

CONTRAFÁCTICOS

IGAL KVART

UNIVERSIDAD HEBREA DE JERUSALÉN

En este artículo presento una aproximación a los contrafácticos basada en la noción de probabilidad objetiva. La misma se halla en el espíritu del modelo metalingüístico, aunque no cae totalmente dentro de él. Por ello, no está desarrollada en términos de mundos posibles o de las nociones que les son parasitarias (p. ej., similitud). Sus rasgos dominantes están arraigados en la probabilidad objetiva y la relevancia causal (analizadas desde la perspectiva probabilística) y, por lo tanto, no se acercan al enfoque de similitud máxima o de cambio mínimo.

1. *La función p.i.*

Si se toma seriamente el carácter implicatorio de la construcción contrafáctica: "Si hubiera sido el caso que A , entonces habría sido el caso que B " (simbolizada como $A > B$), ésta tendría que ser interpretada como una implicación incompleta. Bajo esta interpretación, la cuestión central es: ¿cuáles son las proposiciones que, asociadas con el antecedente, llevan al consecuente si y sólo si el contrafáctico es verdadero? Si las condiciones de verdad de la proposición contrafáctica $A > B$ son función de $A > B$ (aunque no necesariamente exclusivas de ella), una formulación más precisa de la cuestión sería: ¿cuál es la *función* f que, teniendo un contrafáctico como uno de sus argumentos y conjuntos de proposiciones como sus valores, caracteriza de tal modo a la construcción contrafáctica que un contrafáctico es verdadero si y sólo si los valores de f para los argumentos adecuados, junto con el antecedente, implican [*entail*]* el consecuente? Suponiendo que en este contexto las implicaciones [*entailments*] de-

* Cuando "implicar" sea traducción del verbo "*to entail*", pondremos este término inglés entre corchetes; en otro caso será traducción de "*to imply*". Lo mismo respecto del sustantivo correspondiente. [*N. de la T.*]

ben ser nomológicas, es decir, ajustadas a las leyes de la naturaleza,¹ el último debe estar dado entonces por el valor de f . De esta manera, usando el signo " \rightarrow " para la relación de consecuencia lógica, un contrafáctico $A > B$ es entonces verdadero si y sólo si

$$(1) f(A > B, \dots) \cup \{A\} \rightarrow B,$$

lo cual evidencia lo que debe ser denominado el *Esquema Inferencial* de los contrafácticos.²

Esta caracterización esquemática de las condiciones de verdad para los contrafácticos reduce así el problema a la determinación de la función f . Ésta será denominada la *función de las premisas implícitas* (abreviado: la función p.i.) y sus valores para argumentos fijos se llamarán premisas implícitas (de esos argumentos). Los enunciados de f , además de $A > B$, constituirán de este modo los otros parámetros para los cuales las condiciones de verdad de los contrafácticos son una función.

2. Contrafácticos del tipo *d.n.*

La subdivisión natural del problema de determinar las premisas implícitas de un contrafáctico dado indicaría que se consideren por separado el curso real de los acontecimientos anterior al tiempo del antecedente t_A ³ y el curso real de los acontecimientos posterior a ese momento.⁴ En la primera parte de este artículo me concentro en este último, y trato al primero en la sección 13. Tomo la historia del mundo antes del intervalo t como el conjunto de todas las proposiciones (del lenguaje) verdaderas fácticas (es decir, no legaliformes) que se refieran a tiempos enteramente anteriores al comienzo de t .⁵ No obstante, para un importante grupo de contrafácticos, aquellos

¹ Nuestro interés en este artículo se centra en los contrafácticos, no en los contralegales. Para una discusión sobre éstos, cf. [6], cap. 9, IV 2.

² Se puede considerar la posibilidad de debilitar la relación de consecuencia lógica y reemplazarla por una alta probabilidad condicional de B en $f(A > B, \dots)$ y A ; pero no voy a analizar esta posibilidad en este artículo. Para un debilitamiento de este esquema y, por lo tanto, del modelo metalingüístico, véase más abajo, secciones 13 y 16.

³ En términos más precisos: antes del inicio del intervalo t_A . Éste sería el intervalo al que pertenece la proposición A , es decir, el intervalo entre cuyos límites se dice que sucede (en A) el acontecimiento (o situación) descrito por A . Esta especificación de t_A determina el dominio cuantificacional de la variable temporal en una expresión formalmente adecuada de A y se hallaría explícita en A o se la comprendería por medio del contexto. Cf. también [6], cap. 6, I, II.

⁴ Para precisar más: después del inicio del intervalo t_A .

⁵ En este artículo sólo se trata de un lenguaje natural. Desde luego, este concepto se comprende con mucha mayor facilidad en el lenguaje formal.

cuyos antecedentes son compatibles con la historia del mundo que los precede, podrá argumentarse que hay que dejar intacta la historia real anterior al tiempo del antecedente cuando se considera al contrafáctico,⁶ con lo cual únicamente se dejaría abierto el problema de determinar las premisas implícitas del contrafáctico después del tiempo del antecedente. (La compatibilidad con la historia anterior es consistencia con ella junto con las leyes de la naturaleza.) Normalmente, dentro de este grupo de contrafácticos (que llamaremos del *tipo de divergencia natural* o *tipo d.n.*⁷) caen aquellos cuyos antecedentes describen acciones humanas. Como ejemplo (*ejemplo 1*), consideremos a Ralph, que participó en un espectáculo de concursos. Había dos puertas. Detrás de una de ellas estaba un automóvil último modelo. Si Ralph adivinaba cuál era la puerta correcta ganaría el automóvil. Puesto que no tenía ni la menor idea de cuál era la puerta correcta, eligió una arbitrariamente. (Supongamos que su elección no estuvo determinada por factores psicológicos o de otra índole.) Resultó que el automóvil se encontraba detrás de la puerta que Ralph no eligió.

Al considerar el contrafáctico: "Si Ralph hubiera elegido la otra puerta, habría ganado el automóvil", resulta muy claro que la historia anterior del mundo es compatible con el antecedente. Este contrafáctico cae por lo tanto directamente dentro del tipo d.n. De tal manera que, al concentrarnos en los contrafácticos del tipo d.n., abordamos al grupo de contrafácticos más común y (yo diría) el *más* útil. Por esta vía resolveríamos una de las ramificaciones de la citada subdivisión del problema general. (En la sección 13 se trata la otra ramificación de una manera que constituye la extensión natural de la forma en que se trata más abajo la primera ramificación.)

Se pueden hacer algunos comentarios con respecto a las dudas sobre la compatibilidad de diversas proposiciones que describen acciones humanas hipotéticas con la historia previa del mundo. En primer lugar, el alcance de este grupo de proposiciones es amplio para quienes sostienen que el mundo no es demasiado determinista (en los sentidos adecuados) y que es posible que los seres humanos, a nivel general y en virtud de lo anterior, actúen de manera diferente a como efectivamente lo hacen. (Para quienes no sostienen este punto de vista, estos contrafácticos podrían requerir entonces del tratamiento de ambas ramificaciones del problema mencionado

⁶ Pero cf. n. 56, más abajo.

⁷ La elección de este nombre se debe a la interpretación por analogía según las líneas de las descripciones de mundos posibles (cf. más abajo, sección 10): para estos contrafácticos, las descripciones de mundos posibles que son relevantes para su evaluación "divergirán" del mundo real al comienzo de t_A .

al principio de esta sección.⁸) En segundo lugar, la forma en que los hablantes de la lengua procesan contrafácticos particulares en sus marcos de referencia conceptuales bien podría estar cargada de preconcepciones concernientes a diversas características del mundo, incluyendo su grado de determinismo. Ahora que, naturalmente, nosotros esperamos que el resultado de nuestro análisis coincida con las intuiciones reflexivas de estos hablantes y refleja la manera en que procesan en la práctica las proposiciones contrafácticas. Pero entonces, con el fin de lograr la adecuación extensional, los presupuestos externos (como los que se refieren al grado de determinismo de las acciones humanas), que afectan a cuestiones relativas al alcance de un análisis particular (el presupuesto mencionado determina el alcance del conjunto de contrafácticos del tipo d.n.), deben reflejar los correspondientes principios del marco conceptual del sentido común. Por lo tanto, si la explicación de los contrafácticos del tipo d.n. que aquí se propone tiene que reflejar la forma en que los hablantes procesan esos contrafácticos en la práctica, y si esta explicación refleja la forma en que los hablantes procesan contrafácticos cuyos antecedentes *según su juicio* son compatibles con la historia previa, entonces el alcance de esta explicación debe ser el dominio de los casos en que, a la luz del marco conceptual del sentido común, el antecedente es compatible con su historia previa. Pero el presupuesto de que muchas acciones humanas se hallan subdeterminadas en un grado importante por la historia anterior y de que (en su momento) era posible tomar acciones opcionales en su lugar, constituye, al margen de su verdad, uno de esos principios del sentido común y, por lo tanto, tiene que desempeñar un adecuado papel clasificatorio en nuestro modelo cuando confrontamos los ejemplos particulares a la luz de las intuiciones.⁹ (Como lo que buscamos son las condiciones justificadas de aseverabilidad de los contrafácticos en lugar de sus condiciones de verdad, evidentemente muchas acciones humanas se hallan subdeterminadas de manera suficiente por el trasfondo epistémico que poseemos y de aquí que la aplicabilidad del tratamiento presentado en este artículo, con modificaciones obvias, sea amplia.)

⁸ Sin embargo, el tratamiento de los contrafácticos cuyos antecedentes no son compatibles con su historia previa no cae dentro del esquema inferencial, sino que, más bien, requiere de una modificación de éste; *cf.* más abajo, sección 13.

⁹ Y de hecho, este supuesto subyace frecuentemente a diversas discusiones del sentido común en las cuales los contrafácticos desempeñan un papel importante, como las concernientes a la responsabilidad del individuo sobre sus propias acciones, o a las evaluaciones históricas, o a la culpabilidad legal, en las que se da por supuesto que el sujeto tenía la posibilidad de actuar de otra manera.

En consecuencia, los contrafácticos del tipo d.n. figuran entre los más útiles y comunes. El hecho de que, en su evaluación, se preserve la historia previa no debe ser contemplado como si, al ocuparnos de ellos por primera vez, estuviéramos evitando las principales dificultades del análisis de los contrafácticos. El problema de determinar las premisas implícitas del intervalo (t_A, t_B) es difícil de resolver, como se verá en las siguientes secciones. Además, la respuesta a este problema también forma parte, en gran medida, de la respuesta al problema equivalente en el análisis de los contrafácticos que no son del tipo d.n., y también desempeña cierto papel en la respuesta al problema de las premisas implícitas anteriores al tiempo del antecedente (para contrafácticos que no son d.n.). Así mismo, el peligro del callejón sin salida al que Goodman llegó en su investigación se manifiesta plenamente en los contrafácticos d.n., lo cual demuestra cuán amplia es la porción de la tarea general de analizar contrafácticos que comprende su parte. Finalmente, la mayoría de las fallas que encuentro en el análisis de Lewis se evidencian en los casos de contrafácticos d.n. Este tipo por sí solo hace surgir algunos de los problemas más difíciles que afronta el análisis general de los contrafácticos.

Por lo tanto, voy a concentrarme en los contrafácticos del tipo d.n. hasta la sección 13; en ésta muestro cómo generalizar su análisis a contrafácticos de otros tipos.

3. *Contrafácticos - Análisis I*

Un supuesto contrario a los hechos, que constituye el antecedente de un contrafáctico, exige una selección de rasgos de la historia real del mundo, con base en los cuales va a evaluarse la predicción hipotética que constituye la significación de la proposición contrafáctica. La determinación de estos rasgos es el papel desempeñado por la función de las premisas implícitas f en el anterior esquema (1). Considerar el antecedente A de un contrafáctico es considerar el impacto que tiene una transición desde el curso real de los acontecimientos, el cual incluye el suceso $\sim A$, hacia un curso de acontecimientos alternativo A . Hay que considerar este impacto sobre el trasfondo de los rasgos del mundo real que queremos retener. ¿Cuáles van a ser estos rasgos? En el caso del tipo de contrafácticos en discusión, son los que describen a esos y sólo esos acontecimientos que no resultan propensos al cambio como consecuencia de la transición, sucesos cuya ocurrencia es independiente de la transición o que no están amenazados ulteriormente por ella. En términos más precisos, son los sucesos¹⁰ en

¹⁰ Los "sucesos" incluyen en este contexto situaciones, procesos, etc.

cuya ocurrencia la transición de $\sim A$ a A no posee efectos causales negativos de ninguna especie.¹¹

Puesto que, como podemos presuponer aquí, un acontecimiento afectó causalmente a otro posterior,¹² pero no a uno anterior, la historia del mundo previa al tiempo del antecedente t_A ¹³ sigue estando enteramente dentro de la descripción modificada del mundo con base en la cual deben determinarse los efectos de la transición. De igual modo, como para un contrafáctico $A > B$ los sucesos que se producen estrictamente después¹⁴ de t_B (el límite superior de t_B) no tendrían ninguna importancia para la ocurrencia del suceso B (es decir, el acontecimiento descrito por B), la parte restante de la historia del mundo que queremos retener pertenece al intervalo (t_A, t_B) ¹⁵ y se compone de proposiciones que describen los sucesos en cuya ocurrencia la transición no posee ningún efecto negativo.

Los efectos causales de la transición de $\sim A$ a A (es decir, la ocurrencia del suceso A en lugar del $\sim A$) sobre los acontecimientos posteriores pueden ser complejos y transmitirse en más de una forma por más de una cadena causal. Algunos efectos de la transición sobre la ocurrencia de un suceso futuro pueden ser positivos, otros negativos o la transición puede no tener ningún efecto causal en absoluto sobre el suceso futuro. Si la transición no posee un efecto

¹¹ Nótese que considerar la ocurrencia (o los efectos) de la transición de $\sim A$ a A equivale, desde luego, a considerar la ocurrencia (o los efectos de la ocurrencia) del suceso no real A (en lugar del suceso real $\sim A$, en nuestro tipo de casos).

Ocasionalmente me permitiré obviar el nivel del objeto y el meta-nivel, donde se trate de sucesos y de las proposiciones que los describen, cuando no haya riesgo de confusiones, con el fin de evitar formulaciones engorrosas.

¹² Es muy posible que la causación retrospectiva no sea reconocida en nuestro sistema conceptual del sentido común. El impacto de esta falta de reconocimiento junto con el principio de retención mencionado tendría un efecto semejante al de la concepción no determinista del sistema del sentido común que se expuso en la sección anterior. Por lo tanto, no se necesita en este lugar ningún supuesto explícito por mi parte para descartar la causación retrospectiva: el que dicho supuesto sea operativo en un sistema conceptual del sentido común, cuando se combina con el principio de retención citado, bastará para dar forma a las premisas implícitas. De este modo, puedo ignorar aquí las objeciones concernientes a la viabilidad de la causación retrospectiva.

¹³ Es decir, antes del primer extremo de ese intervalo.

¹⁴ Es decir, que se producen en el intervalo temporal posterior al intervalo temporal t_B , y totalmente independientes de él.

¹⁵ Es decir, el intervalo cuyo punto inicial coincide con el de t_A y cuyo punto final coincide con el de t_B . En este artículo, me limito al tratamiento de contrafácticos con un orden temporal normal, es decir, para los cuales el primer extremo de t_A es anterior al de t_B . Para extender el tratamiento que aquí se presenta a contrafácticos que no tienen un orden temporal normal, cf. [6], cap. 9, IV, 1.

causal negativo sobre la ocurrencia de un suceso futuro, puede no tener ningún efecto causal o puede poseer uno puramente positivo. De esta manera, los sucesos reales que son totalmente independientes del acontecimiento $\sim A$ no son candidatos para no ocurrir como resultado de la transición al acontecimiento A . De igual modo, para los sucesos reales cuya ocurrencia está afectada de manera puramente positiva por la ocurrencia del suceso A , su no ocurrencia no es un efecto, sin duda, de la transición al suceso A . Por lo tanto, la ocurrencia de dos tipos de sucesos en el intervalo (t_A, t_B) —y sólo esos dos tipos— no se vería amenazado como resultado de la transición: los sucesos para cuya ocurrencia ésta es causalmente irrelevante y los sucesos para cuya ocurrencia la transición posee una relevancia causal puramente positiva.

Llegamos así al siguiente análisis de los contrafácticos del tipo d.n. (utilizando el signo “ $-L\rightarrow$ ” para la implicación [entailment] a través de las leyes de la naturaleza L):

Análisis I:

Un contrafáctico $A > B$ (del tipo d.n.) es verdadero si y sólo si

(2)

$$\{A\} \cup M_A \cup \left\{ \begin{array}{l} \text{proposiciones verdaderas (no legalifor-} \\ \text{mes) en } (t_A, t_B) \text{ que describen sucesos} \\ \text{para los cuales } A \text{ es causalmente irre-} \\ \text{levante o posee una relevancia causal} \\ \text{puramente positiva} \end{array} \right\} -L\rightarrow B$$

(M_A es la historia del mundo anterior a t_A .)¹⁶

4. *En qué consisten las premisas implícitas*

Quisiera profundizar la argumentación en favor de este análisis mediante la consideración de varios ejemplos. La evaluación de los contrafácticos cuyas consecuencias pertenecen a las diversas categorías

¹⁶ El uso de la conveniente notación “ $-L\rightarrow$ ” no tiene que interpretarse como un menoscabo de la condición de las leyes L por hallarse entre las premisas implícitas.

Para la semántica formal que refleja el análisis de los contrafácticos del tipo d.n., cf. [6], cap. VI, o [4]. Se previene al lector que tenga cuidado con los numerosos errores tipográficos de este artículo.

Los antecedentes de contrafácticos cuya estructura temporal involucra una cuantificación existencial exigen un tratamiento un poco más refinado; cf. [6], cap. 6, II, n. 17.

de proposiciones verdaderas esbozadas más arriba brindan más evidencias en favor de la preservación de las categorías de proposiciones que se sugieren en este análisis dentro del marco del modelo inferencial.

Resulta obvio que la categoría de las leyes no se puede dejar fuera de las premisas implícitas en el modelo inferencial.¹⁷ Sin las leyes, el modelo funcionaría únicamente para contrafácticos muy triviales —los que involucran relaciones de implicación [*entailment*] entre, por un lado, el antecedente y algunas otras proposiciones fácticas y, por el otro, el consecuente.¹⁸ Así mismo, debemos notar que para muchas leyes se pueden construir contrafácticos que son directamente “parasitarios” de estas leyes, es decir, en los que el antecedente especifica todos los parámetros de la ley excepto uno, el cual es especificado por el consecuente según las disposiciones de esa ley. Indudablemente, la preservación de las leyes es esencial para que estos contrafácticos verdaderos resulten verdaderos en el esquema inferencial. Como ilustración (*ejemplo 2*), supongamos que sostuve un gis en mi mano por encima de la mesa durante cierto periodo de tiempo al que pertenece t . No estaba operando ninguna fuerza especial, salvo la gravedad. El contrafáctico: “Si hubiera soltado el gis en t , habría caído sobre la mesa”, es verdadero bajo estas circunstancias. Pero si el consecuente debe inferirse por medio del esquema inferencial, entonces las premisas implícitas deben incluir información sobre la ubicación de mi mano, la mesa y las fuerzas gravitacionales en operación. Pero también debe tenerse información que exprese el contenido de la ley de Newton sobre las condiciones de la caída libre o de alguna otra ley que cubra el mismo aspecto. Sin esta ley no se seguiría el consecuente y el contrafáctico no satisfaría el modelo inferencial.

Con el fin de visualizar la necesidad de retener M_A (la historia —enteramente— anterior al tiempo al que pertenece el antecedente A) entre las premisas implícitas del contrafáctico $A > B$, consideremos el siguiente ejemplo (*ejemplo 3*): Harry era un suicida crónico e incurable. Consiguió una poderosa bomba y un día, en el momento t , la hizo detonar sobre sí y se desintegró por completo. Ahora, consideremos el contrafáctico: “Si la bomba no hubiera explotado, Harry

¹⁷ Nótese que en este lugar únicamente tratamos contrafácticos del tipo d.n. y, desde luego, no contralegales. La cuestión de si hay que preservar las leyes para los antecedentes contrafácticos que son incompatibles con su historia previa puede constituir un punto de controversia; pero no lo es para el tipo de contrafácticos que aquí tratamos.

¹⁸ Recordemos que, por lo general, no se puede esperar que las proposiciones fácticas pertenecientes a un tiempo impliquen [*entail*] una proposición perteneciente a otro tiempo.

se habría suicidado algún tiempo después.¹⁹ Dada su condición psicológica crónica, podemos suponer razonablemente que esto es verdad. No obstante, sin la información relativa a la condición psíquica previa de Harry no podemos fundamentar el citado contrafáctico. Y, sin embargo, no sería posible extraer esta información de la situación después de t_A , por ejemplo, del cerebro desintegrado de Harry. Pero la obtendríamos si hacemos figurar a M_A entre las premisas implícitas.

Desde luego, la mayoría de las leyes y la mayoría de la información de M_A normalmente sobreabundan en el esquema inferencial cuando se considera un contrafáctico particular. En la práctica, cuando se evalúa un contrafáctico dado, sólo se utiliza una parte de las mismas. Pero su inclusión total no perjudica a la adecuación extensiva del modelo inferencial y, por lo tanto, no hay necesidad de separar las que efectivamente necesitan ser tomadas en cuenta para un contrafáctico determinado de las que no es preciso considerar.²⁰

Al contrario del caso de M_A para los contrafácticos del tipo d.n., no todas las proposiciones verdaderas que *no* pertenezcan a tiempos anteriores a t_A pueden conservarse en su totalidad (incluso después de excluir la negación del antecedente). Esto es obvio, pues, de otro

¹⁹ Supongamos que el mecanismo detonador de la bomba era deficiente y comprendía un elemento de azar con respecto a si la bomba iba a explotar al ser detonada. Esto haría plausible el carácter d.n. del antecedente del contrafáctico. (Su plausibilidad puede intensificarse mediante la elaboración de otros detalles según la explicación de la sección 14, más abajo.)

²⁰ Por lo tanto, en principio, aun cuando pudiera bastarnos la información arrojada por una fracción del mundo inmediatamente anterior a t_A , no hay razón para privarnos de información que, aunque pueda ser redundante, resulta útil al elaborar las evaluaciones de ejemplos particulares, en tanto que no perjudica, según lo sostengo, la viabilidad de nuestro análisis.

Habría que agregar que el propósito de estos ejemplos, además de los otros que siguen, consiste en fortalecer la tesis de que M_A , así como las proposiciones de (t_A , t_B) para las cuales A es causalmente irrelevante o posee una relevancia causal puramente positiva, debe conservarse entre las premisas implícitas. En términos estrictos, desde luego, los ejemplos no demuestran que hay que incluir a *todas* esas proposiciones (y no hay razón para creer que *todas deban ser* incluidas para *todo* contrafáctico particular a fin de que se le aplique el esquema inferencial). Pero sostengo que deben ser incluidas en el esquema global que se aplica en términos generales a *todos* los contrafácticos de este tipo. Esta propuesta se apoya también en las consideraciones generales del comienzo de la sección 3 y en la consideración del análisis de los semifácticos (más abajo, sección 8), y se halla justificada finalmente por el argumento de que esta clase de redundancia no menoscaba la aplicación del análisis a contrafácticos particulares. (Quizás habría que tratar de formular una proposición según la cual por cada unidad de información de estas categorías, existe un contrafáctico d.n. verdadero para el cual la presencia de esa unidad de información en el esquema inferencial es crucial a fin de que éste funcione.)

modo, ningún contrafáctico $A > B$ que tenga una B falsa resultaría verdadero en nuestro modelo sin un antecedente inconsistente en el esquema inferencial, y no es posible tolerar esto si $A > \sim B$ va a resultar falso. Por ello, con frecuencia este conjunto resultante sería inconsistente. Goodman, al afrontar un problema semejante, exigió que no se conservaran las proposiciones que no fueran cosustentables [*cotenable*] con el antecedente, por donde para él B no es cosustentable con A en caso de que $A > \sim B$ sea verdadero. Sin embargo, como veremos, no pueden conservarse de esta forma grandes grupos de proposiciones cosustentables con el antecedente. Por lo tanto, en lugar de caracterizar las premisas implícitas con la noción de cosustentabilidad (que, además de lo mencionado, conduce a la circularidad), las caracterizaremos de una manera directa y pormenorizada. Por otra parte, algunas proposiciones no sólo pueden ser incompatibles con A , L y M_A individualmente, sino que ciertas proposiciones lo pueden ser colectivamente.²¹ Esto haría que el problema de la selección fuera aún más apremiante y complejo.

El hecho de que no podamos dejar de incluir como premisas implícitas proposiciones pertenecientes a (t_A, t_B) para las cuales el suceso A es causalmente irrelevante se ilustra mediante el siguiente ejemplo (*ejemplo 4*): consideremos a Jones, quien estando en su agencia de bolsa, deliberó largamente si iba a vender o no sus acciones (pues necesitaba el dinero), hasta que se decidió a venderlas. Unas semanas después, la bolsa de valores subió por las nubes. Puesto que la decisión de Jones (basada en circunstancias personales) no afectó al mercado (pues se trataba de un inversionista bastante pequeño) y suponiendo ciertas condiciones de normalidad en su situación financiera durante este período, habría sido en efecto el caso que:

- (3) Si no hubiera vendido sus acciones, en unas semanas más se habría enriquecido.

Es evidente que la acción de Jones, en las circunstancias descritas, resulta irrelevante causalmente para el alza de la bolsa de valores. También es claro que debe retenerse una proposición que describa el alza entre las premisas implícitas del contrafáctico (3) para que éste resulte verdadero.²²

²¹ Goodman ya observó el problema que presenta la incompatibilidad colectiva.

²² Desde luego, doy por supuesto que el alza de la bolsa de valores no estuvo determinada nomológicamente por la historia previa al antecedente.

Notemos que aun cuando el hecho de que Jones vendiera sus acciones pudo haber tenido un pequeño efecto sobre las condiciones del mercado inmediatamente

Con el fin de visualizar la necesidad de conservar como premisas implícitas de un contrafáctico $A > B$ las proposiciones que pertenecen al intervalo (t_A, t_B) para el que el antecedente posee una relevancia causal puramente positiva, consideremos el *ejemplo 5*: el arquitecto Tom presentó su propuesta al consejo directivo de un banco que quería construir una nueva sucursal y solicitó x dólares por costos de construcción más su comisión. El consejo directivo, que en su conjunto era muy tacaño, favorecía firmemente los bajos costos. Tom obtuvo el contrato, construyó el edificio, pero los gastos de construcción resultaron, en realidad, mayores de lo que había calculado y, por lo tanto, no obtuvo ninguna ganancia. Podría entonces decir: "Si hubiera solicitado mil dólares menos, habría perdido dinero en este edificio", y estaría en lo correcto. Sin embargo, para que este contrafáctico pase el examen del modelo inferencial, debe haber, entre las premisas implícitas, una proposición relativa al hecho de que se le concedió el contrato. Pero el antecedente es en realidad relevante causalmente respecto para esa proposición y, además, lo es positivamente. Pero, lo que es más significativo, su relevancia causal es *puramente* positiva. Es decir, el hecho de que Tom pida mil dólares menos por su contrato afectaría de manera puramente positiva a sus probabilidades de obtenerlo y no las menoscabaría en modo alguno: no tendría ningún efecto negativo parcial sobre sus probabilidades. De aquí que debemos conservar, entre nuestras premisas legítimas del intervalo (t_A, t_B) , las proposiciones que describen sucesos para los cuales el suceso-antecedente posee una relevancia causal puramente positiva.

5. *Proposiciones que no pertenecen a las premisas implícitas*

Hasta aquí hemos tratado de la conservación de ciertas categorías de proposiciones entre las premisas implícitas de un contrafáctico $A > B$ del tipo d.n. Me gustaría argumentar ahora de un modo semejante que esta lista de categorías es completa y que no deben incluirse otras categorías de proposiciones verdaderas.

Como cabe esperar, no podemos reconocer como premisas legítimas las proposiciones que describen sucesos para los cuales el suceso-antecedente posee una relevancia causal puramente negativa. Con el fin de visualizarlo, consideremos lo siguiente (*ejemplo 6*): John, que es una persona normal en términos emocionales, entrevista a Jane para cubrir un puesto en su departamento en el lugar donde trabaja.

después, no es preciso que fuera determinante, y para los propósitos de este ejemplo doy por supuesto que no lo fue, con respecto a la drástica alza de la bolsa de valores que se produjo unas semanas más tarde.

Recibió instrucciones de ofrecer el empleo a la primera persona entrevistada que se encuentre calificada y con quien le agradaría trabajar. Jane, la primera persona que entrevista, está bien calificada, le agrada y, en consecuencia, le ofrece el puesto al final de la entrevista. Ahora consideremos el siguiente contrafáctico:

- (4) Incluso si Jane, sádica y gratuitamente, le hubiera dado una cachetada en el curso de la entrevista, John le habría ofrecido el puesto al final.

Este contrafáctico es evidentemente falso, dada la situación descrita. El hecho de que Jane le dé una cachetada posee, con toda claridad, una relevancia causal puramente negativa con respecto al agrado que debe sentir John para trabajar con ella. Pero este suceso-antecedente es incluso irrelevante causalmente para la situación (en un momento inmediatamente posterior a la hipotética cachetada) en la cual John está consciente de la calificación profesional de Jane y todavía desea dar el cargo a una persona calificada con quien le agradaría trabajar. De esta manera, según nuestro argumento anterior, tal descripción de esta situación califica como premisa implícita para el contrafáctico (4). Ahora que, si también tenemos que reconocer como premisas implícitas las proposiciones que describen sucesos para los cuales el suceso-antecedente posee una relevancia causal puramente negativa, deberíamos reconocer como premisa implícita una proposición que especifique que a John le agradaría trabajar con Jane incluso después de que le dio una cachetada. Pero si procediéramos así, el contrafáctico (4) satisfaría el modelo inferencial y, por lo tanto, resultaría verdadero. En consecuencia, no podemos admitir entre las premisas implícitas el conjunto de proposiciones que describen sucesos para los cuales el suceso-antecedente posee una relevancia causal puramente negativa.

Finalmente, veremos que el último grupo que queda no puede calificar, como grupo, entre las premisas implícitas. Este grupo se compone de proposiciones del intervalo (t_A, t_B) que describen sucesos para los cuales el suceso-antecedente posee un efecto causal mixto —parcialmente positivo, parcialmente negativo, en grados diversos. Para visualizar lo anterior, consideremos (*ejemplo 7*) el caso descrito por el siguiente supuesto contrario a los hechos:

- (5) J. F. Kennedy eligió a Jackson (en lugar de Johnson) como vicepresidente para la fórmula electoral en 1960.

Este suceso tendría una clara relevancia causal para el acontecimiento real descrito por:

- (6) Kennedy ganó las elecciones presidenciales de 1960,

pero de una manera no uniforme. Su relevancia causal sería obviamente negativa por cuanto Kennedy bien podría haber perdido el sur de haber elegido a Jackson; y evidentemente (como lo indicaron los resultados reales de las elecciones) si el sur hubiese apoyado a Nixon, éste bien habría podido ganar. Pero el suceso (5) también habría tenido un efecto positivo en el (6), pues el estado de Washington, que de hecho votó por Nixon, habría podido votar por la fórmula Kennedy-Jackson. De este modo, el suceso (5) es relevante causalmente para el (6) de forma mixta: existen algunos efectos positivos y algunos negativos. Sin embargo, está claro que el contrafáctico:

- (7) Incluso si Kennedy hubiera elegido a Jackson (en lugar de Johnson) como vicepresidente en la fórmula en 1960, Jacqueline Kennedy habría sido primera dama en 1961,

no es verdadero: si el antecedente hubiera sido verdadero, Kennedy tal vez habría perdido las elecciones —esto, entonces, habría sido una posibilidad muy real, y en modo alguno remota—, y por lo tanto Jacqueline no habría sido primera dama en 1961. No obstante, si retenemos (6) como premisa implícita en (t_A, t_B) para el contrafáctico (7), éste ciertamente resultaría verdadero, puesto que de haber Kennedy elegido a Jackson y dado que aun así habría ganado las elecciones, Jacqueline se habría convertido por lo tanto en primera dama. La selección de Jackson en lugar de Johnson es causalmente irrelevante con respecto a las probabilidades de supervivencia física que ella y su marido tenían durante el intervalo (t_A, t_B) y, en consecuencia, el que estuvieran vivos en 1961 puede describirse mediante una premisa implícita. De aquí que la admisión de (6) como premisa implícita provocaría que un contrafáctico falso resultara verdadero en el esquema del modelo inferencial y por ello debe ser excluido.²³ De esta manera, y en términos generales, las proposiciones del intervalo (t_A, t_B) , para las que el antecedente posee una relevancia causal mixta —en parte positiva y en parte negativa— no califican como

²³ Desde luego, en este ejemplo doy por supuesto algo que resulta muy plausible desde la perspectiva no determinista: que para la época de la convención demócrata (cuando Kennedy seleccionó a Johnson como vicepresidente para la fórmula electoral) la elección de Kennedy no estaba determinada nomológicamente por la historia anterior a ese momento (recordemos que, de hecho, las elecciones estaban muy próximas). (Notemos que si *modificamos* el ejemplo bajo un supuesto descriptivo *diferente*, de tal modo que (6) *estaba* determinado por la historia anterior a la convención, mientras que seguimos presuponiendo que (5) puede servir como antecedente d.n., (7) resultaría entonces verdadero. Pero un caso con un supuesto tan poco plausible no constituye un contraejemplo para mi análisis: simplemente no contribuye a fundamentar aquí mi propuesta, mientras que esto se logra mediante otros ejemplos, como los que he seleccionado.)

premisas implícitas para el esquema inferencial.²⁴ Además, la inclusión de esas proposiciones como premisas implícitas concedería la verdad a semifácticos (es decir, contrafácticos con consecuentes verdaderos), cuyos antecedentes son causalmente relevantes para sus consecuentes de una manera mixta. Estos semifácticos resultarían entonces verdaderos en el modelo inferencial, aun cuando muchos de ellos son en realidad falsos, como lo ilustra el semifáctico (7).

La claridad y aceptabilidad del análisis anterior dependen de las nociones de legaliformidad [*lawlikeness*], irrelevancia causal y relevancia causal puramente positiva. Los últimos dos conceptos se tratan en términos de probabilidad objetiva en las secciones 17 y 18.²⁵ De esta manera, el problema de los contrafácticos queda reducido a los conceptos de legaliformidad y probabilidad objetiva, para los cuales el realizar análisis independientes no es menos importante que para la construcción contrafáctica.

6. *Semifácticos*

Los semifácticos son contrafácticos que tienen consecuentes verdaderos (con la forma típica: "Incluso si... , habría sido el caso que..."); los *semifácticos irrel* son aquellos cuyos sucesos-antecedentes son causalmente irrelevantes para sus sucesos-consecuentes, y los *semifácticos p.p.* son aquellos cuyos sucesos-antecedentes poseen una relevancia causal puramente positiva para sus sucesos-consecuentes.²⁶ Mientras que la función de los contrafácticos comunes es predecir (hipotéticamente) los efectos de los sucesos descritos en los supuestos contrarios a los hechos que constituyen sus antecedentes, la función de los semifácticos irrel y p.p. consiste en determinar (por medio de sus consecuentes) los aspectos del curso real de los acontecimientos (en (t_A, t_B)) que deben constituir el trasfondo sobre el cual se van a evaluar los efectos de la transición de $\sim A$ a A (para un antecedente A). Como las proposiciones que pertenecen a este trasfondo tienen que conservarse en la descripción del curso real de los acontecimientos cuando se evalúa el supuesto contrario a los hechos, y por lo tanto se las considera verdaderas cuando se toma a A como

²⁴ Este sigue siendo el caso aun cuando el efecto general del antecedente, aunque mixto, sea positivo; véase la sección 7, más abajo.

²⁵ La noción de probabilidad a la que me refiero aquí y en el tratamiento de los procesos con más probabilidades de haber conducido a A (más abajo, sección 13), es una noción objetiva, emparentada con la de oportunidad objetiva. Para más detalles, cf. sección 17; [6], cap. 4, sección VI, y [8].

²⁶ Con frecuencia estos semifácticos se expresan con la locución auxiliar "*a fortiori*"; es decir, si el suceso C ocurrió después de que el suceso $\sim A$ se produjo, *a fortiori* habría ocurrido si se hubiera producido el suceso A .

un supuesto contrario a los hechos, todos los semifácticos irrel y p.p. deben ser verdaderos (lo cual también es consecuencia inmediata del Análisis I). Que éste es efectivamente el caso —que la irrelevancia causal o la relevancia causal puramente positiva entre los sucesos descritos por el antecedente y el consecuente de un semifáctico son realmente suficientes para que sea verdadero este último— constituye una generalización empírica que se sigue del Análisis I y, puesto que es verdadera, constituye por lo tanto una confirmación de este análisis. Desde luego, los sucesos-antecedentes de los semifácticos restantes poseen, de una u otra forma, cierta *relevancia causal negativa* para sus sucesos consecuentes y por ello pueden ser denominados semifácticos r.n. A diferencia de los otros dos grupos de semifácticos, sólo *algunos* semifácticos r.n. son verdaderos (mientras que algunos son falsos). Al igual que otros contrafácticos del tipo d.n., caen dentro del alcance del Análisis I. Por ello, el último y engorroso miembro de (2) en el Análisis I puede ser representado en términos equivalentes como:

$$\left\{ \begin{array}{l} \text{El conjunto de consecuentes de los semifácticos} \\ \text{irrel y p.p. } A > C, \text{ con } t_C \subseteq (t_A, t_B) \end{array} \right\}$$

Si la irrelevancia causal y la relevancia causal puramente positiva son relaciones entre los sucesos descritos por dos proposiciones A y C y la historia previa M_A (que no se discutirá en este lugar),²⁷ entonces la función $f(A > B, \dots)$ de las premisas implícitas se convierte en $f(A, t_{\bar{B}})$,²⁸ de este modo:

$$f(A, t_{\bar{B}}) = M_A \cup L \left\{ \begin{array}{l} \text{el conjunto de consecuentes de} \\ \text{los semifácticos irrel y p.p. } A > \\ C \text{ para los cuales } t_C \supseteq (t_A, t_B) \end{array} \right\}$$

Vamos a denominar al último miembro de la unión de esta fórmula CIP($A, t_{\bar{B}}$) (para consecuentes de semifácticos irrel y p.p., etc.). No temos también que aunque $f(A, t_{\bar{B}})$ constituye una función de $t_{\bar{B}}$, no es, fuera de esto, una función de B por sí misma: dos contrafácticos $A > B$ y $A > C$ tendrían exactamente la misma premisa implícita si $t_{\bar{C}} = t_{\bar{B}}$. (Esta característica puede estar motivada de forma totalmente independiente del análisis anterior si se reconoce que la construcción contrafáctica sirve para hacer una predicción —especificada

²⁷ Cf. n. 72 más abajo.

²⁸ Recordemos que $t_{\bar{B}}$ es el límite superior del intervalo t_B .

por su consecuente— con base en una selección de características del mundo real que se halla determinada por el antecedente. De esta manera, predicciones hipotéticas diferentes —pertenecientes al mismo momento—, basadas en el mismo supuesto contrario a los hechos, compartirían las premisas implícitas y, por lo tanto, la función de éstas debe ser independiente de B .)

7. Otras ejemplificaciones

Volvamos al ejemplo 4 relativo a Jones y la bolsa de valores y veamos cómo se aplica a él el esquema (2). Como hemos observado, (3) (es decir, “Si Jones no hubiera vendido sus acciones, se habría enriquecido unas semanas después”) era un contrafáctico verdadero. En la explicación de este contrafáctico también observamos que el suceso-antecedente de:

- (8) Incluso si Jones no hubiera vendido sus acciones, la bolsa de valores habría subido,

es causalmente irrelevante para su suceso-consecuente y por lo tanto (8) es un semifáctico irrel. Ahora que, con el fin de fundamentar a (3) necesitamos, desde luego, información en M_A (por ejemplo, que la situación económica de Jones era pasable, que no tenía deudas pendientes, etc.). Así mismo, necesitamos la información que nos brinda el consecuente de (8), además de mucha información relativa a los hechos y situaciones que se produjeron en el intervalo (t_A, t_B) . De este modo, necesitamos saber que Jones no afrontó un desastre financiero durante ese periodo (en cuyo caso una ganancia repentina lo podría haber salvado, pero no lo habría enriquecido), que el poder de compra del dólar no varió sustancialmente en ese intervalo, que Jones no prometió donar toda ganancia inesperada futura a su suegra, etc. Suponiendo que todos estos hechos y situaciones realmente tuvieron lugar, según nuestro ejemplo, el uso de descripciones apropiadas de los mismos como premisas implícitas fundamentaría efectivamente al consecuente conforme al esquema (2). Pero es evidente que estos son hechos o situaciones para los cuales la transición de $\sim A$ a A (es decir, que Jones conserve sus acciones en lugar de venderlas) es causalmente irrelevante y, por lo tanto, sus descripciones apropiadas constituirían consecuentes de semifácticos irrel con el mismo antecedente que el de (3) y formarían parte así de las premisas implícitas del esquema (2) cuando se aplica al ejemplo 4. A partir de estas premisas implícitas, junto con el antecedente A , se seguiría en efecto el consecuente B y, por lo tanto, (3) califica

como contrafáctico verdadero según nuestro análisis, y ciertamente es verdadero.

Notemos que si se violan condiciones como las que hemos supuesto que prevalecían en (t_A, t_B) para este caso (por ejemplo, si Jones incurría en una deuda importante), entonces el esquema (2) no se cumpliría; pero esto resulta como debe ser: el contrafáctico será entonces falso. De este modo, si el hecho de que Jones venda sus acciones posee una relevancia causal positiva para su solvencia financiera durante (t_A, t_B) (por ejemplo, si sus ingresos por las acciones lo salvaron del desastre financiero), no podemos conservar la especificación de que su situación económica fue pasable durante (t_A, t_B) entre las premisas implícitas (puesto que el suceso antecedente poseería entonces una relevancia causal negativa para esta situación) y así (2) no resultaría verdadero en este caso. Pero nuevamente esto es como debe ser: el contrafáctico no tiene que considerarse como verdadero en tal caso. Esto nos brinda un panorama de la dependencia funcional del valor de verdad de un contrafáctico respecto del curso real de los acontecimientos en el intervalo (t_A, t_B) , lo cual es en efecto lo que, entre otras cosas, hace que el contrafáctico verse *sobre* el mundo real en un sentido fuerte e ilumina sus características empíricas.

Recordemos el ejemplo 5 sobre el arquitecto Tom y la nueva sucursal bancaria. El semifáctico: "Incluso si Tom hubiera solicitado mil dólares menos de los que pidió, habría obtenido el contrato", constituye un semifáctico p.p., pues, como vimos en la explicación de este ejemplo, el suceso-antecedente posee una relevancia causal puramente positiva para el suceso-consecuente. No obstante, si consideramos otra vez el ejemplo 7 relativo a Kennedy y Jackson, el semifáctico: "Incluso si Kennedy hubiera elegido a Jackson en lugar de Johnson como compañero de fórmula, habría sido electo para presidente", sería un semifáctico r.n. puesto que, como observamos, el suceso-antecedente posee una relevancia causal mixta para el suceso-consecuente.

Desde luego, cuando se consideran los ejemplos, normalmente sólo se dispone de fracciones de M_B ,²⁹ aunque se deba aludir a diversas características de éste para convencernos de que la elaboración de un ejemplo va a tener un valor de verdad particular para el contrafáctico en cuestión. Como otra ilustración (*ejemplo 8*), supongamos que Daniel ganó una lotería llevada a cabo por un director y su asistente, y donde cada participante tenía 0.001 de probabilidades de ganar el gran premio. Pero Daniel lo ganó porque sobornó al asistente para que lo ayudara. La colaboración del asistente aumentó sus probabilidades en un 10%, pero la ayuda del director las habría aumentado

²⁹ Es decir, la historia anterior a t_B .

en un 30%. Si Daniel hubiera hecho un trato con el director en lugar del asistente, ¿habría ganado? No, incluso con una garantía del 30% podría haber perdido. ¿Habría perdido? Esto tampoco puede afirmarse; podría haber ganado. Ninguno de los dos contrafácticos involucrados es verdadero. Por lo tanto, el simple aumento de las probabilidades del consecuente no basta para convertir en verdadero al contrafáctico.³⁰ Está claro que el suceso-antecedente posee una relevancia causal mixta para el suceso-consecuente en ambos contrafácticos, donde el efecto de la ayuda del director y el efecto de la pérdida de la ayuda del asistente reflejan relevancias causales positiva y negativa respectivamente en el primer contrafáctico, mientras que reflejan relevancias causales negativa y positiva respectivamente en el segundo. Ambos son semifácticos r.n. falsos. Evidentemente, incluir entre las premisas implícitas proposiciones que describan hechos para los cuales el suceso-antecedente posee una relevancia causal mixta —como el que, en este caso, Daniel ganara— también provocaría aquí que un contrafáctico falso (como el primero en este caso) resultara verdadero. De esta manera, tanto $A > B$ como $A > \sim B$ pueden ser falsos, según lo ilustra este ejemplo. En consecuencia, debemos sostener que $\diamond A \supset (\sim(A > B) \equiv A > \sim B)$ es efectivamente falso.³¹

8. *Contrafácticos - Análisis II*

Se puede abordar el problema de los contrafácticos del tipo d.n. de distinta manera. Preguntémonos nuevamente cuáles son las proposiciones pertenecientes al curso real de los acontecimientos (en (t_A, t_B)) que deben conservarse cuando se evalúa un supuesto contrario a los hechos. ¿Cuáles de estas proposiciones verdaderas deben ser aún consideradas como verdaderas dentro de este contexto, contra cuyo trasfondo va a evaluarse uno de esos supuestos? Bien, evidentemente éstas son exactamente las proposiciones que seguirían siendo verdaderas aun cuando tal supuesto fuera verdadero; es decir, los consecuentes (pertenecientes a (t_A, t_B)) de los semifácticos verdaderos que posean el supuesto contrario a los hechos como antecedente. Los semifácticos que tienen el antecedente A sirven así para

³⁰ Esto seguiría siendo así aun cuando en el ejemplo alteráramos el porcentaje de 30% a, digamos, 65%. En consecuencia, el aumento de las probabilidades no sólo sería insuficiente para conceder la verdad a un contrafáctico, sino que tampoco bastaría un aumento de más del 50% de las probabilidades.

³¹ Lo cual es contrario a la postura de Stalnaker. Éste defendió además la validez de este principio en su [14]. Sin embargo, no es posible emprender en este artículo una discusión sobre la defensa de Stalnaker.

demarcar esa parte de la descripción del mundo real (en (t_A, t_B)) que es *invariante* bajo la transición de $\sim A$ a A . Por lo tanto:

Análisis II:

Un contrafáctico $A > B$ (del tipo d.n.) es verdadero si y sólo si
(9)

$$\{A\} \cup M_A \cup \left\{ \begin{array}{l} \text{los consecuentes de los semifácticos} \\ \text{verdaderos } A > C \text{ con } t_C \subseteq (t_A, t_B) \end{array} \right\} -L \rightarrow B$$

Surge así el problema de la relación entre los dos análisis, *I* y *II*, pues los últimos elementos de los antecedentes de (2) y (9) difieren. Además, el Análisis *II* presenta un serio problema por sí mismo, debido a que el último componente de la unión en (9) hace uso de (entre otros) semifácticos r.n. verdaderos. Por lo tanto, la evaluación de un contrafáctico según el Análisis *II* depende de la verdad de los semifácticos r.n. de (t_A, t_B) que, a su vez, cuando se analizan según (9), dependen de la verdad de los semifácticos r.n. del mismo intervalo y así sucesivamente, en lo que parece ser una regresión infinita. (Este problema no se presenta para los semifácticos irrel y p.p., puesto que nos hallamos obligados a darles análisis independientes.) De tal modo, el Análisis *II* parece muy problemático.

Sin embargo, se puede demostrar que el problema de la regresión infinita es espurio. Puesto que los semifácticos r.n. caen bajo el Análisis *I*, es posible probar que ambos análisis son equivalentes:

Teorema: Si los semifácticos r.n. caen bajo el Análisis *I*, entonces: el Análisis *I* y el Análisis *II* son equivalentes.³²

El punto aquí es muy claro: mediante el Análisis *I*, el conjunto de consecuentes de (t_A, t_B) de los semifácticos r.n. verdaderos con antecedente A pertenece a la clausura lógica del resto de la unión que forma el antecedente de (9) en el Análisis *II*, y por lo tanto resulta redundante en éste; de aquí que se evite la regresión infinita. En realidad, se puede demostrar que si el antecedente de (2) en el Análisis *I* se abrevia como $Q(A, t_B)$ y

$$\left\{ \begin{array}{l} \text{el conjunto de consecuentes de los semifácticos r.n.} \\ \text{verdaderos } A > B \text{ con } t_C \subseteq (t_A, t_B) \end{array} \right\}$$

se abrevia como $CCS(A, t_B)$, entonces:

³² Para la prueba formal de este teorema, cf. [6], cap. 2, XII, 1, 2.

$$\text{CCS}(A, t_{\bar{B}}) = (\text{Q}(A, t_{\bar{B}}) - \text{Q}(A, t_{\bar{B}})^T)_{(t_A, t_B)}$$

donde, si S es un conjunto de proposiciones, entonces $S_{(t, t')}$ es el subconjunto cuyos miembros pertenecen a los intervalos temporales dentro del intervalo (t, t') ,³³ y S^T es el subconjunto de proposiciones verdaderas de S . Por lo tanto, los análisis son compatibles.

9. Goodman

Es preciso hacer ahora algunas observaciones con respecto a otras teorías de los contrafácticos. (Ocasionalmente, la exposición se concentrará en estas teorías en tanto que se limitan al tipo de contrafácticos explicados más arriba, es decir, los del tipo d.n.)

Goodman [2] comenzó tratando de encontrar las condiciones relevantes (que deben servir como premisas implícitas —salvo las leyes— según las hemos introducido anteriormente, como en el esquema (1)); luego pasó a una caracterización de la forma: $(ES)(S \& (A \& S - L \rightarrow B) \& C(S))$ ³⁴ (una forma también contemplada por Chisholm [1], aunque Goodman consideró restricciones diferentes para ella); y luego a una caracterización de la forma

$$\begin{aligned} & (ES)(S \& (S \& A - L \rightarrow B) \& R_1(S, A, B)) \\ & \& (ES')(S' \& (S' \& A - L \rightarrow B) \& R_1(S', A, B)) \end{aligned}$$

donde S y S' son conjuntos de proposiciones y $R_1(S, A, B)$ es una restricción de S . Goodman entonces planteó el problema de la cosustentabilidad —cómo excluir proposiciones verdaderas no cosustentables con el antecedente— y terminó reconociendo que su línea conducía a una regresión infinita. Sin embargo, la fórmula de Goodman no es totalmente inadecuada; también puede trivializarse si se toma $\{N \& (A \supset C)\}$ por S y $\{N \& (A \supset \sim C)\}$ por S' , bajo ciertas condiciones que son fáciles de satisfacer.³⁵ Además, es muy improbable que

³³ Nuevamente, t y t' son los límites inferior y superior del intervalo t , respectivamente. El signo “ $\text{CCS}(A, t_{\bar{B}})$ ” significa “consecuentes de los semifácticos del tipo contrafáctico”, que es otro nombre de los semifácticos r.n.: utilizo esta designación porque su comportamiento lógico es afín al de los contrafácticos comunes, más que al de los semifácticos irrel o p.p.

³⁴ Para simplificar la notación, supongo que los conjuntos de proposiciones pueden ser asimilados a las conjunciones de sus miembros. $C(S)$ es una restricción de S .

Las referencias a Goodman que se hacen más abajo corresponden a la 2a. edición, 1965.

³⁵ Si estas condiciones son que N es verdadero, $\sim A \nrightarrow N$ y $\{N, A, C\}$, $\{N, \sim C, A\}$ y $\{N, \sim A, C\}$ son autocompatibles. De aquí que, por ejemplo, se cumpliera la restricción de que ni S ni S' se sigan de A , que es la intención de Goodman en p. 13, n. 7. Pero hay que recordar que Goodman no suscribía oficialmente a estos resultados como soluciones finales.

estas estrategias se puedan corregir mediante otras modificaciones que no sean drásticas, si nuestros análisis de los contrafácticos (y de las funciones de los contrafácticos y los supuestos contrarios a los hechos) son adecuados. Como lo hemos señalado, nuestro análisis implica que la función i.p. debe ser independiente de B . Pero parecería que cualquier intento por reformular la propuesta de Goodman para que se ajuste a la forma del esquema inferencial tiene que usar una función f con B como uno de sus argumentos. En consecuencia, además de estar expuesta a los argumentos de la trivialización mencionados más arriba, la estrategia de Goodman (independientemente de la forma particular en la que él la desarrolló) no es prometedora debido a esta consideración general, si nuestro enfoque está en el camino correcto.

El problema de la cosustentabilidad de Goodman, que él consideró crucial para el problema de los contrafácticos y que lo condujo a la regresión infinita, no es en realidad otra cosa más que el mismo problema de los contrafácticos disfrazado. De este modo, para un determinado antecedente A , el conjunto de proposiciones falsas que, tomadas como consecuentes de A , forman contrafácticos verdaderos, es simplemente el conjunto de negaciones de las proposiciones verdaderas con las cuales A no es cosustentable (puesto que A no es cosustentable con B si y solo si $\sim B$ habría sido el caso si A hubiera sido el caso). Por lo tanto, $A > B$ es verdadero (cuando B es falso)³⁶ si y solo si $\sim B$ no es cosustentable con A . De esta manera, toda teoría que caracterice adecuadamente a los contrafácticos *ipso facto* capta las proposiciones que no son cosustentables con el antecedente como el conjunto de negaciones de los consecuentes falsos de contrafácticos verdaderos, y viceversa: la caracterización de la cosustentabilidad arroja inmediatamente las condiciones de verdad para los contrafácticos.³⁷

No obstante, el camino seguido por Goodman parte de la estrategia de eliminar proposiciones del conjunto de todas las proposiciones fácticas verdaderas con el fin de evitar la incompatibilidad con el antecedente (por medio de las leyes) y con el fin de tener las condiciones relevantes descritas por el conjunto de las proposiciones restantes. Pero es evidente que, después de una eliminación que excluye a todas las proposiciones que no son cosustentables con el antecedente,³⁸ también se elimina cualquier incompatibilidad con el antecedente. De esta manera, la intención del enfoque de Goodman

³⁶ Recordemos que Goodman limitó su análisis principal a los contrafácticos con antecedentes y consecuentes falsos.

³⁷ Nuevamente, con consecuentes falsos.

³⁸ Individual o colectivamente.

equivalía a concebir la resolución del problema de la cosustentabilidad como una resolución del problema de las condiciones relevantes: el problema de eliminar las incompatibilidades con el antecedente.

Vamos a extendernos en este punto concentrándonos en la analogía entre el problema de Goodman de la regresión infinita relativa a proposiciones no cosustentables con el antecedente, y el problema de la regresión infinita que encontramos en el Análisis II, el análisis semifáctico de los contrafácticos (sección 8, arriba). El conjunto de proposiciones verdaderas C pertenecientes a (t_A, t_B) puede dividirse en tres: 1) aquellas para las cuales $A > \sim C$ es verdadero: las que son no cosustentables con A ; 2) aquellas para las cuales $A > C$ es verdadero: consecuentes de semifácticos verdaderos con antecedente A ; 3) aquellas para las cuales ni $A > \sim C$ ni $A > C$ son verdaderos. El problema de Goodman era cómo excluir al primer conjunto; nuestro problema era cómo incluir al segundo. Tanto el primero como el segundo conjuntos están especificados por medio de una forma contrafáctica y, en consecuencia, amenazan con una regresión infinita. Anteriormente expliqué que el problema de Goodman está arraigado en una inadecuación extensional: la simple exclusión de proposiciones para eliminar la no cosustentabilidad con el antecedente no iba a funcionar, pues las proposiciones del tercer grupo, aunque son cosustentables con el antecedente, no deben ser tomadas como premisas implícitas (recordemos el caso de Kennedy en el ejemplo 7 y el de la lotería en el ejemplo 8, expuestos arriba). Así mismo, el problema de Goodman consiste en un producto de la concepción básica que considera que la dificultad central de los contrafácticos es el problema de evitar la incompatibilidad con el antecedente: es esta concepción la que provoca que el problema de la cosustentabilidad se vuelva central. Pero esta fue una *concepción errónea* del problema que también dio origen a la idea de que lo que se necesita es el restablecimiento de la consistencia mediante “omisiones mínimas” o “cambios mínimos” (o recurriendo a una “similitud intuitiva máxima” de otros mundos posibles en los cuales el antecedente es verdadero). Según mi enfoque, lo que se necesita (por simple adecuación extensional) no es evitar la inconsistencia por medio de esos cambios mínimos, sino más bien retener las proposiciones verdaderas a las cuales el antecedente no pone en riesgo, o las proposiciones verdaderas que habrían seguido siendo verdaderas si el antecedente hubiera sido verdadero. Y, en este enfoque, según lo detallamos más arriba, no surge el problema de la cosustentabilidad. Éste se presenta en los enfoques que se centran en el problema de la compatibilidad como dificultad fundamental e intentan resolverlo mediante omisiones o cambios mínimos, puesto que entonces la preocupación es cómo garantizar que todas

las proposiciones que no son cosustentables con el antecedente estén incluidas en esas omisiones.

Pero, al margen del tema de la adecuación extensional, al explicar el Análisis *II* pudimos superar el problema de la regresión infinita relativo al segundo conjunto de proposiciones, mencionado en el párrafo anterior (un problema análogo al que tiene Goodman con respecto al primer conjunto), proponiendo una caracterización independiente para un subconjunto de estas proposiciones —los consecuentes de los semifácticos irrel y p.p.— por medio de un análisis probabilístico, y observando que el resto de este conjunto pertenece a la clausura lógica de este subconjunto, junto con el antecedente y las otras premisas implícitas, los cuales conforman el antecedente del esquema (1) del Análisis *I*. Por lo tanto, nuestra capacidad para evitar la regresión infinita dependió en parte de la viabilidad de un examen independiente de los contrafácticos (el del Análisis *I*), y en parte de la disposición a hacer un análisis independiente de un subconjunto de contrafácticos: los semifácticos irrel y p.p. Goodman no pudo recurrir a otro análisis (abandonó demasiado pronto el modelo inferencial) y no contempló la posibilidad de hacer un análisis independiente para una subclase de contrafácticos. Con esta limitación, es comprensible que no pudiera desembarazarse del problema de la regresión infinita.

10. *Reformulación en términos de descripciones de mundos posibles*

La mayoría de los tratamientos de los contrafácticos de los últimos años se han realizado en términos de análisis de mundos posibles. Sin embargo, aun para quienes, como yo, no simpatizan con el uso de éstos en los análisis filosóficos, resulta importante hacer comparaciones y contrastes. Con este propósito es beneficioso emplear las descripciones de mundos posibles, las cuales no son tan recusables como los mundos posibles. Si tomamos estas descripciones como conjuntos completos³⁹ de proposiciones no legaliformes pero compatibles con las leyes (y por lo tanto consistentes), se puede reelaborar fácilmente el análisis propuesto más arriba para los contrafácticos del tipo d.n. (como otras teorías metalingüísticas) en términos de descripciones de mundos posibles como sigue:

³⁹ En el sentido de que para toda proposición no legaliforme, ella o su negación está contenida en el conjunto (a menos que presuponga una proposición que no esté contenida en él).

Análisis III:

Un contrafáctico $A > B$ (del tipo d.n.) es verdadero si y sólo si B es verdadero en todas las descripciones de mundos posibles en las que A , M_A y el conjunto de proposiciones verdaderas no legaliformes pertenecientes a (t_A, t_B) , para las cuales A es causalmente irrelevante o posee una relevancia causal puramente positiva, son todos verdaderos.

Se puede demostrar que este análisis es equivalente al Análisis I.

Stalnaker [13] propone que un contrafáctico sea verdadero en un mundo (por ejemplo, el real) si y sólo si su consecuente es verdadero en un mundo seleccionado mediante alguna función de selección f (con el contrafáctico como su argumento). Stalnaker insiste en que f recoge un solo mundo, y fue criticado por ello.⁴⁰ Si permitimos que una función selectiva elija un conjunto de mundos, entonces evidentemente la determinación de f constituye el problema análogo, en el contexto del discurso de los mundos posibles, al de la determinación de las premisas implícitas, que es el problema principal dentro del contexto del enfoque asumido en este artículo.

Es fácil representar el Análisis III en términos de esta función selectiva como sigue:

- (10) Un contrafáctico $A > B$ (del tipo d.n.) es verdadero si y sólo si B es verdadero en todas las descripciones de mundos posibles recogidas por la función selectiva $f'(A, t_{\bar{B}})$, donde

$$f'(A, t_{\bar{B}}) = \left\{ \begin{array}{l} \text{el conjunto de descripciones de mundos} \\ \text{posibles en las que } A, M_A \text{ y las} \\ \text{proposiciones verdaderas no legaliformes} \\ \text{de } (t_A, t_B), \text{ para las cuales } A \text{ es} \\ \text{causalmente irrelevante o posee una} \\ \text{relevancia causal puramente positiva,} \\ \text{son verdaderos.} \end{array} \right.$$

11. *Similitud intuitiva máxima y cambios mínimos*

Según Stalnaker, el mundo seleccionado por la función selectiva f para un contrafáctico $A > B$ (evaluado, por ejemplo, en el mundo

⁴⁰ En diversos casos existen muchas razones para esperar un vínculo entre dos o más mundos posibles; cf. [9] de Lewis, pp. 80-81. Un ejemplo (tomado de mi disertación doctoral, Universidad de Pittsburgh, 1957) es: Si el número de cabellos de mi cabeza hubiera diferido en uno del que realmente es... En este caso parece existir un vínculo entre los mundos con un cabello más y los mundos con un cabello menos. Pero cf. la refutación de Stalnaker en [14].

real) debe ser el mundo posible que difiere mínimamente del mundo real en el cual A es verdadero ([13], p. 104). Según Lewis [9], hay que considerar un sistema de esferas de mundos centrado en el real (si el contrafáctico va a evaluarse en éste). Los mundos se ordenan, en este sistema de esferas, según una relación de similitud global. Cuando se trata de antecedentes concebibles (verdaderos en algunos mundos de estas esferas), $A > B$ es verdadero ssi existe una esfera de realización de A ⁴¹ en aquel sistema de esferas, cuyos mundos A son también todos mundos B . Pollock [12] toma por verdadero (en el mundo real) a un contrafáctico $A > B$ ssi B es verdadero en todos los mundos que se hallan más próximos al real, que Pollock considera que son los resultantes de cambios mínimos pensados para adecuar la verdad del antecedente.⁴²

Desde una perspectiva general, estos tres enfoques operan en la misma línea de evaluación para los contrafácticos, la cual recurre a los mundos más similares⁴³ (o efectúa cambios mínimos en el mundo real). En este sentido, son semejantes al de Goodman (sección 9). Para éste, el problema radical consistía en la incompatibilidad del antecedente con la historia real, y la intención principal de su tratamiento se dirigía a realizar omisiones mínimas que restablecieran la compatibilidad (y a confrontar los problemas que fueran derivaciones de esta estrategia). En nuestra exposición de la sección 9, dividimos en tres al conjunto de proposiciones verdaderas pertenecientes a (t_A, t_B) . El tercer conjunto estaba compuesto por proposiciones verdaderas que describían sucesos para los que el acontecimiento A tenía cierta relevancia causal negativa y, sin embargo, no figuraban en la clausura lógica del antecedente del esquema (2) del Análisis I (aunque cada proposición de este conjunto era cosustentable con el antecedente A). Sea $NR(A, t_{\bar{B}})$ el conjunto de proposiciones verdaderas pertenecientes a (t_A, t_B) que son cosustentables con A y que describen sucesos para los cuales el acontecimiento A posee alguna relevancia causal negativa. El tercer conjunto es entonces $(NR(A, t_{\bar{B}}) - CCS(A, t_{\bar{B}}))$.⁴⁴ Un enfoque dirigido a efectuar cam-

⁴¹ Una esfera de realización de A es la que contiene a un mundo en el cual A es verdadera.

⁴² [12], pp. 20, 70, 91.

⁴³ Utilizo aquí similitud en su sentido intuitivo. El libro [9] de Lewis (como ha sido interpretado normalmente) se apega a este uso, aunque en su posterior postura Lewis se adhiere al uso de la similitud intuitiva sólo al comparar mundos bajo aspectos específicos y no como rectora de la relación de ordenamiento global (cf. [10]).

⁴⁴ $CCS(A, t_{\bar{B}})$ se definió (en la sección 8) como el conjunto de consecuentes (perteneciente a (t_A, t_B)) de los semifácticos verdaderos con antecedente A que no son semifácticos irrel o p.p.

bios mínimos o a garantizar la similitud intuitiva máxima procuraría preservar las proposiciones de este tercer conjunto (por ejemplo, mediante la exigencia de que esas proposiciones sean verdaderas en las descripciones de los mundos posibles seleccionados por $f'(A > B)$). Sin embargo, ninguna de las proposiciones de este conjunto se preservaba en mi enfoque. En realidad, una consecuencia de mi análisis consistía en que para cualquier proposición C de este conjunto no sólo era falso $A > \sim C$ sino también $A > C$. En esto, mi enfoque difiere marcadamente en espíritu y dirección de las escuelas de pensamiento con las cuales lo estoy comparando: por lo tanto, el mío no es un enfoque de similitud intuitiva máxima o de cambios mínimos. (Las ilustraciones de esta clase de proposiciones que contribuyeron a fundamentar mi argumentación se dieron en los ejemplos 7 y 8, más arriba.)

12. Otras comparaciones

El enfoque de Lewis (en su *Counterfactuals*) equivale a considerar $A > B$ como verdadero (para un A concebible en el mundo real) ssi B es verdadero en todos los mundos que estén en alguna vecindad A del mundo real (es decir, en todos los mundos A que estén en alguna vecindad de realización de A del mundo real).⁴⁵ Me gustaría mostrar ahora que en mi modelo inferencial, el conjunto de descripciones de mundos posibles asociado con un contrafáctico $A > B$ de tal manera que B debe ser verdadero en todos sus miembros para que $A > B$ resulte verdadero no es un conjunto de mundos A más próximos, es decir, no es una vecindad A de la descripción del mundo real.

Este conjunto, $f'(A, t_B)$ (cf. sección 10), es el conjunto de descripciones de mundos posibles asociado con $A > B$ en la forma indicada en nuestro Análisis III. En todos sus miembros, A , M_A y las premisas implícitas de (t_A, t_B) ⁴⁶ son verdaderos (y cada uno de ellos es compatible con L). Ahora bien, no hemos puesto ninguna res-

⁴⁵ Donde la "proximidad" entre los mundos se entiende en términos de similitud intuitiva global. Ello caracterizaría a la posición de Lewis en su libro [9]. Pero el argumento que sostengo más abajo y según el cual mi enfoque difiere de su postura, también se aplica a otras posiciones afines en espíritu a la suya.

Voy a seguir la denominación de Lewis de un mundo posible (o de una descripción de un mundo posible) como mundo A (o descripción A de un mundo posible) en caso de que A sea verdadero en él.

⁴⁶ Las premisas implícitas pertenecientes a (t_A, t_B) son, desde luego, proposiciones fácticas (no legaliformes); la noción del intervalo temporal al cual pertenece una proposición se introdujo únicamente para esas proposiciones.

tricción a $M'_{(t_B, \infty)}$ ⁴⁷ para que una descripción de un mundo posible M' califique como miembro de este conjunto, excepto la compatibilidad con A , L y las premisas implícitas que pertenecen a tiempos anteriores a t_B . Obviamente, esto deja un amplio margen de libertad para que los miembros de este conjunto se asemejen, más o menos, a la descripción del mundo real por medio de variaciones en sus proposiciones pertenecientes al intervalo (t_B, ∞) . Además, recordemos que $NR(A, t_B) - CCS(A, t_B)$ es el conjunto de proposiciones verdaderas cosustentables con A y pertenecientes a tiempos dentro de (t_A, t_B) para las cuales A posee cierta relevancia causal negativa, pero que no figuran en la clausura lógica del antecedente del esquema (2) en el Análisis I (cf. arriba, sección 11). Por lo tanto, y de manera semejante, puesto que los miembros del conjunto $NR(A, t_B) - CCS(A, t_B)$ no estaban determinados por las premisas implícitas más A (y L), el problema de su inclusión (o la selección de sustitutos para los mismos) entre los miembros del conjunto $f'(A, t_B)$ permitía un correspondiente grado de libertad concerniente a la ubicación más cercana o más lejana de estos miembros con respecto a la descripción del mundo real. Así pues, tomemos una descripción de un mundo posible, m_1 perteneciente a $f'(A, t_B)$, cuya sobreposición con el conjunto $NR(A, t_B) - CCS(A, t_B)$, y con el conjunto de proposiciones verdaderas pertenecientes a (t_B, ∞) , es tal que la hacen tan *desemejante* de la descripción del mundo real como sea posible. Ahora tomemos alguna proposición falsa C , tal que $t_C = t_B$, $A > C$ sea verdadero, pero la posibilidad de inferir C a partir de las premisas implícitas y A pueda subvertirse eliminando un conjunto "menor" de proposiciones de las premisas implícitas que pertenecen a (t_A, t_C) . Ahora tomemos una descripción de un mundo posible, m_2 que incluya a A y a todas las premisas implícitas para $A > C$ *excepto* ese conjunto menor,⁴⁸ que sea compatible con L , que incluya a $\sim C$ (el cual es verdadero) y para el cual la selección de sus otros miembros en (t_A, t_B) y (t_B, ∞) esté pensada con el fin de hacerla lo más similar posible al mundo real. Ahora bien, esta selección de miembros de $NR(A, t_B) - CCS(A, t_B)$ concebida para obtener una similitud máxima con el mundo real en el intervalo (t_A, t_B) , aunque contiene a $\sim C$, debería bastar para hacer que m_2 fuera más semejante a la descripción del mundo real que m_1 , en tiempos previos a

⁴⁷ $M'_{(t_B, \infty)}$ es el subconjunto de proposiciones de la descripción M' de un mundo posible, perteneciente a intervalos temporales que se hallan enteramente dentro de (t_B, ∞) , es decir, enteramente después de t_B .

⁴⁸ Dada la especificación de las premisas implícitas (cf. Análisis I, sección 3) y las consideraciones mencionadas, obviamente las premisas implícitas para $A > B$ y $A > C$ son las mismas.

$t_{\bar{B}}$. No obstante, e incluso al margen de esto, ciertamente el grado de libertad para alejarse, o acercarse, de la descripción del mundo real disponible en el intervalo $(t_{\bar{B}}, \infty)$ es tal que la selección correcta de m_2 la haría evidentemente mucho más semejante que m_1 a la descripción del mundo real. (Recordemos que m_2 está construida para dar una similitud máxima en el intervalo $(t_{\bar{B}}, \infty)$, mientras que m_1 está pensada para dar una similitud mínima en ese intervalo.)

Pero mientras que m_1 fue elegido para que perteneciera a $f'(A, t_{\bar{B}})$, evidentemente m_2 no pertenece a $f'(A, t_{\bar{B}})$, pues hemos eliminado de ella a algunos miembros de las premisas implícitas de $A > B$, miembros pertenecientes a (t_A, t_B) (mientras que $\sim C$ es verdadero en m_2 , aun cuando $A > C$ es un contrafáctico verdadero). Pero hemos visto que m_2 es más semejante (probablemente *mucho* más) a la descripción del mundo real que m_1 . No obstante, para que $f'(A, t_{\bar{B}})$ sea una vecindad A , o un conjunto muy cercano de descripciones A de mundos posibles alrededor de la descripción del mundo real, debe ocurrir que, si alguna descripción A de un mundo posible está más cercana a la descripción del mundo real que otra descripción de un mundo posible de este conjunto, también debe pertenecer al conjunto. Pero m_1 y m_2 violan este requisito. De ahí que $f'(A, t_{\bar{B}})$ no es una vecindad A de descripciones de mundos posibles en torno a la descripción del mundo real.⁴⁹

⁴⁹ Esta nota está dirigida a los lectores de [6], puesto que trata un tema no desarrollado en este artículo. Sólo interesará a quienes hayan prestado una cuidadosa atención a la discusión del cap. 7, III, 5, de este libro, la cual intenta establecer que la teoría de los contrafácticos propuesta en ese lugar (y en este artículo) no es compatible con un enfoque de similitud intuitiva máxima. Como la mayor parte de la nota 33 de esa obra fue recortada por el editor, aprovecho esta oportunidad para agregarla aquí.

En el cap. 7, p. 177 (final) y p. 178 (principio), se daba por supuesto que las premisas implícitas más A más S' axiomatizarse mediante las premisas implícitas más A más un conjunto finito S de S' , para un tipo de cambio mínimo S' ; es decir, que S' podía ser generado por un subconjunto finito de él mismo, S , relativo a las premisas implícitas más A .

Sin embargo, para garantizar la viabilidad de este supuesto, se deben tomar ciertas medidas. Por un lado, la operación de sustracción “-” de la teoría de conjuntos utilizada en la definición de un cambio en “ $Q(A, M_{\bar{B}}) - R$ ”, puede reemplazarse por la operación “ Δ ”: $UD(A, t_{\bar{B}}, M_{\infty})\Delta R$ será $UD(A, t_{\bar{B}}, M_{\infty}) - R$ menos cualquier enunciado de $UD(A, t_{\bar{B}}, M_{\infty})$ del cual se sigue lógicamente un enunciado de R , o del cual se sigue lógicamente la disyunción de un miembro de R y una negación de un miembro de $UD(A, T_{\bar{B}}, M_{\infty})$. (Además, se podría preferir que la operación “ Δ ” se aplique en $(Q(A, M_{\bar{B}}) \cup UD(A, t_{\bar{B}}, M_{\infty}))\Delta R$, no sólo en $UD(A, t_{\bar{B}}, M_{\infty})\Delta R$, en vista de las citadas restricciones.)

La razón para introducir “ Δ ” es la siguiente: $UD(A, t_{\bar{B}}, M_{\infty})$ es infinito, y para cada miembro de R , sus conjunciones con miembros de $UD(A, t_{\bar{B}}, M_{\infty})$ darán un conjunto infinito que debe ser excluido de $UD(A, t_{\bar{B}}, M_{\infty})$ con el fin

13. La interpretación p.c.

Hasta aquí, la mayor parte de este artículo se ha dedicado a los contrafácticos del tipo d.n., es decir, aquellos cuyos antecedentes A son compatibles con la historia previa del mundo M_A . Me gustaría ahora bosquejar (aunque sea brevemente) la forma en que el análisis anterior se extiende a otros tipos. No voy a comentar en este lugar cómo hay que tratar a los contrafácticos con antecedentes imposibles o contrarios a las leyes. Pero entre los que no hemos considerado, el tipo de contrafácticos más común y útil es el de los contrafácticos con antecedentes falsos que, aunque resultan lógicos y metafísicamente posibles y compatibles con las leyes (= *co-legales*), son no obstante *incompatibles* con su historia previa. Para este tipo de contrafácticos contamos con una extensión natural de la interpretación d.n. (es decir, el análisis propuesto más arriba para el tipo d.n.): llamémosla la *interpretación p.c.* (p.c. por procesos conducentes). El hecho de que nuestro análisis de los contrafácticos del tipo d.n. pueda extenderse naturalmente a la interpretación p.c. confirma que el mismo se encuentra en el buen camino.

Normalmente, los antecedentes del tipo de contrafácticos mencionados que están sujetos a la interpretación p.c. no comprenden decisiones humanas o acciones espontáneas, sino más bien diversas situaciones que no habrían podido ocurrir, sin violar las leyes de la naturaleza, inmediatamente después de M_A . Pero habrían podido suceder de haber tenido lugar ciertos procesos no reales que “condujeran” a su ocurrencia. De este modo, consideremos el *ejemplo 9*:

Si Ford hubiera sido presidente en 1978, habría instado al sha de Irán a que empleara su ejército al máximo grado para permanecer en el poder. Para que en este caso se sostenga el antecedente,

de acomodar a \tilde{R} . De igual manera, si $q \in R$ y $p \in UD(A, t_{\tilde{B}}, M_{\infty}) \cup Q(A, M_{\tilde{B}})$, entonces, si $\sim p \vee q$ está en $UD(A, t_{\tilde{B}}, M_{\infty})$, la inconsistencia de $p, \sim p \vee q, \sim q$ contaminará a $UD(A, t_{\tilde{B}}, M_{\infty}) \cup \tilde{R}$ a menos que $\sim p \vee q$ también esté en R .

Sin embargo, esas restricciones de “ Δ ” tal vez no basten para garantizar la viabilidad de la axiomatizabilidad adecuada y finita de un cambio de esa índole (relativo a las premisas implícitas y A). En este caso, debe reducirse la redundancia lógica extensiva entre los miembros de las descripciones de mundos posibles. Por lo tanto, los conjuntos de enunciados utilizados en las descripciones de mundos posibles tendrán que ser reducidos con el fin de lograr la independencia lógica necesaria al interior de cada uno.

Debe considerarse que la discusión de la p. 178 y ss. en [6] procede debe ser contemplada como procedente bajo el supuesto de que es viable tal axiomatizabilidad finita para algún cambio mínimo semejante (relativo a las premisas implícitas y A), ya sea por restricciones como las de la operación “ Δ ” o por una reducción apropiada de la redundancia lógica de los diversos conjuntos en discusión.

algún proceso debería haber conducido a Ford a la presidencia en 1978 (en lugar de ser presidente entonces como resultado de una acto *deus ex machina* o de algún milagro).⁵⁰ De esta manera, el trasfondo natural sobre el que se debe considerar ese supuesto contrario a los hechos se compone de diferentes procesos que divergen del mundo real en algún punto antes de 1978⁵¹ y que “conducen” al antecedente. Cada uno de estos procesos es una serie de hechos o situaciones, y el contrafáctico debe ser evaluado sobre el trasfondo de cada uno de los procesos por separado. Una extensión natural de nuestro tratamiento de los contrafácticos del tipo d.n. nos da la forma en que debe realizarse dicha evaluación. Tendría que seguirse el consecuente B , para una determinada descripción del proceso, de las premisas que incluirían a las leyes, la historia del mundo anterior al proceso, la descripción del proceso, el antecedente A y un

⁵⁰ Como lo señalé en la sección 2, sostengo que el marco natural para el análisis de los contrafácticos comprende el no-determinismo como supuesto operativo. Para más detalles sobre este tema, cf. [6], cap. 9, V.

⁵¹ El momento de divergencia de los procesos escogidos, para un contrafáctico dado $A > B$ del tipo p.c., está determinado según las siguientes restricciones. En primer lugar, consideremos los puntos temporales anteriores a t_A que, como puntos de divergencia potencial de los procesos, permitirían la adecuación de los aspectos *de re* de $A > B$. Es decir, en una interpretación p.c. de $A > B$ y de un proceso que tenga ese punto de divergencia, las ocurrencias *de re* de los términos singulares de $A > B$ (o incluso de términos de clases u otros términos del análisis de Kripke-Putnam) seguirían teniendo una interpretación *bona fide* para esos puntos de divergencia, y el contrafáctico calificaría como lo que denominé *mundo de re*. Los contrafácticos que escapan completamente a la última restricción (o que se interpretan así), deben ser explicados según lo que denominé *mundo de dicto* y no caen dentro de la interpretación p.c. (Para un tratamiento de ellos, cf. [6], cap. 9, IV, 2. Para otros señalamientos con respecto al tema del tiempo de divergencia, cf. [6], cap. 9, nota 48. Sobre las restricciones impuestas por una ocurrencia *de re* de un término singular en un contrafáctico, cf. [5], secciones X y XI, y [6], cap. 9, IV, 3.)

Esos puntos temporales comprenden, entonces, a un intervalo de tiempo ($t(A > B)$, t_A) (cuyo punto inicial, $t(A > B)$, es función del contrafáctico $A > B$). Consideremos ahora los puntos de *este intervalo* para los cuales la función $P(A/M_t)$ posee valores *máximos*. Escojamos entonces al último de ellos. Este el tiempo de divergencia de los procesos seleccionados para $A > B$.

(Esto es así siempre que exista ese último punto; no lo habría si una creciente serie de convergencia de esos puntos —la cual no incluye a su propio límite—, fuera una posibilidad admisible para nuestros propósitos, y si además no existiera ninguno de esos puntos después de dicho límite. Si este caso constituye una posibilidad admisible (y no es totalmente claro que así lo sea) y si ese punto límite temporal es t , entonces sería viable exigir, aparentemente, que los procesos seleccionados diverjan en esos puntos dentro de un intervalo $(t - \epsilon, t)$, para un ϵ muy pequeño. El contrafáctico es verdadero ssi para ese ϵ todos los procesos seleccionados, por medio de sus esquemas respectivos, hacen que la interpretación p.c. arroje un resultado afirmativo.)

incremento de la información fáctica verdadera de la parte de la descripción del mundo real perteneciente al momento del proceso y el intervalo (t_A, t_B) . Este incremento se llevaría a cabo a lo largo de líneas semejantes a las asumidas en la interpretación d.n.: en cada paso del proceso se retendrían proposiciones de la descripción del mundo real para las cuales la parte del proceso que no es posterior a esta etapa resulta causalmente irrelevante o tiene una relevancia causal puramente positiva. El incremento de la información perteneciente al intervalo (t_A, t_B) comprendería a las proposiciones fácticas verdaderas que describen sucesos para los cuales el proceso y el acontecimiento antecedente son causalmente irrelevantes o poseen una relevancia causal puramente positiva. Para cada uno de estos procesos, B debería seguirse de las premisas reunidas en torno a este proceso en la forma descrita para que el contrafáctico sea verdadero.

Como otro ejemplo (*ejemplo 10*), tomemos un caso del tipo Gettier. Supongamos que Juan tiene a dos hombres en su oficina, los señores Nogot y Havit. Este último posee un Ford y el primero no. Pero Juan supone que ninguno de ellos posee un Ford. Ahora consideremos: Si Juan hubiera creído que alguien en su oficina poseía un Ford, lo habría sabido. Esto es falso: aunque Juan habría podido llegar a esa creencia por el proceso de ver al señor Havit conduciendo su propio Ford, también habría podido llegar a ella mediante el proceso de ver al señor Nogot conduciendo un Ford rentado.⁵² Por lo tanto, el consecuente no resultaría verdadero por un proceso así, y el contrafáctico es en realidad falso.

Pero, desde luego, no *todo* proceso conducente a A debe ser tomado en cuenta para la evaluación de $A > B$. Entre tales procesos figuran algunos muy esotéricos, a los cuales preferiríamos ignorar. De esta manera, supongamos que se considera el contrafáctico (*ejemplo 11*): Si Reagan hubiera asumido la presidencia el 21 de enero de 1977, habría comenzado un proceso de propaganda militar masiva a fines de los setenta. Ahora bien, un tipo de proceso conducente al antecedente divergiría de la descripción del mundo real durante la época de la convención del G. O. P. en 1976, en la cual la delegación de Misisipí vacilaba entre Ford y Reagan. En un proceso así, la delegación habría decidido dar su apoyo total a Reagan, otorgándole de este modo la nominación, y Reagan habría procedido entonces a ganar las elecciones generales. Pero en otro proceso que conduce al antecedente, la Unión Soviética habría ocupado a los Estados Unidos, convirtiéndolo en un satélite soviético y nombrando a Reagan como su presidente títere. No es preciso que en este proceso el conse-

⁵² Y, desde luego, se pueden considerar procesos mediante los cuales llegaría a creer que el señor Havit tiene un Ford sin saber que el señor Havit lo tiene.

cuenta resulte verdadero. Pero tampoco es preciso que convierta en falso al contrafáctico de arriba. No debe ser tomado en cuenta entre los procesos conforme a los cuales hay que evaluar al contrafáctico.

Como otra ilustración (*ejemplo 12*), supongamos que Juan está sentado en la habitación *I* en Nueva York en el momento *t*, mientras que en una habitación *II*, adyacente y vacía, hay un cerillo y una caja de cerillos. Consideremos el famoso contrafáctico de Goodman: Si este cerillo hubiera sido frotado, se habría encendido. Ahora bien, un tipo de proceso conducente al antecedente sería que Juan se hubiera levantado, pasara de la habitación *I* a la habitación *II*, tomara el cerillo y la caja y frotara el cerillo. Otro tipo de proceso consistiría en que una inmensa ola proveniente del mar inundara la zona y lanzase al cerillo de tal modo que éste se frotase contra la caja. Seguramente, el cerillo no se encendería en este caso. Pero, nuevamente, no es preciso que este tipo de proceso altere la verdad del contrafáctico citado: es demasiado esotérico como para ser tomado en cuenta.

En consecuencia, sólo hay que seleccionar algunos procesos para evaluar un contrafáctico del tipo en discusión. El procedimiento anterior únicamente debe aplicarse a algunos procesos, de tal modo que el contrafáctico va a resultar verdadero ssi su consecuente es inferible para cada uno de esos procesos. ¿Cuáles deben ser los procesos seleccionados? Esencialmente, *los que presentan más probabilidades de haber conducido al antecedente*. De esta forma, los esotéricos quedan excluidos. Por lo tanto, en el ejemplo 11, no se escogería al proceso en que la Unión Soviética ocupaba a Estados Unidos, pues presenta mucho menos posibilidades de haber conducido al antecedente que el primer proceso descrito en la exposición de ese ejemplo. No obstante, los dos procesos descritos en el ejemplo del tipo de Gettier (*ejemplo 10*) presentarían el mismo grado de probabilidad de haber conducido al antecedente en un marco natural hipotético para este ejemplo y, de hecho, hay que escoger procesos de estos dos tipos para evaluar el contrafáctico del ejemplo 10 (que en realidad es falso bajo el marco al que se aludió).

Es necesario desentrañar dos nociones fundamentales que se utilizan en este análisis, la noción de proceso "conducente" al suceso-antecedente y la noción de procesos con más probabilidades de haber conducido al suceso-antecedente. Con respecto a la primera, un requisito importante que debe satisfacer un proceso descrito por "*P*" con el fin de calificar como proceso que pudo haber "conducido" al suceso *A* es que:

$$P(A/P \& W_P) > P(A/W_P).^{53}$$

⁵³ Según nuestra notación, M_P es la historia del mundo hasta el comienzo del proceso *P*.

No obstante, dado el alcance de este artículo, no es posible exponer aquí un mayor desarrollo de estas nociones.⁵⁴

Aun así, debe notarse que la interpretación p.c. no es más que una extensión natural del análisis para los contrafácticos del tipo d.n. presentado en estas páginas. De tal forma, en primer lugar, el espíritu de preservar $CIP(A, t_B)$ ⁵⁵ se mantuvo en la interpretación p.c. En segundo lugar, puesto que en los casos en que se aplica la interpretación p.c. la historia previa M_A no podía ser preservada (bajo pena de inconsistencia), simplemente pasamos de considerar un esquema bajo el Análisis I a considerar una serie de esquemas análogos. Y en tercer lugar, fiel al espíritu del primer análisis, la caracterización de la selección del proceso debe llevarse a cabo con la ayuda de herramientas probabilísticas.⁵⁶

⁵⁴ Para una mayor elaboración, cf. [6], cap. 9, *I*.

En [6], cap. 9, n. 19, presento una concepción de los procesos en términos de conjuntos ordenados de descripciones de estados, y muestro cómo ello permite una definición muy natural del concepto de los procesos que tienen más probabilidades de haber conducido a A . Otra orientación para buscar la elaboración de la noción de un proceso es la siguiente. Supongamos, en el espíritu de mi enfoque, que reconocemos la noción de probabilidad de un enunciado en el momento t (cuando t es un intervalo temporal o algo semejante), dada la historia del mundo hasta cierto momento anterior determinado. Una forma obvia de concebir un proceso es como una función de los puntos temporales a las descripciones de los estados; sin embargo, otra manera de representar ese proceso consiste en indicar los segmentos temporales (o ciertos conjuntos temporales más complejos) en los cuales los enunciados son verdaderos (en el proceso). De este modo, si está definida la probabilidad de que un enunciado sea verdadero en un intervalo (o algo semejante), dado un tiempo anterior t , también tendremos entonces la definición de la probabilidad correspondiente de un proceso (dado un lenguaje cuantitativo) como la probabilidad de la intersección en que todos sus enunciados son verdaderos en los intervalos (o similares) en que son verdaderos (en este proceso). Si la probabilidad de un proceso está definida, dada la historia del mundo hasta un tiempo anterior t , entonces la noción de los procesos más probables (de un conjunto dado) también puede definirse con facilidad. Se llega a una caracterización de la noción de los procesos que poseen más probabilidades de haber conducido a A (dado un tiempo de divergencia t , $t > t_A$) limitando el conjunto a los procesos que terminan en A . (Comparemos esto con la noción de $P(P/M_t \& A)$.) La elaboración de esta línea de pensamiento exige la demostración de que puede eliminar la dependencia de las probabilidades respecto de la riqueza con que se describan los procesos.

⁵⁵ $CIP(A, t_B)$ fue definido más arriba, sección 6.

⁵⁶ También puede verse claramente que la interpretación d.n. es un caso especial de la interpretación p.c. cuando se advierte que el análisis p.c. arroja el análisis d.n. en el momento en que se selecciona a t_A como tiempo de divergencia, es decir, cuando el único proceso considerado es el proceso vacío.

14. *Ambigüedad d.n./p.c.*

Hasta aquí hemos tomado la interpretación d.n. para aplicarla a contrafácticos cuyos antecedentes son compatibles con su historia previa. Se tomó la interpretación p.c. para aplicarla a casos en que los antecedentes eran incompatibles con sus historias previas.⁵⁷ Ahora bien, resulta evidente que no es posible extender la interpretación d.n. a los casos que son del dominio de aplicación de la interpretación p.c., pues tendríamos entonces un antecedente inconsistente en el esquema inferencial. En tales casos, debe recurrirse únicamente a la interpretación p.c. Sin embargo, ninguna consideración sobre consistencia impide que la interpretación p.c. se aplique a contrafácticos en los cuales el antecedente es compatible con su historia previa. Pero en la gran mayoría de estos casos la interpretación p.c. no tiene ninguna base intuitiva. Además, es a causa de que, en términos de extensión, la interpretación d.n., y no la p.c., da los resultados correctos en estos casos, por lo que se aplica la primera en lugar de la segunda. Desde luego, la interpretación d.n. es mucho más simple y resulta natural esperar que se aplique siempre que sea posible, recurriendo a la interpretación p.c., más compleja, únicamente cuando la primera no pueda utilizarse: cuando el supuesto contrario a los hechos puede expresarse naturalmente en la historia previa real como trasfondo para su evaluación, no tendría sentido recurrir a procesos alternativos, no reales, que hubieran podido conducir al antecedente.

Sin embargo, aunque queda totalmente claro que la interpretación d.n. (no la p.c.) es la interpretación correcta para su área de aplicación, y la única interpretación correcta para la gran mayoría de los casos que se hallan en ella, hay que preguntarse qué ocurre en la zona limítrofe de los dominios de ambas interpretaciones. Esta sería una zona en la que el antecedente A es compatible con la historia previa, pero su probabilidad condicional dada la historia previa, $P(A/M_A)$, es lo suficientemente pequeña. Puesto que la interpretación p.c. puede aplicarse en esos casos que caen dentro de la zona crepuscular de la interpretación d.n., tenemos que considerar la posibilidad de que también pueda percibirse en ellos su viabilidad. Y efectivamente, encontramos que en un cierto tipo de estos casos limítrofes aparece un fenómeno de *ambigüedad*: los contrafácticos pueden explicarse de dos formas diferentes, las cuales pueden arrojar dos valores de verdad distintos. Una observación más detenida revela que la ambigüedad se encuentra entre las interpretaciones

⁵⁷ Ambas interpretaciones son aplicables únicamente a los contrafácticos cuyos antecedentes son proposiciones fácticas que son lógicas, nomológicas y metafísicamente posibles.

d.n. y p.c. En estos casos, ambas poseen una base intuitiva y, en consecuencia, producen dos formas diferentes en las que se interpreta el contrafáctico en cuestión. Consideremos, como ejemplo, el contrafáctico de Downing: ayer Jaime y Jacobo discutieron y Jacobo todavía está muy enojado. Concluimos que si hoy Jaime le pidiera ayuda a Jacobo, éste se rehusaría. Pero veamos: Jaime tiene su orgullo. Nunca pediría ayuda después de semejante discusión; si realmente Jaime le solicitara ayuda a Jacobo hoy, tendría que no haberse producido ninguna discusión ayer. En este caso Jacobo se hallaría en su disposición acostumbrada, generosa. Por lo tanto, si Jaime le hubiera pedido ayuda a Jacobo hoy, éste, después de todo, lo habría ayudado.⁵⁸

Ahora bien, la primera explicación, en la cual el contrafáctico "Si Jaime hubiera solicitado ayuda a Jacobo, éste se habría rehusado" (llamémoslo: $A > B$) resultaría correcto, habría sido arrojada por la interpretación d.n. Dadas las circunstancias en t_A , después de la discusión, con la historia previa M_A intacta, en vista del carácter de Jacobo y en ausencia de circunstancias atenuantes, Jacobo no habría ayudado a Jaime. Pero en la segunda explicación, según la cual $A > \sim B$ resulta verdadero, M_A no quedaría intacta, sino que más bien se considerarían procesos bajo los cuales A es verdadero mientras no se haya producido una discusión. En este caso estaríamos explicando el contrafáctico según los lineamientos de la orientación p.c. En efecto, en la segunda interpretación, y dado el carácter de Jaime, los procesos que presentan más probabilidades de haber conducido al antecedente serían aquellos en los que no tuvo lugar la discusión. Dados estos procesos y el carácter de Jacobo, resultaría $\sim B$ (su ayuda a Jaime). La ambigüedad observada por Downing aparece como una ambigüedad entre las interpretaciones d.n. y p.c.

De esta manera, para casos limítrofes, en los que $P(A/W_A)$ es lo suficientemente pequeño, resulta viable recurrir a procesos alternativos, no reales, con lo cual emerge la interpretación p.c. y, en varios de estos casos, aparece una ambigüedad entre ambas interpretaciones. Nuestra pregunta ahora consiste en si es posible formular una hipótesis más específica sobre el alcance de esta zona limítrofe en la que interactúan ambas interpretaciones. Hasta aquí hemos considerado que $P(A/M_A)$ es suficientemente pequeño. En estos casos el antecedente del contrafáctico aún es estrictamente posible, dadas

⁵⁸ P. B. Downing, citado por Lewis en [10], p. 457. Cf. también [6], cap.9, sección III, para un tratamiento de la ambigüedad d.n./p.c. y de los contrafácticos de Downing. Vale la pena notar que los casos que considero como ejemplos de esta ambigüedad fueron tratados por otros como generadores de retrocesos (según la interpretación que considero como p.c.).

su historia previa y las leyes, pero no pasa de allí. No obstante, esto solo no puede justificar que se recurra a la interpretación p.c.: ésta, más compleja, se aplicará únicamente si pudiese resultar mucho mejor que la interpretación d.n. ¿Qué significa que la interpretación p.c. “resulte mejor”, que haga “más posible” al antecedente? Ciertamente la contraparte en la interpretación p.c. de $P(A/M_A)$ es $P(A/M_t)$, donde t es el momento de la divergencia bajo la interpretación p.c., el momento en que los procesos seleccionados se apartan de la historia real. De esta forma, el retroceso en el tiempo de M_A a una historia previa M_t (con $t < t_A$) puede permitir una mayor compatibilidad del antecedente A con esa historia previa, en el sentido de que $P(A/M_t)$ es más alto que $P(A/M_A)$. Resulta así natural preguntarse cómo exactamente se relacionan ambos cuando se busca una caracterización más precisa de la zona limítrofe.

Una forma natural de formular la condición que especifica la gama de casos sujetos a la ambigüedad d.n./p.c. (que son los casos en los cuales A resulta compatible con M_A) es la siguiente:

- (11) Estos son los casos para los cuales $P(A/M_A) < P(A/M_t)$ para un $t < t_A$ (momentáneo).⁵⁹ De otro modo, si $P(A/M_A) > P(A/M_t)$, para cualquiera de estos t , la interpretación d.n. debe ser la dominante.⁶⁰

Que la condición $P(A/M_A) < P(A/M_t)$ prevalece en el caso del contrafáctico de Downing es bien evidente: el resultado es el caso del tipo ambiguo.

Ahora bien, se puede considerar el caso en que la última parte de esta condición (es decir, que si $P(A/M_A) > P(A/M_t)$, para cualquier $t < t_A$, entonces predomina la interpretación d.n.), se obtiene del análisis en la siguiente forma. Supongamos que $P(A/M_A) > P(A/M_t)$ para cualquier t de esos. Supongamos que admitimos que el proceso vacío $\{A\}$ (divergente en M_A) sea uno de los procesos seleccionados bajo la interpretación p.c. (cuando A es compatible con M_A). Si⁶¹ el momento de divergencia para la interpretación p.c. debe ser el último punto máximo de $P(A/M_t)$, ($t < t_A$), entonces, en el caso en consideración, t_A es el punto de divergencia y el proceso vacío $\{A\}$ es el único proceso seleccionado (es el único

⁵⁹ Sin embargo, t debe permitir que se respeten las ocurrencias *de re* de los términos singulares del contrafáctico; cf. n. 63, abajo.

⁶⁰ Esta condición supera a la considerada en [6] (pp. 248–251) conforme a la cual $P(A/M_A)$ debe ser sopesado con $PC(A)$, la plausibilidad característica de A .

⁶¹ Como se sugirió en ese lugar, cap. 9, n. 18.

proceso que termina en A y diverge en t_A). Ahora bien,⁶² cuando M_A es el momento de divergencia, la interpretación p.c. se viene abajo dando lugar a la interpretación d.n. De aquí que, dado el supuesto anterior, la predominancia de la interpretación d.n. va a mantenerse bajo la forma del colapso de la interpretación p.c. que dará lugar a la interpretación d.n. en el caso en consideración. La interpretación d.n. aparece así como un caso especial de la interpretación p.c. cuando $P(A/M_A) > P(A/M_t)$ para cualquier t, t_A (en cuyo caso A es compatible con M_A , puesto que $P(A/M_A) > 0$). Por lo tanto, que la aplicación de la interpretación d.n. es exclusiva cuando $P(A/M_A) > P(A/M_t)$ para cualquier $t < t_A$, es consecuencia de la misma interpretación p.c.

15. *Contrafácticos con antecedentes verdaderos*

Los contrafácticos con premisas verdaderas son un grupo extraño. Parece existir una convención práctica de no utilizar la construcción subjuntiva cuando el antecedente está reconocido como verdadero. En vista de ello, el hablante no puede enunciar contrafácticos con antecedentes verdaderos, reconocidos como tales, sin que lo que dice suene muy raro. Hasta aquí, hemos restringido nuestro examen a los contrafácticos con antecedentes falsos. No obstante, algunas consideraciones indican que también es deseable asignar valores de verdad a contrafácticos con antecedentes verdaderos. Una de esas consideraciones se presenta en el caso de los contrafácticos anidados. Otra es que esos contrafácticos se emplean en diversas aplicaciones (por ejemplo, el análisis contrafactual del conocimiento realizado por Nozick, en particular su cuarta condición).

Sin embargo, puesto que los contrafácticos con antecedentes verdaderos no parecen evocar intuiciones claras sobre sus valores de verdad, parece mejor, desde el punto de vista metodológico, aplicar también a ellos nuestra teoría de los contrafácticos y permitir que los valores de verdad resultantes se asignen donde caigan, obteniendo así unidad teórica sobre un dominio más amplio (en vez de aventurar un análisis independiente o de eludir cualquier tipo de análisis). En este artículo hemos considerado las interpretaciones d.n. y p.c. Ambas pueden aplicarse naturalmente al caso de los contrafácticos con antecedentes verdaderos, y los resultados se presentan como sigue:

En primer lugar, bajo la interpretación d.n., para un contrafáctico $A > B$ con un antecedente verdadero, $A > B$ es verdadero ssi: $\{A\} \cup M_A \cup CIP(A, t_B) - L \rightarrow B$. Como el antecedente de este esquema se compone de proposiciones verdaderas, resulta evidente que en la

⁶² Como se señaló en [6], cap. 9, I, 5, p. 242.

interpretación d.n. ningún contrafáctico con antecedente verdadero y consecuente falso es verdadero. Los contrafácticos de esta clase, bajo esta interpretación, son aquellos cuyo consecuente B está determinado por ciertas partes de la historia $M_{\bar{B}}$ (a través de L). Sin embargo, estos contrafácticos no resultan distorsionados, es decir, verdaderos ssi su consecuente es verdadero (como resultan en los análisis de Stalnaker y Lewis).

Con la interpretación p.c., el análisis de los contrafácticos con antecedentes verdaderos exige la consideración de una serie de procesos conducentes al antecedente. No obstante, puesto que éste es verdadero, parece inapropiado dejar de considerar a los procesos *reales* con más probabilidades de conducir al antecedente dentro del conjunto de procesos seleccionados, aun cuando tales procesos reales, extraídos del curso real de los acontecimientos, no cuenten entre los procesos que tienen más probabilidades de haber conducido al antecedente. Por lo tanto, el análisis de estos contrafácticos en la interpretación p.c. se presenta como sigue:

Un contrafáctico $A > B$ es verdadero, cuando A es verdadero, sólo en caso de que se observe lo siguiente para cada miembro del conjunto seleccionado de procesos (los que tienen más probabilidades de haber conducido al antecedente, incluyendo a los procesos reales adecuados que conducen al antecedente): Para cada uno, tomemos la historia del mundo anterior al proceso, más el proceso mismo; completemos el proceso con descripciones de los hechos reales (hasta t_A) para los cuales el proceso (tomado gradualmente, paso a paso) es causalmente irrelevante o tiene una relevancia causal puramente positiva; agreguemos los hechos de (t_A, t_B) para los cuales el proceso más el antecedente son causalmente irrelevantes o poseen una relevancia causal puramente positiva; agreguemos también las leyes. El consecuente debe seguirse de todo ello. Esto debe ser verdadero para cada proceso seleccionado.

En lugar de tomar en cuenta sólo partes de la historia real anterior al antecedente por medio del proceso real, como acabamos de hacer, otra opción natural sería tomar únicamente la historia real M_A . De esta manera, con la interpretación p.c. ajustada a esta sugerencia, el esquema p.c. (ahora sin ningún proceso real) más el esquema d.n. deben servir para que los contrafácticos con antecedentes verdaderos resulten verdaderos.

Tanto la interpretación d.n. como la p.c. provocan, como debe ser, que un consecuente falso garantice la falsedad de un contrafáctico con un antecedente verdadero. Pero brindan, respectivamente, un análisis más restringido y uno más rico de los contrafácticos con antecedentes y consecuentes verdaderos: la interpretación p.c. ofrece

un examen más rico y con más restricciones, dificultando el resultado verdadero de este contrafáctico. Como las intuiciones no son aplicables en este caso, hay que recurrir a otras consideraciones que contribuyan a decidir entre ambas interpretaciones. Por un lado, los contrafácticos de esta clase parecen hallarse más emparentados con los del tipo d.n. (debido a que sus antecedentes, por ser verdaderos, son seguramente compatibles con su historia previa). Pero otra vez, como son básicamente forzados, con fundamentos intuitivos endeble, existe la libertad de elegir la interpretación que mejor se acomode a los propios propósitos y aplicaciones teóricas. Es totalmente admisible decretar, en un análisis particular en que se utilizan contrafácticos con antecedentes verdaderos, que hay que tratarlos con la interpretación p.c. si se desea una explicación más rica. No obstante, en un nivel teórico más sistemático, se puede considerar la aplicación del examen desarrollado más arriba (sección 14) concerniente a la ambigüedad d.n./p.c., y por lo tanto la comparación de $P(A/M_A)$ con el último y más alto valor de $P(A/M_t)$ para un t previo a t_A (aunque lo bastante tardío como para permitir un adecuado punto de divergencia)⁶³ con el fin de determinar por cuál interpretación se va a optar.

16. *La teoría de las esferas y la característica de la independencia*

Se pueden distinguir dos tipos de teorías sobre los contrafácticos, al margen de si se formulan en términos de mundos posibles o descripciones de mundos posibles. Una de ellas puede denominarse *teoría de las esferas*. Según ella, existe una relación de tres términos " \leq_i " (para un contexto dado) definida enteramente dentro del conjunto de los mundos en consideración⁶⁴ (que es reflexiva y transitiva), con la relación derivativa " $<_i$ " definida por medio de ella como: $j <_i k \iff j \leq_i k \text{ y } k \not\sim_i j$. Esta relación permite definir la noción de esferas (del tipo lewisiano). Un contrafáctico $A > B$ (para un A concebible) es verdadero en i (en este contexto) ssi B es verdadero

⁶³ De este modo, una ocurrencia *de re* de un término singular denotativo en un contrafáctico impone una restricción a la interpretación p.c. de ese contrafáctico, de ser aplicable, según la cual el punto de divergencia de los procesos seleccionados debe ser, en general, posterior al momento del nacimiento del individuo denotado por ese término singular. Para una mayor elaboración, cf. [5], secciones X, XI.

⁶⁴ Voy a ignorar, para los propósitos de la presente exposición, las relaciones de ordenamiento definidas parcialmente que impiden las comparaciones. Para un tratamiento de las mismas, cf. [11] de Lewis.

en alguna vecindad A de i .⁶⁵ El ordenamiento " \leq_i " puede ser el de la similitud intuitiva global o de similitudes en algunos aspectos que están contrapesados u ordenados por prioridades,⁶⁶ o por cualquier otro ordenamiento.

El otro tipo puede denominarse *teoría de la función selectiva*. Según ésta, existe una función selectiva f , cuyos enunciados comprenden contrafácticos y mundos (o descripciones de mundos posibles y, quizás, un contexto), y que selecciona, para argumentos dados (que incluyen un antecedente A del contrafáctico), un conjunto de mundos A como su valor. Un contrafáctico $A > B$ es verdadero (en el mundo i) ssi B es verdadero en todos los mundos seleccionados por f (para i).⁶⁷

Ahora bien, una aproximación clásica a los contrafácticos es la teoría metalingüística. Sin embargo, ésta no es más que un caso especial de tipos más generales de teorías con las cuales bien podemos estar familiarizados. Según la *teoría metalingüística*, $A > B$ es verdadero ssi B puede inferirse de un conjunto de proposiciones verdaderas (que satisfagan ciertas restricciones) junto con A . Mi propio examen de los contrafácticos del tipo d.n. constituye un caso especial de la teoría metalingüística. Pero en lugar de optar por una simple cuantificación existencial agregando algunas restricciones que calificarían a ese conjunto de proposiciones verdaderas como admisibles, en mi análisis se exige que las premisas implícitas más el antecedente determinen al conjunto de proposiciones requerido: el papel de la función de las premisas implícitas $f(A, t_B)$ consiste en *especificar* este conjunto de proposiciones.

Supongamos que la *teoría metalingüística generalizada* es igual a la teoría metalingüística sin la exigencia de que las "premisas adicionales" requeridas sean verdaderas. Supongamos que la *teoría metalingüística generalizada de multiplicación (TMGM)* exige varios de esos conjuntos Z_i . De esta manera, según *TMGM*, un contrafáctico $A > B$ es verdadero ssi para algunos conjuntos Z_i (especificados de cierta forma) $\{A\} \cup Z_i \rightarrow B$ para cada Z_i . Desde luego, la teoría metalingüística es un caso especial de la teoría metalingüística generalizada que, a su vez, es un caso especial de *TMGM*.

En consecuencia, es evidente que el examen general de los contrafácticos que presenté más arriba (sección 13) como interpretación

⁶⁵ Una vecindad A de i es una vecindad de i en $\mathcal{W}_A \cup \{i\}$, donde \mathcal{W}_A es el subconjunto de mundos A (de los mundos considerados). No tomo en cuenta el caso de los antecedentes inconcebibles.

⁶⁶ Como el modelo de Lewis en [10].

⁶⁷ Si nos limitamos a los antecedentes posibles garantizaremos los mundos posibles en que éstos son verdaderos.

p.c., de la que la interpretación d.n. era un caso especial,⁶⁸ es un modelo de *TMGM*. Cada conjunto Z_i correspondería a uno de los procesos seleccionados que se mencionaban en ese lugar, y tendría que observarse un esquema inferencial modificado apropiadamente para cada uno de ellos a fin de que el contrafáctico resultara verdadero. Por lo tanto, mi enfoque general de los contrafácticos *no* cae dentro del dominio de la teoría metalingüística como se la concibe normalmente.

La aproximación de Lewis cae por completo dentro del dominio de la teoría de las esferas. La de Stalnaker es un tipo de teoría de la función selectiva (para la cual los valores de la función selectiva pueden tomarse como exclusivos). Resulta natural reformular las teorías metalingüísticas como teorías de la función selectiva con la ayuda de las descripciones de mundos posibles. La teoría presentada en esta obra (limitada al tipo d.n.) puede formularse de esa manera, como en (10) de la sección 10 más arriba. Lo mismo también se aplica a los modelos *TMGM* en general y, en particular, al examen p.c. desarrollado arriba (sección 13). Las descripciones de mundos posibles seleccionadas en él serán las que hagan verdadero al antecedente y a las premisas implícitas bajo uno de los procesos seleccionados.

Ahora bien, una característica importante de la teoría de las esferas es que el ordenamiento de mundos es independiente de todo contrafáctico particular. Esto constituye la *característica de la independencia para la teoría de las esferas*. De hecho, los diversos modelos de Lewis que involucran relaciones de similitud intuitiva global o de similitud por aspectos (contrapesados u ordenados por prioridades), son modelos de la teoría de las esferas, y todos ellos poseen en efecto la característica de la independencia. Este rasgo es crucial para la concepción de Lewis, pues resulta esencial para fundamentar la concepción de esferas característica de su teoría, que se establece de una vez por todas (quizás para un contexto dado) al margen de cualquier contrafáctico particular. Si la relación ordenadora fuera función del contrafáctico evaluado, las esferas cambiarían en general con los contrafácticos que se están evaluando.

De igual modo, un rasgo importante que puede o no poseer la función selectiva consiste en que *sólo* la dependencia funcional respecto de $A > B$ es una dependencia respecto de \mathcal{W}_A , el subconjunto de los mundos A (de los mundos considerados). Es decir, si $f(A > B, i)$ es una función selectiva (ignorando otros argumentos posibles), entonces para un $g(\mathcal{W}_A, i)$, $f(A > B, i) = g(\mathcal{W}_A, i)$. Se dice que la función selectiva que presenta este rasgo posee la característica de la independencia.

⁶⁸ Cf. el final de la sección 16 y n. 56.

En la teoría de Stalnaker, la función selectiva $f(A > B, i)$ selecciona al mundo A más cercano a i . De esta forma es equivalente a una función $g(\mathcal{W}_A, i)$ que selecciona al mundo más cercano a i en \mathcal{W}_A . De aquí que la función selectiva de Stalnaker posea la característica de la independencia.

En el modelo inferencial presentado arriba para el tipo d.n., las restricciones de las premisas implícitas eran función del antecedente A . En este modelo, había que preservar al conjunto $CIP(A, t_B)$ (cf. sección 6), que depende mucho de A (y t_B). Dicho conjunto incluía proposiciones especificadas en términos de relevancia causal y relevancia causal puramente positiva que poseía el suceso A con respecto a los acontecimientos descritos por esas proposiciones. Entonces, al seleccionar las descripciones de mundos posibles, según el examen propuesto en este artículo, nos apoyábamos mucho en los rasgos de A que no fueran el requerimiento de su simple inclusión en las descripciones seleccionadas de mundos posibles: las descripciones A de mundos posibles seleccionadas tenían que incluir, además, proposiciones para las cuales A presentaba las relaciones apropiadas de relevancia causal y relevancia causal puramente positiva. Este es también, *mutatis mutandis*, el caso con respecto a la interpretación p.c. más general desarrollada arriba. De esta manera, en contraste con la teoría de Stalnaker, el examen presentado en esta obra *no* posee la característica de la independencia, puesto que la elección de descripciones de mundos posibles dependía de las características del antecedente A *al margen* de su verdad en la descripción de mundos posibles seleccionada, debido al papel central que desempeña $CIP(A, t_B)$ en nuestro examen.

Ahora bien, se puede demostrar que algunos modelos de la teoría de la función selectiva que no poseen la característica de la independencia (en particular, algunos modelos de la teoría metalingüística generalizada y, por lo tanto, de *TMGM*) no tienen modelos equivalentes en la teoría de las esferas.⁶⁹ Como sostuve ampliamente en

⁶⁹ En [6], cap. 8, sección V, 4, demuestro que algunos modelos de la teoría de la función selectiva que no poseen la característica de la independencia, no son equivalentes a ningún modelo de la teoría de las esferas. Para que una función selectiva tenga la característica de la independencia debe ser totalmente independiente del contrafáctico. Pero es fácil ver que si existe esa función selectiva, también existe una función selectiva que no posee la característica de la independencia. De esta manera, la demostración citada utiliza dos contrafácticos $A_i > B_i$, que se convierten en verdaderos por la función selectiva f , que satisface que $A_i \& B_i \rightarrow A_j \& B_j$ ($i, j = 1, 2$). La demostración opera de manera semejante si tomamos una función selectiva f' , que se parece a f excepto en que para $A_i > B_i$ selecciona el conjunto de mundos que escoge f *más* todos los mundos A_i en los cuales todas las proposiciones pertenecientes a t_{B_i} , para las que el suceso A es

estas páginas, resulta crucial permitir que el conjunto $CIP(A, t_{\bar{B}})$ desempeñe un papel central en el análisis de los contrafácticos, y éste es el caso, en efecto, en las interpretaciones d.n. y p.c. Pero para tomar en cuenta seriamente a $CIP(A, t_{\bar{B}})$ se necesitaría una dependencia funcional respecto del contrafáctico $A > B$ al margen de una mera dependencia respecto de \mathcal{W}_A . En consecuencia, no parece que un ordenamiento de los mundos posibles pueda tomar en cuenta adecuadamente a este conjunto dentro de una teoría de la función selectiva que satisfaga la característica de la independencia o de una teoría de las esferas. Pero, según lo señalé, no tomar en cuenta a este conjunto sería una falta grave, para la cual no cabría la compensación de consideraciones de similitud fáctica global que no sean relativas con respecto a los rasgos dependientes del contrafáctico en cuestión.

Por lo tanto, me gustaría señalar que mi examen de los contrafácticos no posee ni un modelo equivalente a la teoría de las esferas ni un modelo equivalente a la función selectiva con la característica de la independencia. También me gustaría indicar que si mis argumentos concernientes a la centralidad de $CIP(A, t_{\bar{B}})$ son correctos, no es realista esperar, en vista de la característica de la independencia en la teoría de las esferas, que un modelo de esta última sea adecuado. Igualmente, no es realista esperar que sea adecuado un modelo de una teoría de la función selectiva que posea la característica de la independencia. Pero, desde luego, abandonar la característica de la independencia en una teoría de las esferas es abandonar lo que yace en el corazón del enfoque de Lewis, es decir, el cuadro de las esferas en torno al mundo en el que se considera al contrafáctico, que es uno y el mismo para todos los contrafácticos (en un contexto dado).

17. *Relevancia causal*

Hasta el momento, en nuestra exposición (especialmente en las secciones 3 y 4) no hemos brindado una explicación precisa sobre el tipo de nociones de relevancia causal y relevancia causal puramente positiva que son adecuadas para el análisis de los contrafácticos que presenté arriba. En esta sección, se detallan explicaciones más específicas sobre estas nociones. Con el fin de asumir una noción objetiva de los contrafácticos, la noción de probabilidad apropiada para nuestra discusión debe ser la de probabilidad objetiva. La noción que propongo es la de que la probabilidad de que un enunciado C

causalmente irrelevante, son verdaderas. Esta función no tendrá la característica de la independencia.

(perteneciente al tiempo t_C) sea verdadero con base en el mismo mundo en t (para un punto temporal $t, t > t_C$), se interprete como el *grado* en que el mundo hasta t determina la verdad del enunciado C . (Pensemos en ella en analogía, en un caso determinista, con la determinación por M_t de que el suceso C se produzca.) Esta noción de probabilidad se halla relacionada con la de una situación de azar sólo que se aplica a un caso singular y toma a toda la historia del mundo hasta cierto momento como una situación.⁷⁰ Dada esta noción de $P(C/M_t)(t < t_C)$, la probabilidad condicional puede ser definida de la forma acostumbrada:

$$(12) \quad P(C/A \& M_t) = \frac{P(C \& A/M_t)}{P(A/M_t)}$$

(para los casos en que A es compatible con M_t y donde $t_A > t$; esto se requiere para los contrafácticos del tipo d.n. y es necesario para definir los términos señalados y para que el denominador no sea 0).

Consideremos ahora, para un contrafáctico $A > B$ del tipo d.n. y para una proposición verdadera C perteneciente a (t_A, t_B) , si A es causalmente relevante para C . Una condición inicial que podría emplearse para caracterizar esta relación entre A y C (dada la historia del mundo) sería (con " M " como abreviatura de " M_A "):

$$(13) \quad P(C/A \& M) = P(C/ \sim A \& M)$$

Es decir, la condición (13) exige que la ocurrencia del suceso A , en cuanto opuesta a la ocurrencia del suceso $\sim A$, dado M , no represente ninguna diferencia probabilística para la ocurrencia del suceso C .

Pero esta condición es demasiado débil. Podría darse el caso de que dos cadenas causales conducentes a C se vean afectadas de manera distinta por la ocurrencia de A en cuanto opuesta a $\sim A$. Podría ser que la ocurrencia del suceso A intensificara una cadena causal pero suprimiera la otra en lo que concierne a sus efectos sobre C (en cuanto opuesta, desde luego, a la ocurrencia del suceso $\sim A$).

En tal caso, A sería causalmente relevante para C , pues su ocurrencia en cuanto opuesta a $\sim A$ filtraría de forma diferente las cadenas causales conducentes a C o $\sim C$ en lo que respecta a su impacto sobre C , y por lo tanto daría por resultado un trasfondo causal diferente para C . Pero bien podría ocurrir que el debilitamiento y el

⁷⁰ Para una mayor elaboración de esta noción de probabilidad, y para una tentativa de introducirla mediante la noción de probabilidades transicionales consecutivas entre estados del mundo para tiempo discreto, cf. [6], cap. 4, secciones I y VI.

fortalecimiento de estas dos cadenas causales por A y $\sim A$ respectivamente se equilibraran por accidente, de tal modo que se diera la condición (13). Por lo tanto, la condición (13) puede darse a pesar de la relevancia causal.

En consecuencia, hay que fortalecer la condición (13). Mi propuesta es hacerlo concentrándonos en sucesos intermedios del intervalo (t_A, t_C) . Estos sucesos, que sirven para transmitir a C los efectos de su trasfondo causal, manifestarían la relevancia causal de A para C . Los acontecimientos de ese intervalo que pueden manifestar la relevancia causal de A para C serían los que representan una diferencia para la ocurrencia de C en el curso de A o en el de $\sim A$; ellos pondrían de manifiesto esa relevancia causal si su propia ocurrencia o su impacto en C se viera afectado por la transición (de la ocurrencia de $\sim A$ a la ocurrencia de A). La ocurrencia de sucesos de este tipo indicaría así la relevancia causal de A para C . Para excluir la relevancia causal de esta índole debemos excluir la ocurrencia de esos sucesos. Tenemos que exigir que, para todo suceso e de (t_A, t_C) cuya ocurrencia representa una diferencia para la ocurrencia de C en el curso de A o en el de $\sim A$, la probabilidad de la ocurrencia de e mismo, así como el impacto de e en C , no deben verse afectados por la transición de $\sim A$ a A . Podemos formular esta última condición para la irrelevancia causal en términos probabilísticos de la siguiente manera:⁷¹

$\{M, A, L, C\}$ es consistente, y para todo suceso real e de (t_A, t_C) , si satisface a

$$(i) P(C/e \& \sim A \& M) \neq P(C/\sim A \& M)$$

o a

$$(ii) P(C/e \& A \& M) \neq P(C/A \& M)$$

entonces satisface tanto a:

$$(iii) P(e/A \& M) = P(e/\sim A \& M)$$

como a

$$(iv) P(C/e \& \sim A \& M) = P(C/e \& A \& M).^{72}$$

⁷¹ Para más detalles, cf. [6], cap. 4, secciones II – V. En este lugar, doy por supuesto que las probabilidades en consideración no son 0 ni 1.

⁷² En este tratamiento de la relevancia causal doy por supuesto que la relevancia causal que posee el suceso A para el suceso C es únicamente función de la historia previa M_A . En particular, en este análisis la historia del mundo perteneciente al intervalo (t_A, t_C) no desempeña ningún papel. Por lo tanto, este análisis refleja una noción de la relevancia causal del suceso A para el suceso C que no depende de la historia del mundo perteneciente a ese intervalo (t_A, t_C) .

Sin embargo, la noción de relevancia causal requerida para el análisis de los

18. *Relevancia causal puramente positiva*

Nos falta especificar mejor la noción de que el suceso A posee una relevancia causal puramente positiva para el suceso C . Podemos empezar con la plausible condición que sigue, la condición del aumento de probabilidades:

$$(14) \quad P(C/A \& M) > P(C/\sim A \& C)$$

Esta condición garantiza que A posee un efecto positivo global sobre C . Pero no llega a garantizar que este efecto sea *puramente* positivo, en cuanto opuesto a ser positivo globalmente, aunque posiblemente mixto. Para asegurar lo anterior, es mejor que recurramos otra vez a los sucesos intermedios del intervalo (t_A, t_C) , y que controlemos su papel en la transmisión de efectos causales a C . Para que A posea una relevancia causal puramente positiva para C , estos sucesos intermedios e deben satisfacer lo siguiente (además del cumplimiento de la condición del aumento de probabilidades): Si un suceso real e de (t_A, t_C) intensifica la transmisión de efectos positivos a C en el curso de $\sim A$, debe seguir haciéndolo en el curso de A , y su contribución positiva en este último no debe disminuir (en comparación con su contribución positiva en el curso de $\sim A$). Además, su probabilidad en el curso de A no debe ser menor que en el curso de $\sim A$. Si la ocurrencia de e no intensifica la transmisión de un efecto causal negativo a C en el curso de $\sim A$, entonces su ocurrencia no debe intensificar la transmisión de un efecto negativo a C tampoco en el curso de A .

Pero si la ocurrencia de e provoca la transmisión de efectos negativos a C en el curso de $\sim A$, entonces, si también tiene una contribución negativa en el curso de A , esta última no debe ser mayor que en el curso de $\sim A$. Y si la ocurrencia de e intensifica la transmisión de efectos negativos a C en ambos cursos, su probabilidad en el curso de A no debe ser mayor que en el curso de $\sim A$.

Esta condición puede formularse ahora en términos de probabilidades condicionales como sigue (con base en la condición del aumento de probabilidades (14)).⁷³

contrafácticos *sí* exige una dependencia respecto de la historia del mundo perteneciente al intervalo (t_A, t_C) , *además* de una dependencia respecto de M_A . La consideración exclusiva de M_A , pero no de la historia del mundo perteneciente a (t_A, t_C) , da por resultado un análisis que es una mera *aproximación* a la noción de relevancia causal requerida para el análisis de los contrafácticos. Para un examen más elaborado de la relevancia causal, que toma en cuenta la historia del mundo perteneciente a (t_A, t_C) y por lo tanto es más adecuado para el tratamiento de los contrafácticos, cf. [8].

⁷³ Como antes, doy por supuesto que las probabilidades no son 0 ni 1. Para más detalles, cf. [6], cap. 5.

$\{M, A, L, C\}$ es consistente, y todo suceso real e de (t_A, t_C) satisfice las siguientes condiciones:

A. (0) Si $P(C/e \sim A \& M) \geq P(C/\sim A \& M)$

entonces:

(1) $P(C/e \& A \& M) - P(C/A \& M) \geq P(C/e \& \sim A \& M) - P(C/\sim A \& M)$

En caso de estricta desigualdad en A (0):

(2) $P(e/A \& M) \geq P(e/\sim A \& M)$.

B. Si

(1) $P(C/e \& \sim A \& M) < P(C/\sim A \& M)$

y si

(2) $P(C/e \& A \& M) < P(C/A \& M)$

entonces:

(3) $P(C/A \& M) - P(C/e \& A \& M) \leq P(C/\sim A \& M) - P(C/e \& \sim A \& M)$

y

(4) $P(e/A \& M) \leq P(e/\sim A \& M)$.

TRADUCCIÓN DE ANA ISABEL STELLINO

REFERENCIAS

- [1] Chisholm, R., "The Contrary-to-fact Conditional", en Feigl y Sellars, *Readings in Philosophical Analysis*, Appleton-Century-Crofts Inc., 1949.
- [2] Goodman, Nelson, *Fact, Fiction, and Forecast*, Bobbs-Merrill, 1955; 2a. edición, 1965.
- [3] Downing, P. B., "Subjunctive Conditionals, Time Order and Causation", *Proceedings of the Aristotelian Society*, LIX, (1958/9): pp. 125-40.
- [4] Kvat, Igal, "Formal Semantics for Temporal Logic and Counterfactuals", *Logique et Analyse*, marzo de 1980, pp. 35-62.
- [5] Kvat, Igal, "Quine and Modalities *de re*: A Way Out?", *Journal of Philosophy*, LXXIX, 6 de junio de 1982, pp. 295-328.
- [6] Kvat, Igal, *A Theory of Counterfactuals*, Hackett Pub. Co., Indianápolis, 1986.
- [7] Kvat, Igal, "Possible Worlds and Causal Empiricism", *Canadian Journal of Philosophy* (pendiente).
- [8] Kvat, Igal, "Causal Independence" (pendiente).

- [9] Lewis, David, *Counterfactuals*, Harvard University Press, 1973.
- [10] Lewis, David, "Counterfactual Dependence and Time's Arrow", *Nous*, XIII, 4, 1979.
- [11] Lewis, David, "Ordering Semantics and Premise Semantics for Counterfactuals", *Journal of Philosophical Logic*, 10, 1981.
- [12] Pollock, John, *Subjunctive Reasoning*, Reidel, Dordrecht, 1976.
- [13] Stalnaker, Robert, "A Theory of Conditionals", en *Studies in Logical Theory*, N. Rescher (ed.), suplemento monográfico del *American Philosophical Quarterly*, 1968.
- [14] Stalnaker, Robert, "A Defense of Conditional Excluded Middle", en *Ifs*, W. L. Harper, R. Stalnaker y G. Pearce (eds.), Reidel, Dordrecht, 1981.