NÚMERO 15

MARÍA INÉS PAZOS

Consecuencia lógica derrotable: análisis de un concepto de consecuencia falible

JUNIO 2006

CIDE

www.cide.edu

Las colecciones de **Documentos de Trabajo del CIDE** representan un medio para difundir los avances de la labor de investigación, y para permitir que los autores reciban comentarios antes de su publicación definitiva. Se agradecerá que los comentarios se hagan llegar directamente al (los) autor(es).

Producción a cargo del (los) autor(es), por lo que tanto el contenido así como el estilo y la redacción son su responsabilidad.

Agradecimientos

Agradezco a David Gaytán y a los integrantes de la División de Estudios Jurídicos del CIDE las valiosas observaciones que hicieron a versiones previas de este artículo.

Resumen

En este trabajo analizo una noción de consecuencia lógica falible que represento en el metalenguaje con el signo "|~" y a la que llamo "consecuencia derrotable". La noción que reviso es la que creo adecuada para dar cuenta de los razonamientos que operan con "normas derrotables" (ciertos condicionales no monotónicos que expreso en el lenguaje objeto con la forma "A>B"). Sostengo que para dar cuenta de las inferencias basadas en este tipo de normas debe apelarse a una noción de consecuencia lógica no deductiva que tiene analogías con la noción de condicional derrotable, pero también diferencias importantes. Ambos se asemejan en que encubren una forma de implicación clásica, monotónica, usada para definirlos junto con una función que llamo "función de selección". Sostengo que difieren en que la noción de consecuencia derrotable cumple con importantes reglas lógicas como la de monotonía cauta y la de oposición de inferencias, que los condicionales derrotables no satisfacen. Finalmente, rechazo para la relación de consecuencia derrotable un principio importante del que carece la lógica de los condicionales derrotables pero que es usualmente aceptado para la inferencia derrotable, el principio de corte cauto. Así, acepto una noción de consecuencia derrotable menos permisiva que otras estándar y a la vez más segura.

Abstract

In this paper I analyze one notion of fallible logical consequence represented in the meta-language with the sign "|~" and that I call "defeasible consequence"This notion is the one that I believe adequate for the reasoning that operates with "defeasible norms" (certain non monotonic conditionals that I express in the object language with the form "A>B"). The notion of defeasible consequence has analogies with that of defeasible conditional, but also important differences. Both conceal a form of classical, monotonic implication, and both can be defined by using a function that I call "choice function". They differ in that the notion of defeasible consequence satisfies important logical rules like cautious monotony and opposition of inferences, which are invalid for defeasible conditionals. I reject for the relation of defeasible inference an important principle that is invalid for the logic of defeasible conditionals, but that is usually accepted for defeasible inference, the principle of cautious cut. Thus, I accept notion of defeasible consequence that is less permissive and more secure than other usual notions of fallible inference.

Introducción

En un artículo anterior (Pazos, 2002) defendí una noción de condicional derrotable como instrumento apto para dar cuenta de enunciados generales con excepciones, en particular de normas jurídicas con excepciones dependientes del contexto. Se trataba de un condicional para el cual no se satisfacían las leyes de Modus Ponens[1] y Refuerzo del antecedente[2]. En ese trabajo avancé una presentación formal para dar cuenta de su uso en razonamientos deductivos. Sostuve también que el uso más interesante de tales condicionales era como premisas de razonamientos no deductivos, específicamente razonamientos no monotónicos. Tal uso requería una noción de consecuencia lógica no deductiva que no presenté de manera formal. En este artículo me aproximaré a una noción de consecuencia lógica falible asociada a condicionales derrotables que llamaré "consecuencia lógica derrotable"; intentaré precisar su significado de manera no formal y analizaré algunas condiciones formales que debe satisfacer contrastándolas con principios análogos para los condicionales derrotables. En particular sostendré que a pesar de las fuertes analogías entre la lógica de los condicionales derrotables y la de la consecuencia derrotable, la última es más fuerte, valiendo para ella las reglas de oposición de razonamientos y monotonía cauta, aunque no tan fuerte como para validar principios importantes que usualmente se le atribuyen, como el principio de corte cauto.

Nota sobre el contexto teórico: Este trabajo se inscribe en el marco de un conjunto de investigaciones acerca de lógicas no monotónicas desarrolladas en los últimos 15 años, principalmente para dar cuenta de la representación del conocimiento y del razonamiento en el lenguaje ordinario. Para un estado del arte puede verse Morado (2004), que contiene una bibliografía básica y otra clásica en lógica no monotónica, y también puede consultarse a Prakken y Vreeswijk (2002). Las lógicas no monotónicas son actualmente uno de los tópicos más estudiados tanto en lógica como en inteligencia artificial. Ellas han sido aplicadas también al derecho por ejemplo en Prakken, (1997) y Prakken y Sartor, (1997).

1.- El sistema CD (condicionales derrotables)

Lenguaje: el de la lógica modal alética más el signo ">" (condicional

[1] $((p \supset q) \& p) \supset q$

[2] $(p \supset r) \supset ((p \& q) \supset r)$

derrotable) y una función binaria $f(x,y)$ usada para definir condicionales derrotables.

Un condicional derrotable p>q se define del siguiente modo

(Df>) $p>q =_{Df} f(p,q) \rightarrow q$

(donde "$\rightarrow$ " representa un condicional estricto)

Lo anterior significa que p implica derrotablemente q cuando el resultado de una función de selección de presupuestos (resultado que implica lógicamente el antecedente de p>q, esto es, $f(p,q)\rightarrow p$) implica estrictamente q, o más sencillamente, que p junto con un grupo de presupuestos implica estrictamente q. Los presupuestos consisten básicamente en una afirmación de normalidad. "p>q" puede leerse como "Si p normalmente q".

La lógica[3] para este lenguaje incluye la del sistema de lógica modal T más, entre otros, los axiomas siguientes:

Expansión $\vdash f(p,q)\_ \ p$

Expansión límite $\vdash \Diamond p \supset \Diamond f(p,q)$

En esta lógica son válidos los siguientes principios lógicos, entre otros, para los condicionales derrotables.

Superclasicalidad $p\rightarrow q \vdash p>r$

Oposición de condicionales restringida
$\Diamond[f(p,q) \& f(p, \sim q)$ entonces $(p>q) \supset \sim(p> \sim q)$

Monotonía cauta restringida $p\rightarrow q, p>r \vdash (p\&q)>r$

Corte cauto restringido $(p\&q)>r, p\rightarrow q \vdash p>r$

No son válidos, en cambio, los principios de

Oposición de condicionales $(p>q) \supset \sim(p> \sim q)$

Monotonía cauta $p>q, p>r \vdash (p\&q)>r$

Corte cauto $(p\&q)>r, p>q \vdash p>r$

[3] El sistema que propuse (Pazos 2002) y que reproduzco parcialmente aquí es una modificación del de Alchourrón (Alchourrón, 1996 y 1991).

A continuación presentaré algunas ideas acerca de la noción de inferencia falible asociada a los condicionales derrotables, a la que llamaré "consecuencia lógica derrotable", intentaré aproximarme informalmente a su contenido y reconoceré algunas analogías con la lógica de los condicionales derrotables, tales como las semejanzas en la definición y la validez de Superclasicalidad. Finalmente analizaré si deben valer para ella reglas análogas a los principios de oposición de condicionales, monotonía cauta y corte cauto. A estas reglas las llamaré, respectivamente, de monotonía cauta y de oposición de razonamientos y regla de corte cauto.

2.- Consecuencia falible, primera aproximación

Los condicionales derrotables aparecen en razonamientos como el siguiente:

Los menores de edad son (normalmente) incapaces	$p>q$
José es menor de edad	p
- -	- - - - -
Luego José es incapaz.	q

La conclusión no se sigue deductivamente de las premisas, sin embargo se sigue derrotablemente (lo que represento mediante la línea punteada), esto es, en condiciones de normalidad, aunque premisas adicionales acerca de casos excepcionales (como "José está casado") podrían hacernos retractar de la conclusión.

Esto muestra que necesitamos un concepto de consecuencia falible que dé cuenta de razonamientos como el anterior. Lo mismo ocurre con otras formas de razonamientos falibles que contienen condicionales derrotables como premisas.

Por otra parte, el razonamiento

Los menores son incapaces	$p>q$
- -	- - - - - - - - -
Luego, los menores casados son incapaces.	$(p\&r)>q$

es claramente incorrecto. Esto muestra la necesidad de analizar el comportamiento de esta noción para determinar qué reglas satisface y cuáles no.

La regla del Modus Ponens parece básica para el funcionamiento de cualquier condicional: la función básica de un condicional en un razonamiento parece ser que posibilite extraer el consecuente bajo ciertas condiciones. Es en ejemplos asociados a esta función donde mejor se ve, en una primera aproximación al concepto que analizamos, tanto la necesidad de rechazar el Modus Ponens en su forma tradicional (la regla deductiva conocida), como la de encontrar reglas no deductivas que den cuenta de la corrección de ciertos razonamientos.

Así, es obvio que la conectiva de condicional derrotable no debe satisfacer la regla del Modus Ponens tradicional:

$$p>q,\ p \vdash q$$

Si lo hiciera, eso significaría que la conclusión es irretractable y así, que el condicional no es genuinamente derrotable sino un condicional clásico.

Sin embargo es altamente contraintuitivo el rechazo del Modus Ponens para una construcción condicional, si se entiende esto como la imposibilidad de inferir el consecuente. Lo que de hecho hacemos ante condicionales derrotables es inferir el consecuente pero no deductivamente sino sólo faliblemente (a falta de información que nos pueda hacer retractar de la conclusión). La conclusión no será segura. Es importante notar que esta inseguridad no se deberá a una falta de confianza en las premisas (podemos estar absolutamente convencidos de la verdad de un condicional derrotable), sino a la falta de necesidad en la relación de inferencia.

Así la regla (deductiva) del Modus Ponens debe ser distinguida de la regla (no deductiva) siguiente:

Modus Ponens Derrotable

$$p>q,\ p \mid\!\sim q$$

donde "$\mid\!\sim$" representa la consecuencia lógica falible, que entenderemos como expresión de que q es "consecuencia derrotable" de las premisas. Esta regla constituye la forma adecuada para representar el razonamiento intuitivamente válido citado.

Debe observarse que la derrotabilidad de razonamientos, asociada a una noción de consecuencia que hemos llamado "derrotable" no representa el mismo concepto que el condicional derrotable del lenguaje objeto. El último, que expresa en el lenguaje objeto un vínculo entre hechos, significa que del antecedente del condicional junto con ciertos presupuestos se sigue estrictamente el consecuente. La relación de inferencia derrotable es un vínculo metalingüístico cuyo significado aún no hemos precisado. No sabemos qué razonamientos valida esta relación lógica además de los que responden a la regla transcripta arriba, esto depende de las propiedades que tenga esta nueva noción que debemos elucidar. Por el momento sólo hemos establecido una condición mínima positiva que ella debe satisfacer: la validez de una forma debilitada de Modus Ponens bautizada aquí "Modus Ponens Derrotable".

3.- Consecuencia lógica derrotable, su significado

No he caracterizado hasta ahora a los razonamientos derrotables y a la noción de consecuencia derrotable asociada a ellos. Sólo mostré que existen ejemplos de ellos tales como cualquier instancia de la regla del Modus Ponens Derrotable. Pero necesitamos una caracterización general de este tipo de razonamientos y para ello debemos precisar el concepto.

¿Qué significa "extraer derrotablemente la conclusión"? ¿Cuál es el vínculo entre premisas y conclusión de una inferencia derrotable válida? ¿Es similar a la relación entre antecedente y consecuente de un condicional derrotable? ¿Qué es lo que justifica, desde el punto de vista metalógico, la obtención de la conclusión en una inferencia derrotable?

Las preguntas anteriores, por una parte acerca del significado y por otra parte acerca de la justificación de las inferencias derrotables están íntimamente vinculadas. Sólo podremos identificar como inferencia derrotable correcta a una forma de razonamiento que consideremos justificada desde el punto de vista metalógico. Por eso, no es posible elucidar el significado de la expresión "inferencia derrotable" sin contar con una teoría de su justificación. Propondré aquí una manera de entender esta relación de inferencia derrotable que la justifica y a la vez proporciona algunas pautas para la identificación de razonamientos derrotables válidos y de las propiedades de la noción de inferencia involucrada. No es, sin embargo, una teoría completa. Una elucidación precisa de esta noción requeriría dar un sistema formal, trabajo que realizaré en este artículo.

Lo primero que debe quedar claro es que un razonamiento derrotable no es uno que arroje una conclusión derrotable. Lo que es derrotable es la relación de inferencia. En una inferencia de este tipo nos es permitido extraer la conclusión de manera derrotable, no deductiva (lo que no excluye la posibilidad de que también pueda extraérsela deductivamente), pero la conclusión misma no tiene por qué ser un enunciado derrotable (aunque podría serlo). La conclusión es al menos insegura, lo que significa que información adicional podría obligarnos a retractarla, pero ella, en principio, podría tener cualquier forma lógica. Por el otro lado, una conclusión que fuese un enunciado derrotable podría seguirse de un razonamiento deductivo y, así, tener ella misma certeza derivada de una confianza plena en las premisas. En ese caso la conclusión sería derrotable sin que lo fuese la relación de inferencia.

Una vez centrada la atención en la relación de consecuencia, debemos analizar qué propiedades (reglas) satisface.

Modus Ponens Derrotable

Imaginemos una inferencia cualquiera de la forma del Modus Ponens Derrotable:

$p>q, p|\sim q$

p>q significa que de p más un conjunto de presupuestos del condicional se sigue estrictamente q. La segunda premisa no afirma que se den todas las condiciones que serían suficientes para q, a saber, p en conjunción con los presupuestos citados (f(p,q)), sino que sólo garantiza la ocurrencia de una parte de tales condiciones: p. Justamente por el hecho de que al inferir q no estamos seguros de que f(p,q) sea verdad, es que no sabemos con certeza que q lo sea, por eso, el razonamiento es deductivamente inválido: la conclusión q no está asegurada por las premisas. El punto que debemos tratar aquí es si el argumento es correcto en algún sentido no deductivo, esto es, si estamos justificados a creer q sobre la base de p>q y p, con alguna confianza menor a la certeza. En otras palabras, sabemos que el razonamiento es deductivamente inválido y queremos mostrar que es no deductivamente correcto.

Replanteemos la pregunta sobre la corrección del Modus Ponens Derrotable. Sabemos que un condicional derrotable encubre un condicional estricto del cual el antecedente de su expresión derrotable es un conjunto. Ahora bien ¿Tenemos alguna razón para confiar en el consecuente de un condicional estricto del que sólo sabemos que ocurre parte del antecedente? Mi propuesta es que lo que sabemos (p en nuestro ejemplo) constituye evidencia de que ocurre una condición H [aquí H=f(p,q)] suficiente para el consecuente. Sostengo además que se trata de una evidencia suficiente para justificar una creencia confiable en H y por tanto también en su consecuencia necesaria, el consecuente q, conclusión del razonamiento derrotable.

El punto en que basaré mi justificación es la afirmación de que en un enunciado derrotable p>q, el antecedente p no es cualquier enunciado que se derive estrictamente de f(p,q) sino uno que constituye un buen indicio de f(p,q). Así, una norma derrotable expresa en su antecedente un hecho especialmente relevante para la obtención del consecuente, uno que crea expectativas razonables de la verdad de f(p,q). Así, si el legislador dispone "los menores son incapaces", la minoridad será un buen indicio de que alguien es, digamos, un menor no emancipado, porque normalmente los menores no están emancipados. Si una ley sobre incapacidad reza "A los mayores de edad se les nombrará un curador", esta norma será entendida en un contexto tal que la mayoría de edad será indicio de incapacidad. El contexto está asociado a la función f (que tiene como argumentos al antecedente y consecuente del condicional derrotable en que consiste una norma derrotable) y por eso no puede variar para distintas instancias de aplicación de una norma: cada norma tiene un único contexto.

La afirmación del contexto único es muy fuerte y requeriría una discusión amplia que no puedo completar aquí, de modo que sólo haré algunas aclaraciones. El contexto al que me refiero es un contexto normativo, esto es, un cuerpo de normas al que pertenece la norma derrotable considerada. El ordenamiento al que pertenece una norma, aunque a lo largo del tiempo puede cambiar en parte de su contenido, es un contexto fijo para las aplicaciones individuales una vez determinados el ordenamiento aplicable y su contenido. Esto es, para cada ordenamiento, en cada momento en que sus elementos (normas) están fijos, el contexto de una norma es el conjunto de disposiciones vinculadas semánticamente a ella, sea porque tratan la misma materia o porque de algún modo completan su sentido y, en suma, ayudan a determinar su significado. Los contextos a los que me refiero no son los contextos de aplicación, las situaciones individuales que son diferentes para cada caso; las circunstancias individuales a las que se aplica una norma cambian, pero no el significado de la norma. Estas circunstancias que varían no son identificadas por la función de selección, que representa una parte del significado de la norma.

Dado este supuesto, la justificación de una inferencia derrotable procede del siguiente modo.

Sabemos que para cualquier hipótesis $H=p_1\&p_2\&...\&p_n$ cualquier consecuencia lógica de H es evidencia para H en mayor o menor medida. Ahora bien, para cualquier inferencia derrotable p>q, p |~ q, la segunda premisa p es consecuencia lógica de un enunciado $H=f(p,q)$, de modo que p, el antecedente del condicional derrotable es evidencia para el antecedente del condicional estricto encubierto $f(p,q)$_q. En la medida en que p apoya $f(p,q)$, también apoya sus consecuencias lógicas, en particular q. De este modo, el antecedente de un condicional derrotable proporciona evidencia para su consecuente.

Por supuesto que una consecuencia lógica de una hipótesis cualquiera H puede no ser una evidencia fuerte para H (aunque siempre será evidencia). Pero dado el supuesto de que el antecedente expreso de un condicional derrotable es un indicio especialmente relevante de su conjunción con los presupuestos del enunciado, porque la conjunción expresa condiciones de normalidad, se sigue que el antecedente de un condicional derrotable es siempre una evidencia especialmente fuerte para el antecedente del condicional estricto encubierto y así, proporciona buenas razones para confiar en la conclusión.

Es importante notar que la realización de un razonamiento derrotable, no involucra un compromiso con los presupuestos. Ellos pueden ser falsos aunque haya evidencia en su favor. Esto diferencia a las inferencias derrotables de los razonamientos entimemáticos donde las premisas implícitas tienen la misma jerarquía que las premisas explícitas, se confía en todas ellas por igual. Por otra parte, en un entimema las premisas tácitas son constituyentes del razonamiento de modo que él mismo cambia si se varían aquellas. En un razonamiento derrotable los presupuestos no forman parte de la identidad del razonamiento y la corrección de la inferencia no depende en absoluto del contenido de los presupuestos. Esto es posible porque la corrección radica en que sin importar cuáles sean los presupuestos, las premisas explícitas los apoyan. Por eso sus conclusiones son inseguras: la confianza en ellas deriva de la que se tiene en los valores de la función f (la conjunción de los presupuestos de las premisas derrotables con sus antecedentes), y este valor tiene sólo la confianza que deriva de las premisas que los apoyan (las que afirman el antecedente de algún condicional derrotable del argumento).

Esta explicación es compatible con la intuición básica de que información adicional invalida la inferencia (Refuerzo del Antecedente): nuevos datos podrían contradecir f(p,q) de modo que p, junto a los nuevos datos, ya no sería evidencia a favor sino que podría ser incluso contraria a la hipótesis. Como la forma de un razonamiento similar al Modus Ponens Derrotable donde p es sustituida por p&&r no garantiza ninguna relación de evidencia (a favor o en contra de la hipótesis) entre p&&r y f(p,q), entonces no puede concluirse nada, ni siquiera tentativamente, acerca de q en tales condiciones con el mismo grado de confianza daría p solo. Acepar este razonamiento implicaría admitir la corrección de razonamientos con una certeza menor a la implicada por los presupuestos de normalidad. De aquí la incorrección del razonamiento.

De este modo un razonamiento derrotable que contiene expreso o implícito entre sus premisas a un condicional derrotable y a su antecedente, supone (en base a la evidencia que proporciona el antecedente expreso del condicional derrotable) la concurrencia de los presupuestos de tal condicional. Esta suposición constituye lo que llamo "presupuestos del razonamiento derrotable". De tales presupuestos, en conjunción con las premisas, se deriva estrictamente la conclusión del mismo modo que en el condicional derrotable de su antecedente en conjunción con los presupuestos del enunciado (contexto) se deriva estrictamente el antecedente. Por eso ambos casos encubren un vínculo clásico: en el caso del condicional, entre la conjunción de sus presupuestos y antecedente por una parte y su consecuente por la otra, en el caso del razonamiento, entre premisas y conclusión.

Análogamente a lo que ocurre con los enunciados derrotables, un razonamiento derrotable puede tener premisas tales que impliquen deductivamente la conclusión, en este caso la inferencia no tendrá presupuestos (no necesarios). Cuando existen tales presupuestos diremos que el razonamiento es genuinamente derrotable.

Condicionales e inferencias derrotables están conceptualmente vinculadas, ambos encubren un condicional clásico y representan intuiciones similares: la de condiciones e información presupuesta. Pero esto no significa que podamos prescindir de alguno de los dos. Sin condicionales derrotables no podemos representar normas jurídicas. Sin razonamientos derrotables no podemos usarlas para tomar decisiones.

Superclasicalidad

Una propiedad general que predicaré de los razonamientos derrotables es que no son siempre razonamientos que puedan ser derrotados, los razonamientos inderrotables también son, al menos, derrotables. La relación de inferencia derrotable es más débil que la inderrotable y por lo tanto es implicada por ella. Esto es, si $p \vdash q$ entonces $p \mid\!\sim q$. Si algo se sigue con necesidad, entonces se sigue el menos de manera incierta. Esto parece en principio razonable, si estamos seguros de algo con base en un conjunto de información, eso implica cualquier grado menor de confianza en la relación de inferencia. De manera análoga a cómo un condicional derrotable se sigue deductivamente del condicional estricto correlativo, deseamos que un razonamiento derrotable sea correcto toda vez que lo sea un razonamiento deductivo correlativo. Por tanto

Si $p \vdash q$ entonces $p \mid\!\sim q$ (regla de superclasicalidad)

Veamos un ejemplo. Sabemos que $p \supset q, p \vdash q$, Por superclasicalidad esto debe implicar $p \supset q, p \mid\!\sim q$. Por ejemplo, sabemos que los que tienen menos de 18 años son menores de edad y que José tiene menos de 18 años, y también sabemos que estas dos premisas implican necesariamente que José es menor de edad. Parece razonable aceptar que "José es menor de edad" se sigue al menos derrotablemente de ese mismo conjunto de premisas.

¿Cómo justificamos esta regla? Así como respecto de la justificación del principio de superclasicalidad para condicionales derrotables mostramos que ella se seguía por definición de la noción de condicional usada, aquí nuestra respuesta dependerá del significado de la noción de consecuencia lógica derrotable.

El consecuente de un condicional derrotable está apoyado por el antecedente en el sentido de que este último más sus presupuestos implican estrictamente al consecuente. El consecuente es tal (es decir, el condicional es verdadero) en la medida en que es el caso que los casos normales en que ocurre el antecedente, que son los casos en que él ocurre junto con los presupuestos (presupuesto de normalidad) son todos casos en los que el consecuente sucede. El condicional es verdadero en el sentido de que afirma la verdad de un condicional estricto acerca de un conjunto de casos normales.

Justamente porque el antecedente de un condicional derrotable implica estrictamente al consecuente dados los presupuestos, es que el antecedente es un buen indicio del consecuente, porque suponemos condiciones normales aunque podemos equivocarnos. Del mismo modo que cuando afirmamos condicionales, cuando razonamos basados en tales condicionales suponemos que ocurren condiciones de normalidad, que los presupuestos del condicional derrotable suceden y que por eso podemos confiar en alguna medida en que el consecuente también pasa. Si las aves (normalmente) vuelan y Twity es un ave, en condiciones de normalidad podemos inferir con seguridad que Twity vuela. Pero con frecuencia no nos preguntamos acerca de tales condiciones de normalidad, lo que hacemos es suponer que ellas concurren e inferir derrotablemente (con un margen de error posible) que Twity vuela. Estamos suponiendo que los presupuestos del condicional derrotable ocurren junto con las premisas del razonamiento. Si ellas se hicieran explícitas, el razonamiento se volvería deductivo.

Como se ve, el razonamiento basado en condicionales derrotables que permite inferencias derrotables supone una forma de necesidad, la necesidad de que la conclusión sucediera si se dieran ciertos supuestos, los que llamamos supuestos de normalidad.

Ahora bien, imaginemos el caso de un razonamiento en el cual las premisas implican la conclusión con necesidad, entonces el razonamiento no tiene presupuestos o, lo que es lo mismo, no existen condiciones de anormalidad en las cuales la conclusión pudiera dejar de extraerse. De modo análogo a cómo los condicionales estrictos implican condicionales derrotables debido a que los primeros son condicionales derrotables sin presupuestos, sucede que los razonamientos deductivos son razonamientos derrotables sin presupuestos. La afirmación de derrotabilidad encubre una afirmación de necesidad: las premisas más los presupuestos del razonamiento implican necesariamente a la conclusión. Cuando el conjunto de presupuestos es vacío, el conjunto de las premisas unido a ese conjunto todavía implica necesariamente a la conclusión. Así, la afirmación de necesidad encubierta sigue siendo verdadera y, por lo tanto, la afirmación de que la conclusión se sigue derrotablemente de las premisas. Esto fundamenta la regla de superclasicalidad para razonamientos derrotables. Veremos la prueba formal en la sección 5.

Por supuesto, la inversa de esta regla no puede ser válida.

4.- Ausencia de analogía

Hasta ahora he señalado varias analogías entre condicionales e inferencias derrotables. Sin embargo, la misma justificación que dimos para la regla del Modus Ponens derrotable, aplicada a otras formas de razonamiento basadas en premisas derrotables, nos mostrará que la analogía no es completa y que por el contrario, condicionales e inferencias derrotables no satisfacen las mismas propiedades.

Oposición de razonamientos

Consideremos el siguiente conjunto de enunciados α={p>q, p>~q, p}. ¿Podemos inferir de aquí q? ¿Y ~q? ¿Son correctos por igual los razonamientos α|~q y α|~ ~q? Veamos.

Dado α, tenemos evidencia dentro del conjunto de premisas tanto para f(p,q) como para f(p,~q). Sin embargo, dado que ellos son inconsistentes (del primero se sigue q y del segundo ~q), sabemos que no pueden ser ambos verdaderos. Por eso, la evidencia que tenemos para cualquiera de ellos se ve disminuida por la evidencia en favor de su falsedad constituida por el apoyo para una proposición opuesta. Así, no podemos obtener de p la confianza requerida ni para f(p,q) ni para f(p, ~q) y en consecuencia, no podemos inferir derrotablemente ni q ni su negación.

La tesis que pretendo defender aquí es que un mismo conjunto de premisas, si es consistente, no puede arrojar correcta y derrotablemente conclusiones contradictorias. Para demostrarlo supongamos que lo hace. En tal caso debería haber entre sus premisas evidencia en favor de cierta hipótesis H1 que implicara cierta conclusión r, y también información contraria a esa hipótesis que apoyara ~r, ya fuera evidencia en favor de una hipótesis contraria H2 antecedente encubierto de un condicional derrotable implicado por las premisas, o información implicada por las premisas que directamente contradijera H1 e implicara ~r. Pero si así fuera entonces la evidencia en favor de r se vería disminuida o incluso anulada por completo por la información implícita en las premisas, y careceríamos de fundamentos para confiar en r.

Por tanto, razonamientos derrotables con premisas consistentes no pueden sustentar conclusiones inconsistentes entre sí.

Así, para los razonamientos derrotables debe ser valer el principio

Si α es consistente, entonces si α|~ r entonces no es el caso que α|~ ~r

Como se habrá advertido, este principio que aceptamos para las inferencias derrotables es análogo al principio de oposición de condicionales, rechazado para los condicionales derrotables. De aquí la ausencia de analogía entre ambos conceptos.

Pero la falta de analogía es más profunda, porque depende del modo de entender las inferencias derrotables. Tal como las hemos presentado, toda inferencia derrotable basada en un conjunto de premisas α tendrá los mismos "presupuestos del razonamiento" que consistirán en el supuesto de verdad del antecedente encubierto de algún condicional derrotable implícito en α. Razonamientos con un mismo conjunto de premisas consistentes no pueden tener presupuestos diferentes, porque la información implícita en ellas sólo apoyará a ciertas hipótesis de manera suficiente. En caso de haber evidencia en favor de hipótesis contradictorias, no apoyará a ninguna de ellas sino cuando alguna se siga deductivamente de las premisas, caso en que sólo se derivará esa hipótesis. De este modo, los presupuestos de un argumento dependen únicamente del conjunto de premisas y no de las consecuencias. Esto muestra una diferencia fundamental con los condicionales derrotables, donde la función de selección es relativa tanto al antecedente cuanto al consecuente. La relatividad de los presupuestos de los razonamientos derrotables al conjunto de las premisas es la fuente de la validez del principio citado al que llamamos aquí "oposición de razonamientos".

5.- Definición y leyes

Una vez establecido que los presupuestos de los razonamientos derrotables son relativos únicamente a las premisas, podemos definir a los razonamientos derrotables usando una función de selección de presupuestos.

Consecuencia lógica derrotable

(Df $|\!\sim$) $\alpha |\!\sim p =_{Df} f(\alpha) \vdash p$

Una definición precisa exigiría dar un conjunto de principios (axiomas) para esta noción de consecuencia, ya fuera indicando leyes válidas directamente para ella o indicando propiedades de la función de selección que figura en la definición.

Un axioma adicional nos permitirá probar inmediatamente la regla de oposición de condicionales. Nuestro axioma afirma que si las premisas de un razonamiento son consistentes, entonces los presupuestos del razonamiento también son consistentes.

Expansión límite Si es posible α entonces es posible $f(\alpha)$

A partir de la definición de Consecuencia Lógica Derrotable y el principio de expansión límite es ya posible demostrar el principio de oposición de inferencias a favor del que argumenté en la sección 4.

Oposición de razonamientos

(OR) Si α es consistente, entonces si $\alpha |\!\sim r$ entonces no es el caso que $\alpha |\!\sim \sim r$

1) α es consistente
2) $\alpha |\!\sim q$
3) $\alpha |\!\sim \sim q$
4) es posible $f(\alpha)$ de 1 por (Expansión límite)
5) $f(\alpha) |\!\sim q$ de 2
6) $f(\alpha) |\!\sim \sim q$ de 3
7) No es posible $f(\alpha)$ (de 4 y 5)

8) No es posible α de 7 y (Expansión límite)
9) No es el caso que $\alpha |\!\sim \sim q$ por reducción al absurdo del supuesto 3.

Un principio obvio para la función de selección, análogo al axioma homónimo para condicionales derrotables y que tomaré también como axioma, es el siguiente:

Expansión $f(\alpha) \vdash \alpha$

Este principio significa que se deducen del resultado de la función (de los presupuestos del razonamiento) todos los elementos de α, es decir, cualquiera de las premisas. En otras palabras, las premisas están entre los presupuestos.

Con expansión es fácil demostrar superclasicalidad.

1) $\alpha \vdash p$
2) $f(\alpha) \vdash \alpha$[4] (expansión)
3) $f(\alpha) \vdash p$ (transitividad de la deducción)
4) $\alpha |\!\sim p$ (definición de consecuencia lógica derrotable)

Monotonía cauta (refuerzo de las premisas con consecuencias derrotables)
Esta regla puede formularse así:

Si $\alpha |\!\sim r$ y $\alpha |\!\sim s$ entonces $\alpha, s |\!\sim r$

Sostengo la invalidez de la regla de monotonía para condicionales derrotables en la forma:

Si p>q y p>s entonces (p&s)> q

(Podemos interpretar el razonamiento anterior como, por ejemplo: los mayores tienen curador, los mayores son capaces, luego los mayores capaces tienen curador).

La regla, que es inválida para condicionales derrotables, afirma que es legítimo reforzar el antecedente con consecuencias derrotables. Su invalidez se debe a que en el caso de los condicionales derrotables la función de selección de presupuestos depende no sólo del antecedente sino del condicional completo, de modo que los condicionales del antecedente de la

[4] La relación de inferencia normalmente se entiende como una relación entre un conjunto de fórmulas y una fórmula, pero también es posible entenderla como una relación entre conjuntos. En el caso de la inferencia deductiva esto no ocasiona ninguna dificultad. Alternativamente se puede tomar al conjunto de fórmulas "consecuencia" como una fórmula única consistente en la conjunción entre las consecuencias del conjunto de las premisas.

regla podrían tener distintos conjuntos de presupuestos, en particular los presupuestos del primero (p>q) podrían no implicar a q, el consecuente del segundo, que podría ser él mismo un hecho derrotable respecto del primero. Si lo fuera, entonces la conjunción p&q, lejos de implicar q debería implicar ~q. En nuestro ejemplo, dado que ser mayor es derrotable respecto de (los mayores tienen curador), es verdad que los mayores capaces no tienen curador (p&s)>~q y es falso (p&s)>q. Esto es suficiente para poner en evidencia la invalidez de la regla de monotonía cauta para condicionales derrotables y justificar que respecto de éstos la regla debe limitarse a reforzar el antecedente con consecuencias seguras, como lo serían las consecuencias estrictas del antecedente del condicional derrotable, esto es, con datos que ya estaban en el antecedente del condicional derrotable originario.

Pero, como hemos visto, en el caso de los razonamientos derrotables los presupuestos son relativos únicamente a las premisas, de modo que dos razonamientos con las mismas premisas tienen los mismos presupuestos y por lo tanto, el conjunto de premisas y presupuestos de ambos deben tener las mismas consecuencias derrotables. Así, agregar a las premisas de un razonamiento las consecuencias derrotables de otro razonamiento que tiene las mismas premisas no puede hacer perder conclusiones, porque esas conclusiones estaban ya necesariamente implicadas por las premisas más los presupuestos del primero.

Formalmente, si 1) $\alpha |\!\sim r$ equivale a $f(\alpha) \vdash r$ (donde $f(x)$ es una función que selecciona los presupuestos de x,que implican estrictamente a los elementos de x) y 2) $\alpha |\!\sim s$ equivale a $f(\alpha) \vdash s$, entonces es claro que s estaba deductivamente implícito en las premisas de 1) y por lo tanto no puede modificar las consecuencias deductivas de éste. Esto es, como para la noción de consecuencia deductiva vale el principio de monotonía a secas, entonces del primero de los dos razonamientos anteriores se sigue $f(\alpha), s \vdash r$.

Debemos aceptar el principio de que si el resultado de una función de selección implica estrictamente algo, entonces la función aplicada a un mismo argumento en conjunción con ese algo (que ya estaba ahí) no varía. Así, asumiremos el siguiente axioma.

Expansión con consecuencias

Si $f(\alpha) \vdash s$ entonces $f(\alpha), s \vdash f(\{\alpha, s\})$ y $f(\{\alpha, s\}) \vdash f(\alpha), s$

Brevemente, la demostración de monotonía cauta es la que sigue:

1) $\alpha |\!\sim r$
2) $\alpha |\!\sim s$
3) $f(\alpha) \vdash r$ (Df$|\!\sim$) en 1
4) $f(\alpha) \vdash s$ (Df$|\!\sim$) en 2
5) $f(\alpha), s \vdash s$ de 4 por lógica deductiva

6) $f(\{\alpha, s\}) \vdash f(\alpha)$, s de 5 y Expansión con consecuencias
7) $f(\{\alpha, s\}) \vdash s$ Transitividad de la deducción en 6 y 5
8) $\alpha, s \mid\!\sim r$ (Df$\mid\!\sim$) en 7.

Monotonía cauta es así válida para los razonamientos derrotables.

Corte cauto, su invalidez

$$\frac{\alpha, q \mid\!\sim r\ ,\ \alpha \mid\!\sim q}{\alpha \mid\!\sim r}$$

Se ha sostenido la validez de este principio para las lógicas no monotónicas argumentando que su fundamento consistía en la sensatez de eliminar premisas consistentes en información que ya estaba en otras premisas que se mantenían. En palabras de Raymundo Morado: "una proposición está *implícitamente* conjuntada con todo aquello que implica",[5] lo que justificaría no hacer explícito eso implicado, es decir, eliminarlo de entre las premisas que lo implicaran. La afirmación de Morado es verdadera cuando se trata de implicación deductiva. En esas condiciones eliminar consecuencias deductivamente implícitas no elimina información, de modo que eliminando las premisas implicadas la conclusión se sigue tanto como se seguía de las premisas originales. Sin embargo, cuando se trata de implicación derrotable, que es una implicación ampliativa (agrega contenido a las premisas), lo implícito "no estaba ahí", o al menos no todo, o no del todo. Eliminar de entre las premisas consecuencias derrotables de otras premisas elimina el contenido adicional que las primeras tenían respecto de las últimas. Así, la conclusión sólo puede seguirse con menos seguridad que originalmente. No podemos usar la misma noción de consecuencia que usábamos en las premisas originales porque el grado de certeza o confiabilidad de la conclusión se debilita. Si originalmente significaba, digamos normalidad, entonces después significa algo menos seguro.

El principio de corte cauto representa un tipo de transitividad que, aunque en menor medida que la transitividad estricta, debilita la conclusión. A través de tal debilitamiento no puede mantenerse el uso de una misma noción de consecuencia. Si la conclusión se sigue después del corte cauto, se sigue mediante una relación de inferencia falible que no es la inferencia derrotable de la que hablamos aquí, ésta supone una cierta fuerza (la que da el concepto de normalidad involucrado en los condicionales derrotables) que es lo que asegura su corrección. Un grado de fuerza menor que ese ya no estaría avalado por la justificación de las inferencias derrotables que he dado, apoyada en condicionales derrotables entendidos como afirmaciones generales de

[5] Raymundo Morado "Problemas filosóficos de la lógica no monotónica", en *Enciclopedia Iberoamericana de Filosofía*, Vol. 27, 2004, pp. 313-344.

normalidad.

Conclusiones

En este trabajo analicé brevemente una noción de consecuencia falible asociada a enunciados derrotables, definidos estos en el sistema CD que presenté parcialmente. Sostuve que para dar cuenta adecuadamente de las inferencias basadas en condicionales derrotables debe apelarse a una noción de consecuencia lógica no deductiva que llamé "derrotable". Ofrecí algunos argumentos para mostrar que tal noción es necesaria y, además, que no es completamente análoga a la noción de derrotabilidad del lenguaje objeto.

Definí a la consecuencia lógica derrotable como una relación deductiva entre las premisas en conjunción con ciertos presupuestos del razonamiento por una parte, y la conclusión por la otra. Señalé además una diferencia muy importante entre la función de selección que debe usarse para definir la noción de consecuencia derrotable y la adecuada para definir condicionales derrotables. En el caso de la noción de consecuencia, sostengo, ella es relativa únicamente a las premisas, lo que evita que razonamientos con premisas diferentes puedan tener diferentes presupuestos.

Si tengo razón en esas propuestas, los razonamientos derrotables se comportan de modo diferente que los condicionales derrotables, básicamente debido a las diferencias en la función de selección. Que ésta sea relativa únicamente a las premisas permite fundamentar dos principios importantes: Monotonía Cauta y Oposición de Razonamientos, cuyos principios análogos no son válidos para los condicionales derrotables en mi concepción de ellos.

Aunque principios como los señalados muestran que la lógica del razonamiento derrotable es más fuerte que la de los condicionales, aún hay principios importantes que no satisface, como argumenté respecto del principio de Corte Cauto.

No he dado un sistema completo, mi objetivo fue sólo proponer que hay analogías importantes y una diferencia básica entre la lógica de los condicionales y la de consecuencia derrotables. Las similitudes son fuertes, como debe esperarse si la última ha de ser la lógica que permite aplicar los condicionales derrotables en razonamientos falibles. La lógica de la inferencia derrotable es más fuerte en el sentido de que valen para ella los principios de la lógica de los condicionales derrotables más otros. En esta medida se aproxima a las concepciones estándar sobre lógicas derrotables como las de Kraus, Lehmann y Magidor (1990) y Morado (1994). Sin embargo creo que tales propuestas son todavía demasiado generosas con la relación de inferencia, permiten demasiado y aceptan inferencias cada vez más débiles. Un razonamiento puede ser falible sin ser débil.

Bibliografía

C. A. Alchourrón (1996), "Defeasible Logics: Demarcation and Affinities" en G. Crocco, L. Fariñas del Cerro y A. Herzig (Comps.) *Conditionals and Artificial Intelligence*, Oxford University Press, pp. 67-106.

_______ (1991), "Philosophical foundations of deontic logic and the logic of defeasible conditionals", en J.J. Mayer & R.J. Wieringa (eds.) *Deontic Logic in Computer Science: Normative Systems Specification*, Wiley & Sons, pp. 11-13.

S. Kraus, D. Lehman y M. Magidor (1990), "Nonmonotonic Reasoning, Preferencial Models and Cumulative Logics", en *Artificial Intelligence* 44, pp. 167-207.

Raymundo Morado (2004), "Problemas filosóficos de la lógica no monotónica", en *Enciclopedia Iberoamericana de Filosofía*, Vol. 27 Filosofía de la Lógica, Raul Orayen y Alberto Moretti (eds)., Trotta-Consejo Superior de Investigaciones Científicas, pp. 313-344.

María Inés Pazos (2002), "Normas Derrotables", en *Perspectivas y horizontes de la Filosofía de la Ciencia a la vuelta del nuevo milenio*, Instituto de Investigaciones Filosóficas, UNAM, Ambrosio Velasco Gómez editor, pp. 181-212.

H. Prakken (1997), *Logical Tools for Modelling Legal Argument. A Study of defeasible Reasoning in Law*. Kluwer Academic Publishers, Dordrecht, Law and Philosophy Library.

H. Prakken and G. Sartor (Eds.) (1997), *Logical Models of Legal Argumentation*, Kluwer Academic Publishers, Dordrecht.

H. Prakken y G. Vreeswijk (2002), "Logics for defeasible argumentation". In D. Gabbay & F- Guenthner (eds.), *Handbook of Philosophical Logic*, segunda edición, Vol. 4. Kluwer Academic Publishers, pp. 219-318.