

EPISTEMOLOGÍA E HISTORIA DE LA CIENCIA

SELECCIÓN DE TRABAJOS DE LAS XVIII JORNADAS

VOLUMEN 14 (2008)

Horacio Faas
Hernán Severgnini

Editores



ÁREA LOGICO-EPISTEMOLÓGICA DE LA ESCUELA DE FILOSOFÍA
CENTRO DE INVESTIGACIONES DE LA FACULTAD DE FILOSOFÍA Y HUMANIDADES
UNIVERSIDAD NACIONAL DE CÓRDOBA



Esta obra está bajo una Licencia Creative Commons atribución NoComercial-SinDerivadas 2.5 Argentina



Contrafácticos y derrotabilidad: ¿es el condicional contrafáctico *might* un condicional derrotable?

María Inés Corbalán*

Condicionales contrafácticos y derrotabilidad (I parte)

Condicionales contrafácticos, invalidez de RA y validez de MP

De acuerdo con la semántica para condicionales contrafácticos, y en términos muy generales, un condicional contrafáctico $A > B$ es verdadero en un mundo w si y sólo si (ssi) B es verdadero en todos los A -mundos más cercanos al mundo w . De acuerdo con ello, algunas tesis válidas para el condicional material y para los condicionales estrictos resultan inválidas para el condicional contrafáctico. En particular, la tesis de Refuerzo del antecedente (RA): $(A \rightarrow B) \rightarrow ((A \wedge C) \rightarrow B)$, resulta inválida para el condicional contrafáctico. En otros términos, $(A > B) \rightarrow ((A \wedge C) > B)$ no es válida en un sistema contrafáctico.

Por otra parte, dada la validez de las tesis $(A \rightarrow B) \rightarrow (A > B)$ y $(A > B) \rightarrow (A \rightarrow B)$ en los diferentes sistemas condicionales, el condicional contrafáctico es caracterizado como un operador intermedio entre el condicional estricto y el condicional material. Es importante notar que esta última tesis es equivalente a la tesis *Modus ponens* condicional (MPC). $((A > B) \wedge A) \rightarrow B$ y con ello, que la regla de *Modus ponens* -de A y $A > B$ inferir B - permanece válida para el condicional contrafáctico.

Otra tesis clásica que permanece válida para el condicional contrafáctico es la tesis de Suficiencia de la conjunción (SC). $(A \wedge B) \rightarrow (A > B)$. Si bien esta tesis podría considerarse antiintuitiva por cuanto, como su nombre lo indica, afirma que la verdad de dos proposiciones cualesquiera A y B es condición suficiente para la aseveración de una vinculación condicional entre ellas, su validación en la semántica formal resulta de la aceptación de la condición de centramiento fuerte (FC) que parece intuitiva si la relación entre los mundos posibles pretende dar cuenta de la similaridad comparativa o cercanía entre dichos mundos. En una semántica fuertemente centrada no hay un mundo posible v que sea tan cercano o tan parecido al mundo w como el mismo mundo w o, en otros términos, cada mundo posible w es más parecido a sí mismo que cualquier otro mundo v . De este modo, si bien la tesis SC podría considerarse indeseable por encontrarse gran cantidad de ejemplos contraintuitivos, resulta validada por las propiedades que la relación de similaridad comparativa satisface intuitiva y formalmente. De este modo, los sistemas para condicionales contrafácticos, que dan cuenta de la comparación entre mundos apelando a la relación de similaridad o cercanía, validan la tesis SC.

Como se ha mencionado, la tesis RA resulta inválida para el condicional contrafáctico. Con ello, podría considerarse que la conexión entre antecedente y consecuente de un condicional contrafáctico es derrotable, esto es, que la vinculación condicional puede resultar rechazada como consecuencia de la adición de cierta información al antecedente. Podría considerarse por ello, que la aseveración de un contrafáctico se realiza atendiendo a ciertas excepciones que, de ser tenidas en cuenta, podrían hacer inaceptable la aseveración del consecuente. Pese a ello,

* UNLP

Makinson (1993: 363) observa que la conectiva contrafáctica constituye un condicional no derrotable y que la semántica propuesta por D. Lewis trata al condicional contrafáctico como un condicional "estricto", esto es, como un condicional que no admite ninguna excepción. Tal como nota Makinson, la verdad (no vacua) de un condicional contrafáctico ($A > B$) al igual que la verdad de un condicional necesario o estricto ($A \Rightarrow B$) requiere que el consecuente B sea verdadero en todos los A-mundos. Si bien la verdad del condicional $A > B$ no exige la verdad de B en todos los A-mundos accesibles a w, sino sólo la verdad de B en todos los A-mundos que son accesibles y más cercanos a w, ésta condición resulta, en opinión de Makinson, demasiado estricta. De allí que Makinson sostenga que el condicional contrafáctico no toma en consideración las posibles excepciones y, por tanto, constituya un condicional no derrotable.

Condicionales e inferencia derrotable

Las condiciones semánticas que, de acuerdo con Makinson, ha de satisfacer un condicional contrafáctico para ser derrotable resultan del estudio de la noción de inferencia no monótona. Mediante la ampliación del lenguaje a partir de la introducción de una conectiva condicional derrotable y la transformación de la cláusula para la evaluación de la regla $A | \sim B$ en una cláusula para la evaluación de la fórmula condicional $A \sim > B$, Makinson logra establecer las condiciones de verdad del condicional derrotable. Así, la regla:

$A | \sim B$ se cumple en el modelo de normalidad absoluta² $\langle W, <, \models \rangle$ ssi $v \models B$ para todo mundo $v \in W$ que es mínimo bajo la relación $<$ en el conjunto de todos los mundos que satisfacen A
se transforma en:

$w \models A \sim > B$ ssi $v \models B$ para todo mundo $v \in W$ que es mínimo bajo $<$ entre todos los mundos que satisfacen A.

El elemento fundamental de esta semántica es la relación de orden $<$. Mediante dicha relación se establece un ordenamiento de los elementos de W respecto del grado de normalidad. La relación de normalidad suele ser entendida como una relación irreflexiva y transitiva y $v < w$ es leída como 'v es más normal que w' o 'v es preferida a w'.³ El ordenamiento que se establece a partir de la relación de normalidad es absoluto, es decir, el orden entre los mundos no varía en función del mundo desde el cual una fórmula es evaluada. En ambos casos, tanto si se evalúa el valor de verdad de la fórmula $A \sim > B$ en un mundo o el valor de verdad de la aserción $A | \sim B$ en un modelo, la evaluación de B se realiza atendiendo a los mundos que son mínimos de acuerdo al orden de normalidad establecido mediante la relación $<$.

Así estipuladas, la condición de verdad para el condicional derrotable y para el condicional contrafáctico resultan diferentes por cuanto requieren de ordenamientos diferentes de los mundos. El ordenamiento entre los mundos en la semántica adecuada al condicional derrotable es independiente del mundo desde el cual se realiza la evaluación del condicional. El ordenamiento de mundos en la semántica para el condicional contrafáctico varía en función del mundo en el cual la evaluación toma lugar. Así, en los modelos de normalidad variable, la relación de orden $<$ es sustituida, para cada mundo w, por la relación de orden $<_w$ de modo que:

$w \models A > B$ ssi $v \models B$ para todo mundo $v \in W$ que es mínimo bajo $<_w$ entre todos los mundos (que son accesibles desde w bajo la relación R y) que satisfacen A.

Modus ponens condicional

Resulta claro que esta diferencia entre la semántica de normalidad absoluta, adecuada a los condicionales derrotables y la semántica de normalidad variable, adecuada a los condicionales contrafácticos redundaría en la consideración de la invalidez de ciertas tesis. En particular, la tesis de *Modus ponens*, válida en una semántica de normalidad variable cuya relación de cercanía está débilmente centrada⁴, resulta inválida desde una semántica de normalidad absoluta.⁵

Se ha mencionado que Makinson considera que el análisis de los condicionales contrafácticos propuesto por D. Lewis, si bien invalida la tesis de RA o monotonía no logra dar cuenta de un condicional contrafáctico derrotable. La falla de *Modus ponens* y la falla de monotonía o refuerzo del antecedente constituyen a decir de Makinson, “manifestaciones de la derrotabilidad”. La caracterización de un condicional contrafáctico derrotable exige por ello, tanto la invalidez de la tesis de monotonía o RA como la falla de la regla de *Modus ponens*. Dado que el condicional contrafáctico caracterizado por D. Lewis no falla en la regla de *Modus ponens* inválida para el condicional derrotable, no constituye por ello, un condicional derrotable.⁶

A fin de obtener las condiciones de verdad de un condicional contrafáctico y derrotable, Makinson modifica la semántica de los condicionales contrafácticos introduciendo el ordenamiento entre mundos propio de la semántica para los condicionales derrotables. Así, la semántica adecuada a los condicionales contrafácticos derrotables presenta dos ordenamientos independientes de los mundos: por un lado, el ordenamiento característico de las semánticas para condicionales contrafácticos, esto es, un ordenamiento que atiende al grado de similaridad y que es relativo al mundo desde el cual un cierto condicional es evaluado. Por el otro, el ordenamiento característico de las semánticas para condicionales derrotables, esto es, un ordenamiento que atiende al grado de normalidad en términos absolutos. De acuerdo a ello, la condición de verdad del condicional contrafáctico derrotable es:

$w \models A \approx B$ ssi $v \models B$ para todo $v \in M$ que es el más cercano a w entre los mundos más normales que satisfacen A .⁷

La intuición contra *Modus ponens*

Tanto la falla de RA como la falla de MP expresan que las situaciones antecedentes consideradas por el hablante para la afirmación de un condicional contrafáctico no son todas las situaciones posibles. La falla de RA es clara en ese sentido: la adición de cierta información al antecedente del condicional podría ocasionar el rechazo del consecuente. Así, quien afirmara el contrafáctico *Si hubiera tomado un vaso de agua, habría saciado mi sed*, no aceptaría dicho condicional si supiera que el vaso contenía sal. La situación posible en la cual el vaso contiene, no sólo agua sino también sal no es una de las situaciones posibles consideradas por el hablante. De acuerdo con la semántica formal para los condicionales contrafácticos, las posibles situaciones consideradas por el hablante para la preferencia de un condicional son las más parecidas a la situación actual y, en particular, la propia situación actual.

Si bien esta selección parece dar cuenta del carácter no necesario del contrafáctico, Makinson propone que imaginemos la siguiente situación. Un hablante indica que *Si Mozart hubiera vivido hasta el año 1800, habría escrito más sinfonías*. Dicho hablante se entera luego de que Mozart de hecho vivió hasta pasado el año 1800 y no murió en el año 1791 como había creído. Sin embargo, Mozart había falsificado los documentos de su defunción para escapar a sus

acreedores y había viajado al sur de Argentina a buscar oro. Así, si bien es cierto que Mozart vivió hasta el año 1800, es falso que escribió más sinfonías. Sin embargo, es claro que dicha situación, aunque real, es tan extraña que no habría sido considerada por el hablante. De acuerdo con Makinson, quien profiere el contrafáctico *Si Mozart hubiera vivido hasta el año 1800, habría escrito más sinfonías* considera que Mozart habría escrito más sinfonías si, bajo el supuesto de que los acontecimientos en la vida del compositor siguieron el curso normal, Mozart hubiera vivido hasta el año 1800. Queda claro a partir del ejemplo que ciertas situaciones, si bien reales, por ser tan atípicas o excepcionales, no son tenidas en consideración por quien formula dicho condicional.

Lo interesante de considerar a los condicionales contrafácticos como condicionales derrotables es que la existencia de situaciones atípicas o excepcionales no convierte al contrafáctico *Si Mozart hubiera vivido hasta el año 1800, habría escrito más sinfonías* en un enunciado falso. Al igual que en el caso de la inferencia derrotable, en el caso de los condicionales contrafácticos se considera que la vinculación entre antecedente y consecuente puede tener excepciones, que su aserción se realiza bajo el supuesto de normalidad.

Adoptando la propuesta de la inferencia derrotable para los condicionales contrafácticos podríamos continuar considerando verdadero el contrafáctico aun cuando en el mundo actual el antecedente resultara verdadero y el consecuente falso si la situación actual es atípica o anormal. Por ello, en opinión de Makinson, debe rechazarse, para los condicionales contrafácticos derrotables, esto es, para los condicionales contrafácticos que expresan que la vinculación entre antecedente y consecuente se da en condiciones normales, el esquema de *Modus ponens*. Por el contrario, quien sostuviera la validez de MP y hubiera proferido el condicional contrafáctico *Si Mozart hubiera vivido hasta el año 1800, habría escrito más sinfonías* debería retractarse si, posteriormente averigua que Mozart falsificó los documentos que probaban su defunción y no murió en 1791 sino en el año 1810 pero no escribió ni una traza de música. Por el contrario, de acuerdo con Makinson, el rechazo del esquema MP permite que el hablante considere el condicional contrafáctico *Si Mozart hubiera vivido hasta el año 1800, habría escrito más sinfonías* verdadero aun si se muestra que efectivamente Mozart vivió más allá de 1791 pero no escribió ninguna sinfonía a causa de su nueva y atípica vida en el sur argentino. Esto es así porque si la aserción condicional se realiza bajo el supuesto de normalidad, la verdad del antecedente no es condición suficiente para la verdad del consecuente, aun cuando la situación en la cual el antecedente es verdadero sea la misma situación actual. De este modo, la invalidez de MP indica que no cualquier situación que haga verdadero al antecedente del condicional es la clase de situación considerada en la afirmación del condicional.

De este modo, la expresión de la derrotabilidad del condicional no depende sólo de la falla de RA sino también de la invalidez de MP. Mientras que la falla de RA indica que no todas las situaciones posibles son consideradas por quien profiere un condicional, la falla de MP expresa que la situación actual podría ser una de tales situaciones no consideradas. En consecuencia, y en concordancia con la condición de verdad anteriormente indicada, las situaciones posibles que el hablante considera cuando profiere un condicional contrafáctico derrotable son, a la vez que similares a la situación real, normales o típicas.

Condicionales contrafácticos y derrotabilidad (II parte)

Condicional *might* y las tesis RA, MP y SC

El sistema de D. Lewis cuenta con un segundo operador contrafáctico, el condicional *might* $A\Diamond\rightarrow B$, que es leído como *Si se hubiera dado A, se podría haber dado B*. D. Lewis caracteriza este condicional a partir de la negación del condicional *would*. Considerando que $(A\Diamond\rightarrow B) = \text{df. } \neg(A\Box\rightarrow\neg B)$, las condiciones de verdad para el operador *might* quedan establecidas:

$w \models A\Diamond\rightarrow B$ ssi hay A-mundos accesibles y para todo A- \neg B-mundo accesible, hay algún AB-mundo accesible al menos tan cercano a w.

Dado que el operador *might* es más débil que el operador *would* y que la verdad de $A\Diamond\rightarrow B$ en w no exige la verdad del consecuente en todos los A-mundos más cercanos a w, en lo que sigue analizaré la posibilidad de considerar al condicional *might* un condicional contrafáctico derrotable.

Ya hemos visto que, de acuerdo con Makinson, la semántica apropiada para dar cuenta de un condicional contrafáctico derrotable invalida, en tanto condicional contrafáctico, la tesis de RA, y, en tanto condicional derrotable, la tesis MP.

Respecto de la tesis de monotonía o RA, ya se ha indicado que todo sistema para condicionales contrafácticos la rechaza. Asimismo, la contraparte *might* de dicha tesis, *i.e.* $(A\Diamond\rightarrow C)\rightarrow((A\wedge B)\Diamond\rightarrow C)$, resulta inválida. Ello es así dado que la verdad de $A\Diamond\rightarrow C$ en w sólo exige que existan A-mundos y que existan AC-mundos más cercanos a w que todos los A- \neg C-mundos y ello no implica ni que haya AB-mundos ni que, de haber AB-mundos, los ABC-mundos sean más cercanos a w que los AB- \neg C-mundos. Así, RA falla también para el operador *might*.

De acuerdo a las condiciones de verdad del operador *might*, la tesis MP para el condicional *might* resulta válida en modelos fuertemente centrados pero inválida en modelos débilmente centrados.⁸ Si bien D. Lewis reconoce que la noción de similitud o cercanía debe satisfacer alguna condición de centramiento, considera que la condición FC puede descartarse manteniendo sólo WC. Luego, si se mantiene WC pero se descarta la condición de centramiento fuerte, la tesis MP se invalida para el operador *might*.

Así, dado que es posible obtener una semántica dentro de los límites aceptados por el propio D. Lewis tal que resultan inválidas las tesis RA y MP, parecería que el operador *might* constituye un condicional contrafáctico derrotable.

Sin embargo, es interesante notar que la propuesta semántica de Makinson para los condicionales contrafácticos derrotables al seguir de cerca la semántica para la inferencia derrotable invalida, no sólo la tesis de RA sino asimismo SC: $(A\wedge B)\rightarrow(A>B)$. Luego, a fin de poder afirmar que el condicional *might* constituye un condicional derrotable, parecería que debe analizarse la validez de dicha tesis para el operador *might*.

Si la semántica está fuertemente centrada, entonces, si $w \models (A\wedge B)$ y $w \models A$ entonces, $w \models B$. Luego si se acepta FC, $((A\wedge B)\rightarrow(A\Diamond\rightarrow B))$ resulta una tesis válida. Por otro lado, en una semántica con un orden entre mundos sólo débilmente centrado, no resulta válida la tesis SC para el operador *would* pero su contraparte *might* permanece válida.⁹ Simplemente, si $w \models (A\wedge B)$, aún cuando no se de que $v \models B$ para todo v que es tan cercano a w como w mismo $w \models A\Diamond\rightarrow B$. Por ello, a los fines de invalidar esta tesis para el operador *might* deberíamos modificar el

ordenamiento de mundos modificando la condición de centramiento, eliminando en este caso, la condición WC.

Sin embargo, suele reconocerse que en una semántica que pretenda dar cuenta de los condicionales contrafácticos, donde la relación de orden sea entendida como una relación de similitud o cercanía, alguna condición de centramiento -débil o fuerte- debe ser adoptada. Así, WC es la mínima condición exigible y por ello, no puede ser eliminada.¹⁰

En consecuencia, si bien es posible invalidar la tesis MP para el operador *might* mediante la modificación de la semántica para los condicionales contrafácticos dentro de los límites aceptables, con ello no logra invalidarse SC para el operador *might*.

Conclusiones

Se ha mostrado que la tesis RA resulta inválida para el operador *might* independientemente la condición de centramiento adoptada y que la invalidez de la tesis MP resulta de la eliminación de la condición FC. Asimismo, se ha notado que las modificaciones en la semántica necesarias para invalidar SC requieren rechazar la mínima condición de centramiento exigida para que la relación entre mundos sea entendida como una relación de cercanía o similitud. De modo que una semántica cuyos mundos se ordenan a partir de la relación de similitud y, por ende, está débilmente centrada, valida SC para el contrafáctico *might*. De este modo, si, de acuerdo con Makinson, un contrafáctico derrotable, al igual que la inferencia derrotable debe invalidar RA, MP y SC, el contrafáctico *might*, aun siendo más débil que el condicional *would* no constituye un condicional derrotable.

Notas

¹ Sólo cuando sea necesario distinguir los dos operadores contrafácticos de Lewis -- *might* y *would*, se utilizará el signo $\sqsubset \rightarrow$ para éste último.

² Un modelo de normalidad absoluta es un tripleto $\langle W, <, \vDash \rangle$ donde W es un conjunto de ítems arbitrarios, comúnmente denominados "mundos", $<$ es una relación binaria sobre W y \vDash es una relación de satisfacción entre mundos y proposiciones del lenguaje, que se comporta veritativo funcionalmente para las conectivas clásicas y que satisface la condición del límite. Makinson, 1993. 345s.

³ La relación $<$ establece así un orden parcial.

⁴ Así como la validez de SC se corresponde con la condición FC, la validez de MP se corresponde con la condición WC. ningún mundo es más cercano a w que w mismo, i.e. no se tiene que $v <_w w$

⁵ Para un contraejemplo en un modelo de normalidad absoluta, elijase un mundo w tal que $w \vDash A$ y $w \vDash A \rightarrow B$ pero que no sea mínimo bajo la relación $<$. Así, aún cuando para todo $v \in W$ que sea mínimo bajo $<$, $v \vDash B$, no se da que $w \vDash B$

⁶ De acuerdo con las semánticas para contrafácticos propuesta por Lewis y Stalnaker, si el condicional contrafáctico $A \rightarrow B$ y el antecedente A resultan verdaderos en el mundo actual, entonces, B es verdadero en todos los A -mundos más cercanos al mundo actual, lo cual incluye, por WC o por FC, al mismo mundo actual. Es decir, si $w \vDash A \rightarrow B$ y $w \vDash A$, entonces $v \vDash B$ para todo v más cercano a w . Luego, $w \vDash B$

⁷ Dado que el ordenamiento que responde al grado de similitud puede estar fuerte o débilmente centrado Makinson propone alternativamente dos condiciones de verdad para un contrafáctico derrotable. Aquí sólo se indica la condición que no depende de la condición de centramiento.

⁸ Demostración. si $w \vDash A \rightarrow B$ y $w \vDash A$ entonces por FC, para cualesquiera $v, w \in W$ tal que $v \neq w, w <_w v, w \vDash B$. Por el contrario, si el modelo está débilmente centrado, entonces, si $w \vDash A$ y $v \vDash B$ y (wRv) y v es tan cercano a w como w mismo, entonces $w \vDash A \rightarrow B$ aún cuando no se de que $w \vDash B$

⁹ El sistema VC de Lewis contiene como tesis a SC y, por ende, la condición de centramiento es fuerte. El sistema VW, por su parte, se obtiene eliminando dicha tesis y su semántica requiere el debilitamiento de la condición de centramiento.

¹⁰ Así, por ejemplo Nute, aunque pretende rechazar la tesis SC, indica: "Surely, the actual world is more similar to itself than is any other world; otherwise, the notion of the similarity of worlds would be incomprehensible () We may try to

avoid CS [SC], but this amounts to accepting the questionable thesis that two distinct worlds might be as similar to each other as either is to itself" (1980:69s).

Referencias

Lewis, D. (1973) *Counterfactuals*, Oxford, Blackwell

Makinson, D. (1993) "Five faces of minimality" en *Studia Logica* 52, 339 – 379

Nute, D. (1980) *Topics in conditional logic*, Dordrecht, Reidel.