Gladys. Condicionales contrafácticos: condiciones de verdad y semántica de mundos posibles. Acerca de las teorías de R. Stalnaker y D. Lewis. En Crítica, Volúmen 12, Número 34. IIF, UNAM. P. 31.

si se cumple la segunda premisa siga siendo verdadero A. Más brevemente, se pide que el B-mundo antecedente de la segunda premisa sea así mismo un A-mundo. Desde un enfoque distinto, podría afirmarse que el antecedente A del primer contrafáctico fija o retrotrae toda la situación contrafáctica a los mundos en los que él se verifica[17]

En otras palabras, Palau pone como restricción para la evaluación de un argumento de transitividad contrafáctica que al asignar un mundo antecedente a la primera premisa y asumir en él la verdad del antecedente, esa asunción tenga efecto en el mundo en el que se evalúa la segunda premisa. De esta manera se establece un contexto contrafactual apropiado y se ligan las premisas del argumento transitivo. Nótese que cuando Palau habla de la "primera premisa" en la cita anterior se refiere a la premisa que aparece escrita o enunciada en la argumentación que es la que según ella fija el contexto contrafáctico, pero como las premisas en un argumento tienen la propiedad conmutativa, en el caso de una inferencia de tipo transitiva, se puede considerar a cualquiera de las dos premisas la "primera premisa", por lo que de manera más general podemos decir que el contexto contrafáctico no sólo es fijado por la premisa que aparece primero en la enunciación sino que se establece asumiendo la verdad de ambos antecedentes en los mundos hipotéticos, es decir los mundos en los que se evalúan las premisas deben ser mundos donde se verifican ambos antecedentes (esta tesis es mía no de Palau). De igual forma para la conclusión se pide que esta sea evaluada en mundos donde los supuestos del contexto contrafáctico persisten. En ese sentido, en casos de inferencias contrafáctica la máxima de que para cada enunciado contrafáctico se seleccione el mundo o la esfera de mundos más cercanos a él se sustituye por la máxima de que en una inferencia contrafáctica los mundos

[17] Ib., p.31

seleccionados para cada enunciado sean mundos o esferas que respeten el contexto contrafáctico de la inferencia, en palabras de Palau:

"*El punto central de nuestra discrepancia con ambas teorías es que en ellas se exige la elección siempre del mundo antecedente más cercano posible respecto del mundo real* i *sin tomar en cuenta la situación contrafáctica global o contexto contrafáctico. Con nuestra restricción, los mundos antecedentes seleccionados son cada vez más lejanos respecto de* i"[18]

Como muestra de lo anterior, en el ejemplo que consideramos previamente para mostrar la invalidez de la transitividad, es evidente que al seleccionar un mundo para evaluar la conclusión en el que tanto 'T' como 'M' son verdaderos, seleccionamos un mundo menos similar al real que el mundo en el que ´T´ es verdadero pero ´M´ es falsa puesto que en el mundo real (al ser 'T' y 'M' antecedentes de enunciados contrafactuales) ninguna de las dos proposiciones es verdadera.

Así pues, es patente ahora que si bien la maquinaría de la teoría semántica de Stalnaker-Lewis permite asignaciones de mundos en donde podemos verificar las premisas de una inferencia de transitividad al mismo tiempo que podemos falsear la conclusión de la misma, estas asignaciones no corresponden a los parámetros semánticos implícitos en el uso competente de este tipo de inferencias. Por lo que podemos afirmar que la noción de que la transitividad es una inferencia invalida, es resultado de un uso equivoco de los elementos teoría

Ahora, regresemos al principio de esta sección donde discutíamos la invalidez de la inferencia por refuerzo del antecedente. Palau se pregunta después de haber

[18] Palau Gladys. *Condicionales contrafácticos: condiciones de verdad y semántica de mundos posibles. Acerca de las teorías de R. Stalnaker y D. Lewis*. En Crítica, Volúmen 12, Número 34. IIF, UNAM. P. 32.

afirmado la validez de la transitividad si esto no obliga a aceptar la validez de la inferencia por refuerzo del antecedente (lo cual no es deseable) ya que la prueba de su invalidez descansaba en la posibilidad de elegir mundos antecedentes diferentes para cada enunciado de la inferencia. Adicionalmente Palau presenta un esquema inferencial (formulado aquí con la notación de Stalnaker) con la siguiente estructura que parece implicar la validez de la inferencia por refuerzo del antecedente dada la validez de la transitividad:

1. $(A \wedge C) > A$

2. $A > B \,/\, (A \wedge C) > B$

A primera vista parece que la conclusión de este esquema puede ser inferida de las premisas por transitividad, sin embargo si apelamos a las restricciones acerca del contexto contrafáctico tenemos que el mundo o esfera de mundos en que debe ser evaluada la premisa 2 debe ser un mundo donde tanto ‘A’ como ‘C’ sean verdaderas, lo cual posibilita que en ese A-C mundo el consecuente de la segunda premisa sea falso. Ahora bien dado que la premisa 1 es una verdad lógica podemos eliminarla de la deducción e inferir la conclusión únicamente de 2. Nótese que el contexto contrafáctico cambia al desaparecer la premisa 1 por lo que las condiciones de verdad de la premisa 2 tomada aisladamente son diferentes a sus condiciones de verdad en conjunción con la premisa 1. No obstante aún sin la premisa 1 la inferencia resulta inválida ya que aún cuando la conclusión cumpliría el requisito de que su antecedente sea un A-mundo (dado el contexto fijado por la premisa 2) ese A-mundo debe ser obviamente también un C-mundo, lo que posibilita la falsedad del consecuente en la conclusión.

CAPÍTULO 3

ANÁLISIS DE UNA ALTERNATIVA A LOS MUNDOS POSIBLES

3.1 Cuestionando el modelo Stalnaker-Lewis

Hasta este punto en el presente trabajo hemos visto cómo es problemático considerar a todas las afirmaciones condicionales del lenguaje natural como condicionales materiales de la lógica clásica ya que las condiciones de verdad de estos últimos parecen no adecuarse a las propiedades semánticas de muchas de las afirmaciones condicionales del lenguaje natural, específicamente a las propiedades semánticas de los condicionales de tipo contrafáctico, por lo que he intentado buscar una explicación alternativa de en qué circunstancias este tipo de condicionales son verdaderos o falsos, en otras palabras he tratado de encontrar una teoría que establezca de manera correcta las condiciones de verdad de los enunciados condicionales contrafácticos.

La teoría que hemos revisado es el modelo semántico compuesto por las teorías de Robert Stalnaker y David Lewis, el cual toma como punto de partida la semántica de mundos posibles desarrollada por Saul Kripke. Este modelo ha sido cuestionado por filósofos que consideran o bien que una teoría como la de Stalnaker y Lewis presenta problemas en su planteamiento, o bien que es posible formular una teoría explicativa acerca de la semántica de los condicionales contrafácticos sin necesidad de apelar al aparato semántico de los mundos posibles. Para la siguiente sección de esta tesis voy a presentar una revisión de las ideas acerca de los condicionales contrafácticos que el filósofo Michael Woods expone en su libro *Conditionals*. Woods señala por una parte varios problemas que tiene el modelo explicativo para contrafacticos basado en mundos posibles tal como lo plantean Stalnaker y Lewis, y

por otra parte señala también que el planteamiento general de dicho modelo es inadecuado y que el uso de mundos posibles como herramienta teórica es dispensable. Finalmente, Woods propone un modelo diferente de análisis de condicionales en el que rechaza la distinción entre condicionales indicativos y contrafácticos. Lo que lo lleva a concluir que los condicionales que tradicionalmente son llamados contrafácticos tienen en realidad una forma lógica distinta a la asumida en el análisis por mundos posibles y por ende sus propiedades semánticas son explicables de mejor manera en términos de su verdadera forma lógica.

Es de mi interés en este trabajo el discutir una postura que esté en conflicto con el paradigma Stalnaker-Lewis porque así podré someter a prueba este modelo y evaluar la viabilidad de un análisis de condicionales contrafácticos basado en mundos posibles en comparación con un análisis basado en un aparato semántico diferente. Por lo que en la última parte de este capítulo voy a tratar de examinar las objeciones de Woods, así como su propia teoría contrastándola con las de Stalnaker y Lewis para poder dictaminar si el modelo propuesto por estas últimas se sostiene y así poder presentar mis conclusiones en la última sección de esta tesis.

3.2 Las objeciones de Michael Woods a la semántica de mundos posibles para condicionales contrafácticos

En su libro *Conditionals*, Michael Woods elabora una lógica de los enunciados condicionales en general, en la que rechaza (como veremos más detalladamente después) la distinción tradicional entre condicionales indicativos y subjuntivos (contrafácticos) mediante la negación de la presuposición de que existen condicionales que implícitamente asumen la falsedad del antecedente. Después,

propone una clasificación propia basada en las diferencias semánticas que entre los condicionales provocan características como el tiempo verbal. Por último, propone también un análisis lógico de los condicionales en sus propios términos. Este análisis es radicalmente diferente al modelo Stalnaker-Lewis puesto que no parte del test de Ramsey ni usa la teoría de mundos posibles como maquinaria semántica. No obstante, Woods reconoce que en el uso cotidiano, un hablante que enuncia un condicional 'contrafáctico'[19] usualmente cree que el antecedente de dicho condicional es falso y que -por esa misma razón- es entendible la apelación a una semántica de mundos posibles para entender estos condicionales, pues -señala- es natural pensar en situaciones posibles imaginarias cuando se trata de condicionales cuya hipótesis es la verdad de una proposición que se piensa es de hecho falsa, Es decir que aún cuando reconoce que pensar los condicionales contrafácticos en términos de mundos posibles es razonable, Woods rechaza esta estrategia e intenta en cambio construir una teoría alternativa.

I argued that one traditional piece of doctrine about conditionals is false: that there is a class of conditional statements -traditionally labelled 'counterfactual' or 'subjunctive'- which convey the implication, by the tense and mood of the verb that the antecedent is false. (...) Although I have rejected the Counterfactuality Thesis, I have recognized that 'counterfactual' conditionals are typically used by someone who believes the antecedent to be false. So an analysis in terms of possible worlds, in

[19] Desde el punto de vista de Woods los condicionales que tradicionalmente son clasificados como contrafácticos en realidad no lo son, de ahí que use comillas para hablar de ellos. En lo subsecuente, si al exponer sus ideas uso comillas para hablar de condicionales contrafácticos será porque estoy expresando su opinión acerca de los contrafácticos. Cuando hable de mis ideas acerca de los condicionales no usaré las comillas pues yo no estoy de acuerdo con él en este aspecto.

which provision is made for possible situations distinct from the actual, is naturally applied to them.[20]

Sin embargo Woods se opone explícitamente al uso de mundos posibles como herramienta de análisis de condicionales, por lo que es menester para su postura señalar los problemas que tiene la semántica de mundos posibles para el análisis de condicionales, específicamente los condicionales que se consideran usualmente como 'contrafácticos'. La objeción central de Woods es que la noción de similitud como criterio de selección de mundos posibles en los que la verdad de un contrafáctico es evaluada en el modelo Stalnaker-Lewis está insatisfactoriamente definida, y -por consiguiente- el análisis de contrafácticos puede ser reformulado en otros términos que no recurran a los mundos posibles. En última instancia Woods sostiene que, para hacer un análisis adecuado de las diferencias semánticas de los condicionales 'contrafácticos', se debe atacar el problema de cuál es la forma lógica de estos enunciados. Esto lo veremos cuando revisemos la propuesta de Woods en la sección 3.2.

Comenzaremos a revisar el ataque de Woods al modelo Stalnaker-Lewis señalando los problemas que detecta en este modelo. Algunos de estos problemas ya los habíamos revisado en una sección anterior pero aquí los revisaremos de la manera en que Woods los presenta. Los problemas están relacionados a diferentes características del modelo, estas son: (i) el supuesto de unicidad, (ii) el supuesto del límite, (iii) el supuesto de que el mundo más cercano al mundo que de hecho es el caso es idéntico al mundo que de hecho es el caso (puesto que ningún mundo es más cercano al mundo real que el mundo real mismo) y (iv) la forma en que Lewis trata los condicionales con antecedente necesariamente falsos que en su teoría son

[20] Woods, Michael. *Conditionals*. Oxford University Press. Oxford. 1997. P.40

verdaderos por vacuidad, puesto que si no hay un mundo posible en el que el antecedente de un condicional contrafáctico sea verdadero entonces es imposible que se cumplan las condiciones de falsedad de ese contrafáctico.

Como hemos visto, (i) y (ii) son supuestos que están presentes en la teoría de Stalnaker pero no en la de Lewis, (i) consiste en el supuesto de que dado un condicional contrafáctico y un mundo base hay un solo mundo posible que es el más cercano a ese mundo base en el que el antecedente es verdadero. El supuesto (ii) consiste en que dado un condicional contrafáctico y un mundo base no se da el caso de que haya siempre un mundo cada vez más cercano al mundo base de manera infinita. Woods prontamente señala que sostener (i) es poco plausible pues dada una asunción contrafáctica es posible que esta se cumpla de diferentes maneras igualmente cercanas al mundo base, lo cual haría imposible la selección de un mundo sobre otro para la evaluación de un contrafáctico. Él mismo ejemplifica:

I may suppose, contrary to fact, that I encountered some acquaintance on a journey, which might have come about in countless ways that would have preserved close similarity to the actual world; but it is hardly plausible to hold that we recognize just one of the scenarios that would have made the antecedent true as closer to the actual than any other.[21]

De la misma manera tampoco es plausible sostener (ii) pues parece perfectamente posible tener una cadena infinita de mundos cada vez más cercanos a un mundo base si consideramos contrafácticos como "Si hubiera corrido más de un kilómetro en la mañana me sentiría muy animado para ir a trabajar". Woods entonces apunta que Lewis ya había reformulado las condiciones de verdad de los contrafácticos para no estar obligados a aceptar (i) y (ii) diciendo que en vez de que

[21] Woods, Michael. Conditionals. Oxford University Press. Oxford. 1997. P.45

un contrafáctico sea verdadero si se da el caso que el consecuente es verdadero en el mundo más cercano al mundo base en el que el antecedente es verdadero, la verdad de un contrafáctico dependa en cambio de si el consecuente es verdadero en algún mundo posible en el que el antecedente sea verdadero y que a su vez este mundo esté más cerca del mundo base que cualquier otro en el que el consecuente sea falso. Esta reformulación no se compromete ni con (i) ni con (ii) y evita los problemas de aceptar que haya más de un mundo posible que sea más cercano al mundo base y que incluso pueda haber un número infinito de mundos cada vez más cercanos al mismo. Sin embargo, señala Woods, esta reformulación permite que en la teoría de Lewis sea posible que dado un condicional contrafáctico con antecedente P y consecuente Q, haya mundos cercanos en alguna medida al mundo base en los que P es verdadero y Q es verdadero, y haya mundos cercanos en la misma medida al mundo base en los que P es verdadero y Q falso. Lo cual haría que fueran falsos al mismo tiempo los condicionales "Si P hubiera sido el caso, Q habría sido el caso" y "Si P hubiera sido el caso, Q no habría sido el caso". Esto significaría que en la teoría de Lewis sería falso el principio de tercero excluso condicional, que es la tesis de que dadas dos proposiciones P y Q cualesquiera o bien es verdadero que "Si P hubiera sido el caso, Q habría sido el caso" o bien es verdadero que "Si P hubiera sido el caso, Q no habría sido el caso". Woods intenta entonces salvar el supuesto de unicidad, lo que permitiría que el principio de tercero excluso condicional se mantuviera en pie, pues si solamente existe un mundo que es el más cercano al mundo base en el que el antecedente de un contrafáctico es verdadero, entonces en ese mundo el consecuente es inequívocamente o bien verdadero o bien falso. Woods sostiene que aunque el supuesto de unicidad es difícil de aceptar como tal, se puede señalar que un hablante en contextos normales de enunciación, cuando usa un condicional contrafáctico,

generalmente asume que hay un mundo en el que se da esa situación contrafactual de la siguiente manera:

…we may think that, even though that Assumption is hardly defensible in general, it is presupposed by someone who uses a 'counterfactual' conditional that there is some one possible Word closest to the actual in which the antecedent in question is true. When I say that if Jones had been present, he would have voted for the proposal, I am saying that in the situation that most resembles the way things actually are in which Jones is present, assuming that there is one, he voted in favour. If that assumption does not hold, the conditional in question (and likewise its contrary) is neither true nor false. This suggestion surely has considerable plausibility.[22]

En lo que respecta a (iii) Woods señala lo que en la teoría de Lewis sucede cuando el antecedente de un condicional contrafáctico es verdadero. Aunque Lewis sostiene que los condicionales contrafácticos son usados apropiadamente cuando se presupone la falsedad del antecedente, esta no forma parte integral del enunciado mismo, sino sólo de su uso, por lo que aún cuando un condicional contrafáctico tenga un antecedente verdadero éste es susceptible de ser tratado de la misma manera que un condicional contrafáctico con el antecedente falso. En este caso si aplicamos el aparato de Lewis, el condicional sería verdadero si el consecuente es verdadero en el mundo base y falso si el consecuente es falso también en el mundo base, pues parece razonable sostener (iii) es decir que el mundo más cercano en el que es verdadero el antecedente de un condicional contrafáctico que de por sí tiene el antecedente verdadero desde el mundo base es el mundo base mismo. A Woods le preocupa que estas condiciones de verdad obliguen a aceptar condicionales cuyo antecedente y

[22] Woods, Michael. Conditionals. Oxford University Press. Oxford. 1997. P.46

consecuente no estén semánticamente relacionados, problema que se da incluso con el condicional material como lo señala a continuación:

The main argument against counting all 'counterfactual' conditionals with true antecedents as true if their consequents are true is similar to the standard arguments against treating Simple Conditionals as material conditionals: it would require us to hold as true 'counterfactual' conditionals with unrelated antecedent and consequent that we should not normally use[23]

Finalmente acerca de (iv) Woods apunta que es un error de la teoría de Lewis el asignar el valor de verdad verdadero a todos los condicionales contrafácticos con antecedente imposible pues hay instancias en que usamos de manera inequívoca contrafácticos cuyo antecedente sabemos es imposible y sin embargo no aceptaríamos su verdad necesariamente. Un ejemplo claro es el razonamiento matemático por reducción al absurdo en el que usualmente para probar una afirmación se asume como verdadero algo imposible.

…it appears that we do use such conditionals, knowing the antecedent to be impossible. One example, mentioned already, is in mathematical reasoning: 'If the square root of two had been rational…', where a certain hypothesis is reduced to absurdity. Again, someone may initiate a speculation with 'If I had been born in the eighteenth century…' ; yet, if one accepts the claim that what a thing comes from it necessarily comes from, maybe there is no possible world in which I am born in the eighteenth century. We do use such conditionals, however, and we adopt a discriminating attitude towards their truth-values: not every such conditional with an

[23] Woods, Michael. Conditionals. Oxford University Press. Oxford. 1997. P.46

impossible antecedent is thought to be true, as is evident from the mathematical examples.[24]

Ahora bien, además de los problemas ya señalados, el rechazo de Woods hacia una semántica de mundos posibles como aparato de análisis para las condiciones de verdad de los condicionales contrafácticos (o incluso para cualquier otro tipo de condicional) se funda en las dificultades que para él presenta en este modelo la relación de similitud como criterio de selección de mundos posibles relevantes en los que se evalúa un condicional. A ojos de Michael Woods la explicación de la noción de similitud, particularmente en la teoría de Lewis no es satisfactoria. De manera general, dice Woods, se puede cuestionar la noción de similitud señalando que al escoger un mundo posible basándonos en su similitud con el mundo que de hecho es el caso no queda claro en qué forma debe ser este mundo posible similar al mundo base pues los aspectos en que un mundo es similar a otro son diversos y dependen de nuestros propósitos cuando enunciamos un contrafáctico. Dicho de otro modo Woods considera que no queda claro cómo escoger un mundo posible en lugar de otro para evaluar un condicional contrafáctico usando la similitud como criterio de selección dado que es difícil decidir cuándo un mundo es más similar a otro si hay tantas cosas que pueden variar entre mundos, además de que parece difícil también decidir cuáles son las diferencias relevantes según la hipótesis que se haga en el antecedente. Lewis considera[25] que el mundo más relevante para evaluar un condicional contrafáctico y preservar las intuiciones que tenemos acerca del valor de verdad de estos es uno cuya similitud con el mundo que de hecho es el caso se dé de la siguiente manera: que coincida con el mundo base en todos los aspectos hasta el momento en que se cumple

[24] Woods, Michael. Conditionals. Oxford University Press. Oxford. 1997. P.47

[25] 'Counterfactual Dependance and Time´s Arrow', Nous, 13 (1979), 455-76, reimpreso en Lewis, *Philosophical papers*, ii. 32-52. La referencia aparece en el texto original de Woods.

la hipótesis del antecedente, momento después del cual el mundo hipotético diverge del mundo base, conservando además el mundo hipotético las mismas leyes naturales que el mundo base. Un mundo como este, dice Lewis, se adecua a nuestros propósitos mucho mejor que por ejemplo un mundo que sea similar al mundo base solo en cierta medida pero consistentemente a través del tiempo. Es decir, es preferible un mundo que sea exactamente igual dentro de un periodo de tiempo limitado que uno que sea sólo aproximadamente similar en todo momento, además de que conserve también las mismas leyes naturales en lugar de que coincida en hechos específicos.

[The world that] seems to be the world we consider when evaluating the conditional [is the] one in which everything occurs as it actually does until the time of the antecedent, and then events unfold, in accordance with laws, in a way that deviates dramatically from the actual. (...) such a world is the one that most resembles the actual.[26]

Pongamos como ejemplo el siguiente condicional contrafáctico:

"Si Kurt Cobain no se hubiera disparado en abril de 1994, entonces Nirvana seguiría dando conciertos".

Dadas las características del mundo relevante para evaluar un condicional como éste tendríamos que seleccionar un mundo en el que se preservaran las mismas leyes naturales que tiene el nuestro y en el que además todo es exactamente igual hasta el momento anterior al que Kurt Cobain se dispara a sí mismo en abril de 1994. Según este criterio este mundo es suficientemente similar al nuestro para evaluar la verdad del consecuente dada la hipótesis del antecedente y así determinar el valor de verdad del condicional completo. De manera que si, dada la coincidencia del mundo

[26] Woods, Michael. Conditionals. Oxford University Press. Oxford. 1997. P.49

hipotético con respecto del mundo real hasta el momento en que se da el antecedente y el efecto de las leyes naturales de nuestro mundo, es verdadero que Nirvana sigue dando conciertos, entonces es verdadero el condicional contrafáctico completo. En caso contrario el condicional contrafáctico es falso.

Este criterio de selección de mundos presenta problemas evidentes para Woods. Woods sostiene que los dos requisitos del mundo hipotético entran en conflicto entre ellos, es decir que la idea de un mundo hipotético que es exactamente igual al mundo real hasta un momento dado y que luego diverge abruptamente es inconsistente con la preservación de las leyes del mundo real en ese mundo hipotético. Esto se hace patente si consideramos casos en los que no interviene la voluntad humana para que el antecedente sea satisfecho. Por ejemplo considérese el siguiente condicional:

"Si no hubiera temblado el 19 de septiembre de 2017 en México, la ciudad de Jojutla no hubiera sufrido ningún daño".

En este caso si suponemos un mundo hipotético que coincide con el nuestro hasta un momento antes del temblor del 19 de septiembre de 2017 parece muy difícil que podamos aceptar que el temblor no tenga lugar sin que se violen las leyes naturales puesto que la secuencia de eventos geológicos (que están gobernados por las leyes naturales) que dieron como resultado el movimiento telúrico tendrían lugar tanto en el mundo real como en el hipotético dada su coincidencia hasta antes del sismo. Woods señala que ante esta objeción Lewis sostiene que en estos casos postulamos un mundo hipotético que coincide con el real en sus leyes y hasta el momento antes de que suceda el antecedente y luego el antecedente se vuelve verdadero por obra de un "milagro pequeño". En otras palabras una desviación pequeña de las leyes naturales del mundo real en el mundo hipotético es requerida para la construcción de un mundo hipotético posible en donde el antecedente del condicional contrafáctico

correspondiente sea verdadero y por consiguiente se pueda en ese mundo evaluar la verdad del consecuente también.

De acuerdo con Woods, Lewis sostiene que comparativamente un mundo en el que ocurre una sola desviación mínima de las leyes naturales es más cercano al mundo real que un mundo en el que ocurre una desviación significativa o uno en el que ocurren numerosas desviaciones mínimas. En el planteamiento de Lewis el "milagro pequeño" permite que el antecedente sea verdadero aún cuando el mundo hipotético coincida con el mundo real hasta antes de que se cumpla el antecedente. Habiéndose cumplido el antecedente sus consecuencias se ramifican de manera que posteriormente el mundo hipotético difiere del real en gran medida y continua divergiendo de él conforme pasa el tiempo. Esto quiere decir que para que hubiera coincidencia del mundo hipotético con el real en cualquier momento posterior al antecedente tendrían que suceder muchas desviaciones de las leyes naturales del mundo real en el mundo hipotético, lo cual significaría que un mundo así estaría cada vez más alejado del mundo real, por lo que Lewis considera el mundo hipotético que coincide con el real hasta el momento antes del antecedente y difiere de él subsecuentemente como el mundo hipotético más cercano al mundo real.

Lewis construye la teoría de esta manera, sostiene Woods, porque le interesa justificar el rechazo de cierto tipo de condicionales contrafácticos que expresan un tipo de asimetría temporal. Consideremos los siguientes dos condicionales:

1) Si Lulú no hubiera ido a la fiesta donde nos conocimos hace dos años, ella no habría ido al cine conmigo hoy.

2) Si Lulú no hubiera ido al cine conmigo hoy, ella no habría ido a la fiesta donde nos conocimos hace dos años.

En el caso de 1) parece no haber problema alguno para evaluarlo con el aparato de mundos posibles. Podemos suponer sin ningún problema que el mundo en que Lulú no fue a la fiesta donde nos conocimos, ella tampoco fue al cine conmigo el día de hoy. Pero si consideramos el condicional 2) puede parecernos extraño formular algo así, donde parece que el hecho de que Lulú no fue al cine conmigo hoy provocó de alguna manera que faltara a la fiesta donde nos conocimos hace dos años. Woods explica la extrañeza de este tipo de condicional:

The bizarreness seems to be attributable to the fact that the relation of dependence (and therefore also the temporal order) between the states of affairs specified in the antecedent and consequent is reversed.[27]

Ahora bien, uno puede estar tentado a parafrasear este tipo de condicionales de manera que no resulten tan contraintuitivos, por ejemplo si tenemos:

3) Si Juan hubiera llegado a tiempo al trabajo, el habría salido de su casa temprano.

Parece que el hecho de haber llegado a tiempo provocaría el hecho previo de que Juan salió de su casa, pero esto no es el caso para la siguiente reformulación:

4) Si Juan hubiera llegado a tiempo al trabajo eso habría sido porque salió temprano de su casa.

Sin embargo tal paráfrasis convierte el condicional inicial en un tipo de condicional totalmente diferente, en palabras de Woods:

Such sentences do not say that if something had been the case at one time, something else would have been the case earlier; they say that if something had been the case, it

[27] Woods, Michael. Conditionals. Oxford University Press. Oxford. 1997. P.51

would (then) have had a certain explanation, or that (then) something would have needed to have happened earlier.[28]

Lewis justifica el por qué este tipo de condicionales que expresan una asimetría temporal son rechazados en el modelo de mundos posibles que propone argumentando lo siguiente: dado que al evaluar un condicional de este tipo postulamos un mundo hipotético que coincide con el mundo real hasta antes del antecedente, y que además temporalmente los sucesos que tienen lugar después de un momento dado dependen de él de una manera que los sucesos anteriores a ese momento no dependen; entonces en el aparato semántico que Lewis propone un condicional retroactivo sería imposible de analizar pues debido al requisito de que ambos mundos, el hipotético y el real, coincidan antes de ese periodo no es posible considerar las diferencias que habrían sucedido en el periodo antes del antecedente si este se hubiera cumplido.

En otras palabras, siendo que el consecuente de un condicional retroactivo hace referencia a diferencias entre el mundo hipotético y el mundo real en el periodo de tiempo previo a su antecedente, y que al suponer el mundo hipotético es justo en ese periodo de tiempo en el que el mundo hipotético y el real coinciden en todo aspecto, es claro cómo es que las condiciones de verdad dadas por el análisis de Lewis impiden considerar divergencias entre los mundos hipotético y real de la manera en que un condicional retroactivo lo solicita.

La noción de similitud entre mundos de Lewis, tal como la dibuja Woods es insatisfactoria para este último debido a que, como vimos, Lewis se ve obligado a admitir una desviación de las leyes naturales del mundo real en el mundo hipotético para así garantizar la verdad del antecedente en el mundo real y la máxima cercanía

[28] Woods, Michael. Conditionals. Oxford University Press. Oxford. 1997. P.52

de este con respecto del mundo real. Adicionalmente Woods apunta que si la relación de similitud (sobre la cual descansa el modelo de mundos posibles) no es satisfactoriamente definida, realmente no tenemos por qué estar atados a evaluar los condicionales contrafácticos en términos de mundos posibles y podemos en cambio buscar formular una teoría alternativa sin la necesidad de hacer referencia a posibilidades imaginarias que nos permita entender de mejor manera las particularidades de los contrafácticos:

Is the reference to possible worlds still essential? As Bennett recognizes, it depends whether the relation of similarity is taken as fundamental. If the relation of closeness between possible worlds is not explained in terms of similarity, the theory can be reformulated as a metalinguistic theory (...) why do we engage in the practice of considering possible worlds alternative to the actual in the way that we do? We need, first, to consider the reasons we have for examining, retrospectively, contrary-to-fact possibilities: what is the point of asking ourselves how things would have been if certain matters had been different? Further, we need to attend more closely to the specific features of 'counterfactual' conditionals: what is their logical form?[29]

Es por esto que Woods construye un modelo alternativo para explicar la semántica de los contrafácticos, no ya bajo el planteamiento intuitivo del test de Ramsey sino siguiendo una estrategia de análisis en la cual su pregunta central apunta a encontrar la forma lógica de los enunciados tradicionalmente llamados contrafácticos. Este nuevo planteamiento lo veremos en la siguiente sección.

[29] Woods, Michael. Conditionals. Oxford University Press. Oxford. 1997. Pp. 55 y 57.

3.3 La Teoría de Michael Woods de condicionales contrafácticos

Como vimos con anterioridad, Woods sostiene que no hay una distinción entre condicionales indicativos y contrafácticos, y rechaza lo que él llama la Tesis de la contrafactualidad, lo cual significa que rechaza que existen condicionales que presuponen la falsedad del antecedente en oposición a los condicionales indicativos en donde el antecedente puede ser verdadero o falso. Woods reconoce no obstante que cuando se usan condicionales que tradicionalmente se conocen como contrafácticos, el hablante sí que tiene generalmente la creencia de que el antecedente es falso, pero que dicha creencia no tiene que ver con la estructura del condicional como tal. Woods sí reconoce entonces una oposición entre estos condicionales que usualmente se les llama contrafácticos y los condicionales que él llama condicionales simples (que son condicionales en los que dos enunciados independientes ocupan la estructura de la forma 'Si, entonces') Sin embargo, estas clases de condicionales se oponen no por la falsedad de su antecedente sino en la medida en que son semánticamente diferentes dado que, Woods afirma, no son enunciados que tengan la misma forma lógica. De manera que su estrategia para el análisis semántico de los condicionales contrafácticos va a ser la de tratar de describir cuál es su forma lógica.

Al igual que Stalnaker y Lewis, Woods parte de una observación intuitiva para construir su modelo semántico. La hipótesis de Woods es que cuando una persona enuncia una oración como i) 'si la bomba hubiera explotado ayer habría muerto mucha gente' esa persona expresa lo mismo que expresaría otra persona que en el pasado dijera ii) 'si la bomba explota mañana, morirá mucha gente'. En otras palabras, Woods parte de la observación de que las dos oraciones expresan el mismo hecho aunque cada una lo hace desde una ubicación temporal distinta. La oración i) reporta un hecho acerca del pasado desde un momento posterior a la enunciación de

ii) y a su vez ii) reporta un hecho acerca del futuro desde un momento anterior a la enunciación de i).

Lo anterior significa para Woods 2 cosas: la primera es que en realidad las oraciones condicionales contrafácticas son versiones en tiempo pasado de oraciones condicionales prospectivas[30] como la oración ii) del ejemplo anterior. La segunda es que la forma lógica de las oraciones condicionales contrafácticas no es la de una oración condicional simple en donde dos oraciones ocupan la estructura 'Si, entonces'. Esto es así porque, como veremos después con más detalle, para Woods el tiempo verbal de la oración funciona en este caso como una especie de operador modal que gobierna sobre la oración completa, por lo cual la estructura condicional queda subordinada al tiempo verbal y por consiguiente en lo que respecta a condicionales contrafácticos ni si quiera podríamos hablar en sentido estricto del antecedente y consecuente del condicional. En palabras de Woods

…we may take the sentence containing 'would' as a past tense version of the one containing 'will'. If we take that position, it is already clear that a 'would'-sentence like 'If it had rained yesterday afternoon, the match would have been cancelled' is not straightforwardly of the 'If…the…' form, involving as it does the application of a past tense to something already of conditional form. (…) it is not constructed from

[30] En el texto en inglés el nombre que Woods da a este tipo de condicionales es 'will conditionals' debido a la presencia en el idioma inglés del auxiliar 'will' para indicar el tiempo futuro de la oración, como se puede ver si traducimos al inglés la oración ii) de nuestro ejemplo como ' If the bomb explodes tomorrow, many people will die'. Más adelante en su texto, ya cerca del final, también se refiere a ellos como 'forward looking conditionals', de ahí que yo haya elegido nombrar este tipo de condicionales como 'condicionales prospectivos'. Pienso sin embargo que es más afortunado para la teoría la versión en inglés debido a que Woods usa el auxiliar 'will' como operador lógico separado del resto de la oración de forma que es mucho más clara la estructura lógica que quiere mostrar, lo cual se pierde en la traducción al español.

two sentences, each independently past-tensed, by applying the 'If…then…' construction to them…[31]

Si consideramos lo anterior podríamos parafrasear la oración contrafáctica i) como:

'En el pasado fue cierto que: si la bomba explota mañana, morirá mucha gente'.

En donde la expresión "en el pasado fue cierto que" funciona de la misma manera que como funciona por ejemplo en lógica alética la expresión "es necesario que". En consecuencia, dado que la oración i) del ejemplo es en realidad la oración ii) bajo el dominio del tiempo verbal, debemos entonces para entender los condicionales contrafácticos, sostiene Woods, entender primero los condicionales prospectivos, y de la misma manera que con los condicionales contrafácticos para entender los condicionales prospectivos debemos preguntarnos por su forma lógica:

What is the logical form of such sentences? I suggest that they are best understood as containing an application of the present tense 'will' to a sentence or predicate that is already of the 'If…then…' form. If so, the final sentence will not itself be of that form. The 'If…then…' construction is the same, semantically as the one that we find in sentences of the 'If…then…' form built up from two arbitrary sentences, but we have to regard the 'If…then…' as having been introduced at an earlier stage in the construction of the sentence…[32]

Lo que Woods sostiene es que los condicionales prospectivos que son la clave para entender los condicionales contrafácticos son al igual que estos resultado de la aplicación de un tiempo verbal a una oración previa más simple con la estructura

[31] Woods, Michael. Conditionals. Oxford University Press. Oxford. 1997. P.79
[32] Woods, Michael. Conditionals. Oxford University Press. Oxford. 1997. P.81

condicional de la forma ‘Si, entonces’ por lo que a los condicionales prospectivos no puede considerárseles condicionales en el mismo sentido.

La tesis de Woods es entonces que, así como los contrafácticos, los condicionales prospectivos tampoco tienen la forma de un condicional de estructura ‘Si, entonces’ y que son en cambio oraciones compuestas de la aplicación de un tiempo verbal a un condicional simple. Podemos analizar la composición de los condicionales prospectivos de la siguiente forma:

Considérense las siguientes dos oraciones:

‘Pedro hace su tarea’

y

‘Pedro aprueba el curso’

Si tomamos esas dos oraciones y las introducimos dentro de la estructura ‘Si, entonces’ tenemos el siguiente condicional simple:

‘Si Pedro hace su tarea, entonces Pedro aprueba el curso’

Ahora, si a esa oración completa le agregamos la modalidad de tiempo verbal de la siguiente forma:

‘Será el caso que: Si Pedro hace su tarea entonces Pedro aprueba el curso’[33]

[33] En el texto original hay un paso intermedio entre el condicional simple y el condicional que empieza con ‘será el caso que’ en el que Woods usa el auxiliar ‘will’ para alterar la oración completa. Traduciendo nuestro ejemplo al inglés tendríamos el condicional simple ‘if Pedro does his homework, then Pedro passes the course’ y luego añadiendo ‘will’ como modalidad tendríamos ‘Will (if Pedro does his homework, then Pedro passes the course)’ que finalmente se lee de manera más natural como ‘it will be the case that if Pedro does his homework, then Pedro passes the course’. Si bien este paso no es indispensable para entender cómo Woods construye un

Podemos entonces parafrasear la oración resultante como:

‘Si Pedro hace su tarea entonces aprobará el curso’,

que es un condicional prospectivo. De esta manera, podemos ver que el tiempo verbal tiene efecto sobre la oración completa aún cuando en el español más natural su presencia sólo se nota en el consecuente y por ende es el “operador” principal tal como ocurriría con una oración que antes lleve el operador modal ‘necesariamente’.

[The sentence] ‘It will be the case that, if John catches the train, he arrives in time.’ (…) appears in natural English as: ‘If John catches the train, he will arrive in time,’ in which ‘will’ is present-tensed, and governs the whole sentence. I claim that the ‘will’, though it occurs as an auxiliary forming the main verb in the apodosis, in fact has as its scope a sentence or at least a predicate, so that it has wider scope than the grammatical ‘If’ clause. If ‘will’ is treated as a modal expression, we should not be surprised that, although it occurs grammatically as part of the apodosis of the ‘If…then…’ sentence, its scope is the whole conditional. Exactly the same phenomenon occurs with the modal expression ‘ncecessarily’ in: ‘If Peter has a driver’s licence, he is necessarily over seventeen.’…[34]

Así pues Woods concluye que tanto los condicionales contrafácticos como los condicionales prospectivos son enunciados complejos cuya estructura no es la de un enunciado de la forma ‘Si, entonces’ sino que son producto de la aplicación de un tiempo verbal a otro enunciado. La relación que guardan entre ellos estos dos tipos de enunciados nos permite entender el problema en torno al cual ha girado nuestra discusión en este trabajo. Su relación consiste en que ambos expresan desde una

---

condicional prospectivo a partir de un condicional simple, si deja más claro de qué manera Woods piensa en el auxiliar ‘will’ como un operador modal.

[34] Woods, Michael. Conditionals. Oxford University Press. Oxford. 1997. P.82

locación temporal distinta (pero refiriéndose cada uno al mismo momento en el tiempo) la presencia de una tendencia o una disposición, es decir, afirman que en un momento determinado se dan las condiciones para que un evento o hecho tenga lugar, aunque ese hecho o evento no sea el caso en ese momento, tomando en cuenta el cumplimiento previo de una condición.

If we wanted to sum up what is reported in the present by 'will', and in the past by 'would', we could say it is the presence of a tendency or disposition. If it is true that if it rains the match will be cancelled, matters are so disposed as to ensure cancellation in the event of rain.[35]

Es la afirmación en ambos condicionales de que hay una disposición que se puede cumplir la que lleva a Woods a concluir que en realidad enunciados como 'si la bomba explota mañana, morirá mucha gente' pertenecen a la misma categoría que enunciados como 'si la bomba hubiera explotado ayer habría muerto mucha gente', es decir, que ambos son enunciados 'contrafácticos' (o sea, enunciados de la clase de condicionales que tradicionalmente se han entendido como contrafácticos).

Apoyándonos en todo lo anterior, podemos ahora formular las condiciones de verdad para enunciados de tipo contrafáctico de la siguiente manera:

La oración: Si P hubiera sido el caso, Q habría sido el caso,

Es verdadera si y sólo si en un momento anterior (determinado por el contexto) a su enunciación, la oración:

Si P es el caso, Q será el caso, fue asertible desde una posición epistémica ideal

[35] Woods, Michael. Conditionals. Oxford University Press. Oxford. 1997. P.83

En otras palabras ‘Si P hubiera sido el caso, Q habría sido el caso’ reporta la existencia en un momento anterior a su enunciación de una disposición que basta para el cumplimiento de Q dado el cumplimiento previo de P. O en palabras de Woods:

A statement like [that] is about the past, relative to the time of utterance, but the specification of what is the case in the past involves some time in the future and also a time that is past relative to that.[36]

En esta formulación lo que hace verdadero al condicional contrafáctico no es como en la teoría de mundos posibles una situación imaginaria que tengamos que postular, en cambio, bajo este modelo explicativo lo que hace verdadero a un condicional contrafáctico es lo que de hecho es el caso, sólo que el estado de cosas que lo hace verdadero esta ubicado temporalmente en un momento distinto al momento de su locución. Bajo este modelo explicativo de los condicionales contrafácticos, si la oración ‘Si P hubiera sido el caso, Q habría sido el caso’ es verdadera, sabemos que en el estado de cosas relevante ubicado en un momento previo a su locución se dan las condiciones suficientes para el cumplimiento del consecuente de la oración ‘Si P es el caso, Q será el caso’ en el caso de que se cumpla su antecedente. Y sabemos que se dan esas condiciones porque ‘Si P es el caso, Q será el caso’ fue asertible desde una posición epistémica ideal, lo que significa que fue asertible desde el punto de vista de alguien que conoce todos los hechos relevantes acerca de cómo es el estado de cosas hasta el momento del antecedente, de manera que puede dar cuenta de la existencia de la disposición expresada en la oración.

Esto quedará claro si regresamos a uno de los primeros ejemplos que nos sirvió para mostrar la problemática de las condiciones de verdad de los condicionales

[36] Woods, Michael. Conditionals. Oxford University Press. Oxford. 1997. P.89

contrafácticos. Consideremos el siguiente par de condicionales, siendo el condicional I un condicional simple y II un condicional contrafáctico:

I) Si Cortázar no escribió Rayuela, alguien más lo hizo, y

II) Si Cortázar no hubiera escrito Rayuela alguien más lo habría hecho.

En el caso de I) podemos afirmar su verdad dado nuestra posición epistémica desde la cual sabemos que alguien de hecho escribió Rayuela. En este tipo de condicional nuestra perspectiva epistémica es una perspectiva factual o sea que el que enuncia el condicional de hecho tiene el conocimiento relevante para su aserción, en contraste con el condicional II) en el que la perspectiva epistémica del que enuncia es una perspectiva ideal o hipotética como veremos a continuación.

De acuerdo con la teoría de Woods el condicional II) es verdadero si y sólo si en algún momento anterior a que Rayuela fuera escrita, la oración:

III) Si Cortázar no escribe Rayuela, alguien más lo hará

Fue asertible desde el punto de vista epistémico de alguien que conoce todas las condiciones y hechos relevantes que aseguran el cumplimiento del consecuente de III) si es que se cumple también el antecedente, en otras palabras, de alguien que sabe que en ese momento en el tiempo existe la disposición para el cumplimiento del consecuente. Este punto de vista epistémico es ideal e hipotético, es decir que la teoría no afirma que de hecho alguien deba tener el conocimiento relevante para afirmar el consecuente. En palabras de Woods:

We are not envisaging an actual epistemic point of view –a state of knowledge that someone might actually have been in antecedently; rather, we are describing things from an ideal epistemic perspective.[37]

En cambio la perspectiva epistémica necesaria para la aserción de III) es hipotética desde el punto de vista y ubicación temporal del que enuncia el condicional II) y lo es porque al expresar la existencia de una disposición en el pasado y no de algo que de hecho fue el caso no podemos sino sólo reconstruir de manera ideal que en algún momento existieron las condiciones que permitieron la ocurrencia del consecuente.

Dado el modelo anterior, Woods ofrece soluciones alternativas a los problemas que él mismo señaló de la explicación de condicionales contrafácticos basada en mundos posibles. En primer lugar revisa el problema de cómo justificar el rechazo de los condicionales retrospectivos que parecen presentar problemas para el modelo de mundos posibles que defienden Stalnaker y Lewis. Estos condicionales quedan fuera del modelo de Woods porque para que tuvieran cabida en él tendríamos que afirmar que es posible tener una disposición de un estado de cosas para la ocurrencia previa del consecuente con respecto de la ocurrencia del antecedente como lo requeriría un condicional retrospectivo. Pero eso es imposible, sostiene Woods dada la dirección de la causalidad. O sea que, dado que un contrafáctico de la forma 'Si P hubiera sido el caso, Q habría sido el caso' es verdadero si en el pasado puedo afirmar 'Si P es el caso, Q será el caso' desde una perspectiva epistémica ideal y antes de que ocurra P, y puesto que sólo lo puedo afirmar si es que en ese momento en el tiempo hay una disposición que garantiza la ocurrencia de Q si P ocurre, entonces para el condicional contrafáctico retrospectivo 'Si Q hubiera sido el caso, P habría sido el caso' es

[37] Woods, Michael. Conditionals. Oxford University Press. Oxford. 1997. P.85

imposible asignar un valor de verdad en los términos antes descritos dada la dirección de la causalidad que impide que Q suceda en un momento anterior a P.

En segundo lugar, Woods sostiene que en su teoría quedan capturadas todas las propiedades semánticas de la noción de similitud de la teoría de Lewis sin los problemas que esta presupone (como la apelación a una desviación de las leyes naturales) y sin la necesidad de apelar a mundos posibles. En la teoría de Lewis, el mundo hipotético en el que se evalúa un condicional es un mundo en el que todo es exactamente como en el mundo real hasta antes de que se cumpla el antecedente y en el que se preservan también sus leyes naturales. Esto se traduce en el nuevo modelo semántico de acuerdo con Woods, en la forma de la perspectiva epistémica ideal desde la cual se enuncia el condicional prospectivo que reporta la existencia de la disposición que hace verdadero al condicional contrafáctico correspondiente. En esta posición epistémica hipotética todos los hechos relevantes para el cumplimiento del consecuente del condicional prospectivo hasta antes del cumplimiento del antecedente son accesibles al igual que las leyes naturales del mundo real. Este análisis evita la referencia a mundos posibles imaginarios porque lo que hace verdadero a un condicional contrafáctico en este modelo es lo que de hecho ocurrió y lo que de hecho fue susceptible de ocurrir de acuerdo a las leyes naturales.

Finalmente Woods vuelve al problema que para la semántica de mundos posibles presentaba el manejo de contrafácticos con el antecedente imposible. En este la teoría de Woods no se presta tan claramente al análisis de condicionales cuyo antecedente es la negación de una verdad matemática por ejemplo, puesto que tal negación no es un evento en particular que ocurre en un momento en el tiempo. Sin embargo sí podemos ver cómo en este modelo, dado que no estamos presuponiendo la falsedad del antecedente, y en cambio consideramos la disposición para la ocurrencia de otro

hecho en caso de que este sea verdadero, no estamos obligados a aceptar que un condicional contrafáctico es verdadero por vacuidad y en contraste podemos examinar y derivar que consecuencias tendría la verdad de un antecedente imposible si quisiéramos por ejemplo hacer un razonamiento por reducción al absurdo.

Habiendo expuesto su teoría para el análisis de condicionales contrafácticos y habiendo señalado también cómo esta resuelve los problemas que la semántica de mundos posibles para contrafácticos acarreaba, Woods se limita a señalar a manera de conclusión por qué es preferible en términos generales su explicación a la de Stalnaker y Lewis o cualquier otra que haga uso del aparato semántico de mundos posibles, con lo cual responde a la pregunta que planteó al principio acerca de por qué explicar los contrafácticos en términos de situaciones imaginarias. Para Woods postular mundos posibles para la evaluación de condicionales contrafácticos sugiere un interés en construir fantasías en las que detallamos como pudieron ser las cosas si aspectos del mundo hubieran sido diferentes. En cambio si movemos el núcleo de la explicación de los condicionales contrafácticos a la explicación de los condicionales prospectivos los cuales tienen que ver con cómo será el mundo en el futuro dadas ciertas condiciones entonces el análisis de los contrafácticos nos será útil en última instancia para anticipar posibilidades futuras y planear para contingencias que pudieran suceder. Bajo el modelo de Woods pues, preguntarse por los condicionales contrafácticos es preguntarse por disposiciones que existieron en el pasado y al dar cuenta de los condicionales que en el pasado reportaban la existencia de estas disposiciones, adquirimos también conocimiento acerca de los condicionales que en el presente reportan la existencia de disposiciones y condiciones para la ocurrencia de hechos futuros.

3.4 Crítica a las tres tesis de Michael Woods

A lo largo de este capítulo hemos cuestionado y puesto a prueba (siguiendo el trabajo de Michael Woods) la idea de una semántica basada en el aparato formal de mundos posibles como análisis de los condicionales contrafácticos. Woods se ha encargado no sólo de señalar problemas inherentes a la teoría sino también de mostrar que el planteamiento general de una semántica de mundos posibles para contrafácticos es erróneo, producto de un entendimiento equivocado de las propiedades semánticas de los condicionales contrafácticos, Woods ha sostenido que los condicionales contrafácticos no son enunciados con estructura 'si, entonces' y que son en cambio versiones en pasado de enunciados que declaran la existencia de una disposición en un estado de cosas, lo que lo lleva a afirmar en última instancia que en realidad no hay tal cosa como la *contrafactualidad*, es decir que no existe una clase de enunciados condicionales que se comprometen con la falsedad de su antecedente. Ahora bien, ¿es exitoso el ataque de Woods? ¿Ha tirado Woods el modelo Stalnaker-Lewis? O por el contrario ¿se mantiene en pie el análisis basado en mundos posibles? En la sección final de esta tesis voy a tratar de contestar estas preguntas evaluando primero las objeciones de Woods para después contrastar las explicaciones de Woods, Stalnaker y Lewis para finalmente dictaminar cuál de entre ellas es la mejor explicación de los condicionales contrafácticos.

Comenzaré con los problemas del presupuesto de unicidad y la asunción del límite. Al respecto de estos problemas Woods realmente no esgrime un argumento muy elaborado ni muy convincente y basa su objeción en consideraciones muy generales (véase el ejemplo que pone acerca de las maneras en que me puedo encontrar con un conocido que cito textualmente en 3.2). Los problemas que estos supuestos acarrean sí son un reto a la teoría de mundos posibles pero Woods no los

fórmula de manera que se muestren sus repercusiones. Y de la misma manera, la solución que propone (también citada en 3.2) es más bien un esbozo de lo que podría salvar la idea de que haya un solo mundo que es más similar al mundo real al evaluar un contrafáctico. En la sección 2.3 de esta tesis expongo de manera más profunda las implicaciones que el supuesto de unicidad conlleva y la relación que este tiene con el principio de tercero excluso condicional, así como la reformulación de la teoría de Stalnaker por parte de Lewis usando nociones modales y cuantificacionales para rechazar la idea de que hay un único mundo más similar al mundo base, preservar el principio de tercero excluso condicional y tener una teoría funcional. Woods también menciona que Lewis reformula las condiciones de verdad para contrafácticos y que esa reformulación no se compromete con el supuesto de unicidad y el supuesto del límite, lo cual evita los problemas que estos supuestos acarrean, sin embargo Woods rechaza el planteamiento de Lewis porque este implicaría la falsedad del principio de tercero excluso condicional. Lo que Woods no considera es que para Lewis hay dos tipos de contrafácticos interdefinibles: los contrafácticos que afirman que algo sucede necesariamente dada una hipótesis contrafactual y los que afirman que algo sucede posiblemente. Y que con estos dos condicionales contrafácticos es posible formular exitosamente el principio de tercero excluso condicional (véase el apartado 2.3 de este trabajo). De manera que en lo que respecta a los presupuestos del límite y de unicidad, la objeción de Woods no se sostiene. Ahora bien en lo que respecta al supuesto de que el mundo más cercano al mundo que de hecho es el caso es idéntico al mundo que de hecho es el caso, Woods no elabora más respecto de cómo es que nos obligaría a aceptar como verdaderos condicionales sin relación alguna entre antecedente y consecuente, ni tampoco provee ejemplos que hagan patente lo que sostiene. Acerca del problema de considerar verdaderos a todos los contrafácticos con antecedente lógicamente imposible tengo que decir que en el caso del razonamiento

por reducción al absurdo no funciona en la manera que Woods lo presenta. Cuando se razona suponiendo la verdad de una contradicción lógica lo que se pretende es derivar otra contradicción para mostrar justamente lo contrario a lo que se había asumido, en ese caso el condicional en donde el antecedente es la hipótesis imposible y el consecuente es la contradicción derivada es verdadero, pues de hecho la contradicción sí se seguiría de la hipótesis. En el caso de que no fuera posible derivar una contradicción de la hipótesis eso implicaría que la hipótesis inicial no es en realidad lógicamente imposible y entonces no tendríamos problema alguno.

La objeción que sí plantea un reto para el modelo Stalnaker-Lewis es la que Woods hace a la noción de similitud. Plantear la similitud del mundo hipotético con respecto del mundo real exigiendo que ambos mundos sean exactamente iguales hasta el punto en que sucede el antecedente respetando las leyes naturales sí acarrea los problemas que Woods apunta y tales dificultades son insalvables a menos que se modifique la noción de similitud propuesta, más aún, la solución propuesta (la ocurrencia del "milagro pequeño") implica aún más problemas para la noción de similitud como ya vimos en 3.2. Si hemos de conservar el modelo Stalnaker-Lewis, los requerimientos de similitud del mundo hipotético deben ser otros, de manera que no tengamos que comprometernos con algo tan cuestionable como una violación hipotética a las leyes naturales por menor que sea. Woods sin embargo, toma estas dificultades como punto de partida para replantear totalmente el modelo semántico con el que entendemos los condicionales que, según él, erróneamente llamamos contrafácticos.

La teoría de Woods para condicionales contrafacticos es ciertamente atractiva. La idea de que lo que hace verdadero a un contrafáctico es una disposición que existió en el pasado y que es reportada por un condicional que habla acerca del futuro es un

planteamiento muy inteligente, aunado a que su retórica acerca de lo que en esencia implica preguntarse por condicionales contrafácticos y postular mundos posibles sí que provoca el cuestionamiento de la importancia de comprometerse con una tarea como esa: ¿es acaso el postular mundos posibles contrafácticos equivalente a sumergirnos en fantasías fútiles acerca de cómo hubieran podido ser las cosas en circunstancias diferentes? Mi respuesta es que no, que la referencia a mundos posibles es necesaria. Contrario a lo que sostiene Woods, yo voy a sostener que el modelo Stalnaker-Lewis basado en la semántica de mundos posibles es una mejor explicación de las propiedades semánticas de los condicionales contrafácticos que su teoría. Más específicamente, voy a defender que las tres tesis que considero las tesis centrales de la teoría de Woods de enunciados contrafácticos son falsas. Estas tesis son: i) No hay condicionales que reporten situaciones contrafactuales lo que hay son condicionales que reportan disposiciones, ii) lo que hace verdadero a un condicional que reporta disposiciones son los hechos no un estado de cosas hipotético, de lo cual se deriva la tesis de que iii) no es necesario apelar a mundos posibles para saber cuándo un condicional de los que usualmente se denominan contrafácticos es verdadero.

La tesis i) es la tesis que expresa el rechazo de Woods a la distinción tradicional entre condicionales indicativos y subjuntivos (contrafácticos). En el texto de Woods no hay argumento explícito para rechazar esta distinción, pero en el transcurso del desarrollo de su teoría se entiende que lo hace porque considera que los contrafácticos son versiones en pasado de los condicionales prospectivos (los que Woods llama en el texto en inglés ‘will conditionals’) es decir que los condicionales contrafácticos son reductibles a condicionales que no presuponen la falsedad del antecedente y que en cambio parecen más bien expresar una disposición de lo que puede pasar si se cumple el antecedente. En este planteamiento no hace falta

comprometerse con la falsedad del antecedente que parece sugerir el condicional contrafáctico, que es a su vez lo que motiva a un análisis semántico mediante mundos posibles, de manera que al no haber condicionales que reporten situaciones contrafactuales sino más bien condicionales que reportan disposiciones la semántica de mundos posibles es dispensable. Pero no todo condicional contrafáctico es la versión en pasado de un condicional prospectivo. Considérese el condicional 'si fuera astronauta, entonces viviría en la Luna', este condicional es claramente diferente al condicional 'si hubiera sido astronauta entonces viviría en la Luna' en tanto que el primero no es acerca del pasado sino que claramente habla del presente en contraste con el segundo que reporta algo acerca del pasado. Tampoco es un condicional prospectivo o alguna versión de este, en tanto que no declara cómo serán las cosas en el futuro de cumplirse el antecedente, claramente el consecuente es acerca de un hecho en el presente. Pero más importante, es un condicional contrafáctico. Ciertamente reporta una disposición, nos dice que el estado de cosas actual es tal que de ser verdadero el antecedente, el consecuente sería verdadero también, sin embargo sí que hay un compromiso con la falsedad del antecedente en el presente, en el momento de la locución del condicional. Cuando digo 'si fuera astronauta, entonces viviría en la luna' estoy dando a entender que en realidad no soy astronauta, que ese no es el caso, pero que en caso de que lo fuera viviría en la luna, o lo que es lo mismo, estoy diciendo que cierta disposición se cumpliría si una situación contrafactual se cumpliera, si algo que de hecho no es el caso lo fuera. Un condicional de este tipo reporta a la vez una disposición y una situación contrafactual. Por lo tanto, sí hay condicionales contrafácticos. Por lo que la tesis i) es falsa.

La tesis ii) es la razón por la cual Woods se siente atraído por reducir los condicionales contrafácticos a condicionales disposicionales: Si lo que hay son condicionales disposicionales y estos son verdaderos en virtud de los hechos entonces

no es necesaria la referencia a mundos posibles para saber si un condicional de este tipo es verdadero. Lo que sostiene Woods es que un condicional que reporta una disposición (como lo es, de acuerdo con Woods, 'si hubiera sido astronauta viviría en la Luna') es verdadero o falso dependiendo de si los hechos son tales que aseguran o no el cumplimiento del consecuente en el caso de que se cumpla el consecuente. Sin embargo al tratarse de condicionales que reportan disposiciones (o sea tendencias de cómo podría ser el mundo aunque en el momento no sea así) y no estados de cosas que de hecho son el caso, lo que los hace verdaderos es lo que potencialmente puede suceder y no lo que en realidad sucede. Es decir, por mucho que los hechos tal y como son en el momento relevante aseguren la ocurrencia de ciertos otros hechos, lo que hace verdadero a un condicional disposicional es la propensión de las cosas que son el caso a ser de otro modo y no las cosas mismas. En otras palabras apelar a disposiciones es también apelar a situaciones imaginarias en tanto que son situaciones no actualizadas. Tanto para un condicional contrafáctico como para un condicional disposicional se cumple que al preguntarse por sus condiciones de verdad se pregunta por un estado de cosas posible y no por uno actual. Por lo que la tesis ii) es falsa y la tesis iii) por consiguiente también debido a que preguntarse por un estado de cosas posible es preguntarse por un mundo posible.

Adicionalmente quiero defender que el rol de la posición epistémica ideal de la teoría de Woods es el mismo que juega la idea de similitud en el modelo Stalnaker-Lewis pues ambas son herramientas teóricas para lidiar y dar cuenta de todos los otros hechos adicionales al antecedente que aseguran el cumplimiento del consecuente de un condicional contrafáctico. Lo que nos deja en la posición de poder afirmar que la teoría de Woods en realidad apela a las mismas intuiciones que el modelo Stalnaker-Lewis solamente que reformulado en otros términos para evitar hacer referencia a mundos posibles. Sin embargo como las intuiciones son las mismas

Woods no pudo escapar, a mi juicio, de hacer referencia a situaciones no actualizadas, las cuales son mejor manejadas por medio de la semántica de mundos posibles. Intuyo que Woods pretendía elaborar una explicación más económica que pudiera evadir las complicaciones de la semántica de mundos posibles, aunque para ello tuvo que tirar nociones tan básicas como la noción de contrafactualidad e incluso negar que los condicionales contrafácticos fueran, por su forma lógica, enunciados condicionales en absoluto. Pero aún cuando pueda ser muy compleja la semántica de mundos posibles y aún cuando nociones como la de similitud necesiten ser pulidas, pienso que las características semánticas de los condicionales contrafácticos son tales que es preferible permanecer dentro de este modelo y tratar de refinarlo que buscar construir un paradigma alternativo.

Conclusiones

A lo largo de este trabajo hemos visto cómo los condicionales contrafácticos del lenguaje natural no comparten las condiciones de verdad del condicional material de la lógica clásica, lo cual motivó la búsqueda de una teoría que diera cuenta de esas condiciones y que explicara la semántica de los condicionales que implícitamente presuponen la falsedad de su antecedente. Las teorías que elegí para su análisis fueron las teorías de mundos posibles de Robert Stalnaker y David Lewis pues la literatura que revisé constantemente apuntaba al modelo Stalnaker-Lewis como el punto de partida de la manera en que se entienden los condicionales contrafácticos en la actualidad, y la teoría de Michael Woods de contrafácticos porque rechaza y cuestiona este modelo, lo cual me permitió ponerlo a prueba. En el tándem compuesto por las teorías de Stalnaker y Lewis tenemos que para saber cuándo un condicional contrafáctico es verdadero en el mundo real, se postula un mundo posible hipotético (o múltiples mundos en el caso de Lewis) el cual es el mundo más parecido posible al mundo real en el que el antecedente del contrafáctico es verdadero. Si en ese mundo hipotético el consecuente es verdadero, el condicional será verdadero y si el consecuente es falso el condicional será falso. Puntualmente, señalamos los problemas inherentes al modelo explicativo basado en mundos posibles y tratamos de buscar soluciones a esos problemas: vimos en 2.3 que la teoría de Stalnaker presupone que hay un único mundo posible hipotético que es el más similar al mundo base en el que se enuncia un contrafáctico, vimos cómo esta suposición es insostenible y mediante la reformulación de Lewis vimos como eso no representaba una dificultad si admitíamos dos tipos de condicionales contrafácticos: los que su consecuente es verdadero necesariamente y los que su consecuente es verdadero posiblemente, dada la suposición contrafáctica del antecedente, lo cual también nos permitía formular el principio de tercero excluso condicional. Posteriormente vimos

cómo hay cadenas inferenciales que en la lógica clásica son válidas mientras que en la semántica de mundos posibles resultan inválidas, debido a que, por el requerimiento de seleccionar el mundo máximamente similar para cada contrafáctico de la cadena, asignamos mundos hipotéticos diferentes para cada uno de ellos. Dado que es deseable, por la manera en que razonamos cotidianamente, que estas cadenas sean inferencias aceptables, tratamos de establecer ciertos requerimientos, siguiendo a Gladys Palau, basados en la noción de un contexto contrafáctico para que la asignación de mundos hipotéticos sea una asignación relevante al contexto inferencial. Después en el tercer capítulo analizamos la teoría de Woods. Woods sostenía que podíamos replantear la semántica de los contrafácticos sin recurrir a los mundos posibles debido a que los contrafácticos son en realidad versiones en pasado de condicionales prospectivos y que estos lo que hacen es reportar una disposición para el cumplimiento del consecuente, dada la verdad del antecedente. De acuerdo con Woods los mundos posibles son dispensables porque lo que hace verdadero a un condicional disposicional es lo que de hecho es el caso y no lo que podamos suponer de manera hipotética. Por último defendí la semántica basada en mundos posibles para los condicionales contrafácticos argumentando que no todo condicional contrafáctico es en realidad una versión en pasado de un condicional prospectivo disposicional. Más aún, defendí que hay condicionales que claramente suponen la falsedad de su antecedente. Defendí también que tanto los condicionales contrafácticos como los condicionales disposicionales son verdaderos en virtud no de lo que es el caso sino de un estado de cosas posible, lo cual hace necesaria la referencia a mundos posibles para formular sus condiciones de verdad.

Desde la introducción de esta tesis me planteé la tarea de responder la pregunta de cuáles son las condiciones de verdad de los condicionales contrafácticos, pero también la de determinar si el modelo Stalnaker-Lewis se sostenía después de

cuestionarlo. Creo que aún no logro responder la pregunta por las condiciones de verdad de los condicionales contrafácticos pero si he logrado defender el modelo Stalnaker-Lewis. Pienso que he sido capaz de mostrar que el modelo Stalnaker-Lewis es un modelo viable para explicar la semántica de los contrafácticos y que es una mejor teoría de los condicionales contrafácticos que la de Michael Woods por las razones expuestas en 3.4. Si bien la pregunta de cuáles son las condiciones de verdad de los enunciados contrafácticos queda sin respuesta definitiva, así como la cuestión de si puede haber un análisis que los abarque a todos o si debemos usar teorías distintas para afirmaciones distintas. En el transcurso de esta tesis he señalado algunas de las dificultades de este modelo y creo que he podido responder a cada una de estas dificultades (como se vio en 2.3, 2.4 y 3.4) de manera que el modelo Stalnaker-Lewis se sostiene. También hice patente en 3.4 que el recurrir a la herramienta de los mundos posibles no sólo surge de una intuición como la del test de Ramsey sino que además es necesario apelar a estados de cosas hipotéticos para saber cuándo un condicional contrafáctico es verdadero o falso. Cierto es que la adopción del modelo Stalnaker-Lewis acarrea problemas que aún deben solucionarse (como los problemas de la noción de similitud, que dada su dependencia del contexto de enunciación bien podría servirse de herramientas de lógicas relevantistas) sin embargo es preferible el camino de los mundos posibles y no el que Woods propone el que debemos transitar si aspiramos a entender de manera adecuada los condicionales contrafactuales.

Bibliografía

-Arlo-Costa, Horacio. *The Logic of Conditionals*. The Stanford Encyclopedia of Philosophy. (Winter 2016 Edition) Edward N. Zalta (ed.) URL= <https://plato.standford.edu/archives/win2016/entries/logic-conditionals/>

-Bennett, Jonathan Francis. A Philosophical Guide to Conditionals. Oxford University Press. Oxford, 2003

-Gomez Torrente Mario. *La teoría semántica de las lógicas condicionales de Stalnaker y Lewis*. URL= https://minerva.usc.es/xmlui/bitstream/10347/1036/1/pg_059-080_agora12-2.pdf

- Goodman N. "The Problem of Counterfactual Conditionals", The Journal of Philosophy, Vol. 44, No. 5.

- Harper, William L. (et al.) *A sketch of some recent developments in the theory of conditionals*. En Ifs. Springer, 1980.

-Jackson, Frank. Conditionals. Blackwell. Oxford, 1987.

-Kripke, S, *Semantical Analysis of Modal Logics, I.* Zeitschrift für mathematische Logik und Grundlagen der Matematik, vol. 9. 1963.

-Lewis, David. Counterfactuals. Blackwell. Oxford, 1973.

-Lewis, David. *Counterfactuals and comparative possibility*. En Ifs. Springer, 1980

-Lewis, David. *Counterfactual Dependance and Time´s Arrow*, Nous, 13 (1979), reimpreso en Lewis, *Philosophical papers*, ii.

- Palau Gladys. *Condicionales contrafácticos: condiciones de verdad y semántica de mundos posibles. Acerca de las teorías de R. Stalnaker y D. Lewis*. En Crítica, Volúmen 12, Número 34. IIF, UNAM.

- Quine, W.V.O. Methods of logic. London, 1974.

-F.P. Ramsey, *General propositions and Causality*. En Ramsey, F.P. Foundations of Mathematics and other Logical Essays. New York, 1950.

-Sanford, David H. If P then Q: Conditionals and the foundations of reasoning. Routledge. London, 2003.

-Stalnaker, Robert. *A Theory of conditionals*. En Rescher, Nicholas. Studies in logical theory. Oxford, 1968.

-Woods, Michael. Conditionals. Oxford University Press. Oxford. 1997.