

OBLIGACIONES *PRIMA FACIE* Y DERROTABILIDAD

CARLOS A. OLLER

Universidad de Buenos Aires y Universidad Nacional de La Plata, Argentina

Resumen

Este artículo examina algunos de los problemas que presenta el tratamiento lógico de las obligaciones *prima facie* en la obra de Carlos Alchourrón. Por una parte, señalaremos que su sistema para los condicionales derrotables *DFT* no formaliza adecuadamente la noción intuitiva de condición contribuyente que Alchourrón utiliza para elucidar la de condicional derrotable. Por otra parte, argumentaremos que la noción de deber *prima facie* de David Ross no queda adecuadamente formalizada en el sistema *AD* de lógica deóntica que Alchourrón construye como una extensión de *DFT*.

PALABRAS CLAVE: obligación *prima facie* - condicional derrotable - condición contribuyente.

Abstract

This paper deals with some problems concerning the logical treatment of *prima facie* obligations in the work of Carlos Alchourrón. On the one hand, we will argue that his *DFT* system for defeasible conditionals does not adequately formalize the intuitive notion of contributory condition used by Alchourrón to elucidate the notion of defeasible conditional. On the other hand, we will argue that David Ross' notion of *prima facie* duty is not adequately formalized in Alchourrón's deontic system *AD*, an extension of his *DFT* logic.

KEY WORDS: *prima facie* obligation - defeasible conditional - contributory condition.

1. Introducción: derrotabilidad y lógicas condicionales

Carlos Alchourrón publicó a mediados de los años noventa una serie de artículos sobre la lógica de los condicionales derrotables o anulables (*defeasible conditionals*) [Alchourrón, 1993, 1994, 1996]. En estos trabajos ofrece una elucidación filosófica de la noción de derrotabilidad (*defeasibility*) que aplica para aclarar conceptos deónticos tales como el de obligación *prima facie*.

El problema de la derrotabilidad de los enunciados y de las inferencias es generado, según Alchourrón, por ciertas oraciones del lenguaje ordinario que expresan un conocimiento incompleto. Los seres humanos extraen conclusiones razonables, pero derrotables, a partir de información incompleta en ausencia de evidencia que haría implausibles a esas

inferencias. Son capaces de extraer conclusiones a partir de la información de que disponen en un momento dado, y de actuar de manera acorde, pero también suelen retractar esas conclusiones si éstas resultan incompatibles con la nueva información que reciben.

El problema del razonamiento a partir de información incompleta puede tratarse desde por lo menos dos puntos de vista diferentes, pero que Alchourrón juzga equivalentes. Por una parte, se puede utilizar una noción de consecuencia no-clásica que refleje el carácter derrotable de este tipo de inferencias. Por otra parte, se puede apelar a una conectiva condicional no-clásica para formalizar aquellas oraciones que, a pesar de formularse como si fueran enunciados universales, tienen excepciones sobreentendidas que no aparecen explícitamente en su formulación. De esta manera, se logra tratar el problema de la derrotabilidad sin abandonar la lógica deductiva utilizando una lógica condicional que es una extensión de la lógica elemental clásica.

Alchourrón favorece el segundo enfoque para el tratamiento formal de la derrotabilidad de los enunciados y de las inferencias, y prefiere elucidar el razonamiento derrotable en el marco de la lógica deductiva. Este enfoque difiere del que, por los años noventa, era ya predominante en la literatura sobre las lógicas no-monótonas –de las cuales las lógicas derrotables son una variedad–. Sin embargo, la posición de Alchourrón puede entenderse teniendo en cuenta las posiciones que había defendido en sus trabajos anteriores acerca del razonamiento normativo y su firme convicción de que la lógica deductiva es una herramienta básica e indispensable para el análisis filosófico.

El análisis deductivista que Alchourrón propone de la derrotabilidad ha sido objeto de reparos de diverso tipo [Loui, 1997], aunque no nos ocuparemos aquí de esas reacciones críticas. Dedicaremos este artículo a presentar algunos problemas que, a nuestro entender, tiene el tratamiento lógico de las obligaciones *prima facie* en el marco de la lógica condicional *DFT* que Alchourrón propone para reflejar el comportamiento inferencial de los condicionales derrotables, continuando así con lo desarrollado en [Becher et al., 1999]. Por una parte, señalaremos que este sistema no formaliza adecuadamente la noción intuitiva de condición contribuyente que Alchourrón utiliza para elucidar la de condicional derrotable. Por otra parte, argumentaremos que la noción de deber *prima facie* de David Ross [Ross, 1939] no queda adecuadamente formalizada en el sistema *AD* de lógica deóntica que Alchourrón construye como una extensión de *DFT*, su lógica para los condicionales derrotables.

2. Condicionales derrotables: el sistema *DFT* de Alchourrón

Alchourrón sostiene que las construcciones condicionales del lenguaje ordinario son usadas a menudo de manera tal que el antecedente A , junto con un conjunto de supuestos tácitamente aceptados en el contexto de emisión del condicional, constituyen una condición suficiente para el consecuente B . En estos casos A es, de acuerdo a la terminología que Alchourrón toma de von Wright [von Wright, 1951], una condición contribuyente para B —i.e. una condición necesaria de una condición suficiente para B —.

Alchourrón simboliza “ A es una condición contribuyente para B ” con “ $fA \Rightarrow B$ ”, donde fA se usa para simbolizar la aserción conjunta de A y el conjunto de supuestos que lo acompañan tácitamente y \Rightarrow es el operador de implicación estricta del sistema de lógica modal *S5*. Además, como abreviatura de “ $fA \Rightarrow B$ ” utiliza “ $A > B$ ”, donde la conectiva $>$ es el condicional derrotable cuyo comportamiento inferencial se trata de determinar.

Las propiedades lógicas del operador de implicación estricta \Rightarrow y del operador f determinan una lógica condicional, el sistema *DFT* de Alchourrón para los condicionales derrotables. Además, dadas esas propiedades, el operador f es un tipo especial de la función de revisión caracterizada por *AGM*, la teoría del cambio racional de creencias desarrollada por Alchourrón, Gärdenfors y Makinson en 1985 [Alchourrón, Gärdenfors & Makinson, 1985]. De esta manera, Alchourrón integra en su teoría de los condicionales derrotables uno de sus aportes anteriores a la teoría lógica.

El sistema axiomático $DFT_{>}$ aísla las fórmulas válidas de la lógica *DFT* en términos del operador condicional derrotable $>$. La lógica $DFT_{>}$, como Alchourrón lo señala, puede considerarse una variante notacional de la lógica de las obligaciones condicionales *DSDL3* de B. Hansson [Hansson, 1969]. Este sistema puede caracterizarse mediante las siguientes definiciones, axiomas y regla de inferencia:

- (Def. \Box) $\Box A =_{\text{df}} (\neg A > \perp)$
 (Def. \geq) $(A \geq B) =_{\text{df}} (\Box \neg (A \vee B) \vee \neg ((A \vee B) > \neg A))$
 ($>.1$) $\vdash (A > A)$ (Identidad)
 ($>.2$) $\vdash (A > (B \wedge C)) \equiv ((A > B) \wedge (A > C))$ (Distribución en el consecuente)
 ($>.3.1$) $\vdash ((A > B) \wedge (A > C)) \supset ((A \vee B) > C)$ (Aglomeración en el antecedente)
 ($>.3.2$) $\vdash (A \geq B) \supset (((A \vee B) > C) \supset (A > C))$ (Distribución parcial restringida en el antecedente)
 ($>.4$) $\vdash (\neg A > \perp) \supset A$ (Subordinación modal)
 (Regla de extensionalidad) Si $\vdash (A \equiv B)$, entonces $\vdash (A > C) \equiv (B > C)$ y $\vdash (C > A) \equiv (C > B)$

Los modelos para el sistema $DFT_{>}$, que son los mismos que los de DFT , pueden caracterizarse como triplos ordenados $M^\alpha = \langle W, [], Ch^\alpha \rangle$, donde W es un conjunto no-vacío de mundos, $[\]$ es una función de interpretación del lenguaje que correlaciona cada oración A con un subconjunto de W (el conjunto de mundos en los que A es verdadera) y Ch^α es una función de elección que adjudica a cada oración un subconjunto de W (el conjunto de mundos en que A y los supuestos que acompañan a A según el agente α son todos verdaderos). La función de selección Ch^α debe cumplir con las siguientes restricciones:

- ($Ch^\alpha.1$) $Ch^\alpha(A) \subseteq [A]$ (Expansión)
 ($Ch^\alpha.2$) Si $[A] = [B]$, entonces $Ch^\alpha(A) = Ch^\alpha(B)$ (Extensionalidad)
 ($Ch^\alpha.3$) Si $[A] \neq \emptyset$, entonces $Ch^\alpha(A) \neq \emptyset$ (Expansión limitada)
 ($Ch^\alpha.4$) $Ch^\alpha(A)$ ó
 $Ch^\alpha(A \vee B) = Ch^\alpha(B)$ ó
 $Ch^\alpha(A) \cup Ch^\alpha(B)$ (Ordenación jerárquica)

La cláusula de verdad para la conectiva $>$ es :

(C. $>$) $w \in [(A > B)]$ si y sólo si $Ch^\alpha(A) \subseteq [B]$

2. Condiciones contribuyentes y condicionales derrotables

Alchourrón toma el concepto de condición contribuyente –la noción central en su reconstrucción de la lógica de los condicionales derrotables– de la teoría de las condiciones de von Wright [von Wright, 1951]. Sin embargo, como se verá a continuación, la noción intuitiva de condición contribuyente no parece quedar adecuadamente reflejada por la lógica DFT para los condicionales derrotables.

En primer lugar, la definición de la condición contribuyente usada por Alchourrón – A es una condición contribuyente de B si y sólo si A es una condición necesaria de una condición suficiente de B – hace que esa noción sea trivial, si no se hace ninguna otra especificación. En efecto, dado cualquier par de las oraciones A y B , A es siempre una condición contribuyente de B porque $(A \wedge B)$ es una condición suficiente de B , y A es una condición necesaria de $(A \wedge B)$.

Además, la noción intuitiva de la condición contribuyente no parece validar algunos de los principios sancionados por la lógica de los condicionales derrotables DFT de Alchourrón. Para ver esto, supóngase que A es lógicamente posible y que, como bien puede suceder, las dos oraciones siguientes son verdaderas:

- (1) $(A \wedge B) \Rightarrow C$
 (2) $(A \wedge \neg B) \Rightarrow \neg C$

La interpretación de A como “ a es un número par”, la de B como “ b es un número impar” y la de C como “ $a + b$ es un número impar” nos proporciona un ejemplo de un par de oraciones verdaderas que tienen esa forma. En este caso A es una condición contribuyente de C y es también una condición contribuyente de $\neg C$. Por lo tanto, la oración siguiente es también verdadera:

- (3) $(A > C) \wedge (A > \neg C)$

Sin embargo, según la semántica de Alchourrón para los condicionales derrotables, cuando A es lógicamente posible, (3) es siempre falso.

Dado que (3) no es inválido según la noción intuitiva de condición contribuyente, el siguiente axioma del sistema $DFT_{>}$ de Alchourrón no puede ser válido en una lógica de los condicionales derrotables inspirada en esa noción intuitiva:

- (4) $(A > (B \wedge C)) \equiv ((A > B) \wedge (A > C))$

Porque, si (4) fuera válido, entonces (5) podría también ser verdadero, dado que (3) puede serlo:

- (5) $A > (C \wedge \neg C)$

Pero, (5) puede solamente ser verdad si fA es una contradicción, lo que no puede ser el caso según la noción intuitiva de la condición contribuyente, si A es lógicamente posible.

Así pues, la clase de derrotabilidad involucrada en la noción de condición contribuyente no parece ser identificable con la clase de derrotabilidad formalizada por los sistemas como DFT . En particular, no parece identificable con la noción de condicionalidad expresada por los condicionales de la clase “Normalmente, si A entonces B ” y “Típicamente, si A entonces B ”.

Una noción formal de condicionalidad que parece satisfacer al menos algunas de las características de la noción intuitiva de condición contribuyente puede ser caracterizada semánticamente como sigue: A es una condición contribuyente de B si y sólo si la intersección del conjunto de los A -mundos y el conjunto de los B -mundos no es vacía. Esta definición no trivializa la noción de la condición contribuyente, y se puede

considerar como calificación de la noción intuitiva usada por Alchourrón. De hecho, es equivalente a la definición siguiente: A es una condición contribuyente de B si y sólo si existe un conjunto de C -mundos tal que este conjunto está incluido tanto en el conjunto de los A -mundos como en el conjunto de los B -mundos y el conjunto de los C -mundos no es vacío, es decir, si y sólo si A es una condición necesaria de una condición suficiente no-vacua para B .

La conectiva que expresa esta noción de condicionalidad es derrotable, es decir no satisface ni el refuerzo del antecedente ni el *modus ponens*. Además, no satisface la distribución consiguiente y, por lo tanto, una oración lógicamente posible A puede ser una condición contribuyente tanto para una oración B como para su negación. Pero, esta noción es también simétrica –es decir, si A es una condición contribuyente de B , entonces B es una condición contribuyente de A –. Aunque ésta parece ser una propiedad contraintuitiva para una noción de la condición contribuyente, conjeturamos que no es posible encontrar una mejor contraparte formal no trivial de la noción intuitiva de la condición contribuyente.

3. Obligaciones *prima facie* y condicionales derrotables

Alchourrón sostiene que la noción de obligación *prima facie* de David Ross [Ross, 1939] se puede formalizar en un sistema de la lógica deóntica construido como una extensión de *DFT*, su lógica para los condicionales derrotables.

En su artículo “*Detachment and defeasibility in deontic logic*” [Alchourrón, 1996] utiliza la expresión “deber *prima facie*” para denotar las obligaciones derrotables incondicionales. El operador de obligación de esta clase se define en su sistema *AD* (*actual and defeasible duties*) –que resulta de añadir el operador de obligación O del sistema de lógica deóntica estándar a *DFT*– como un condicional de *DFT* que tiene una oración que expresa una obligación no derrotable o efectiva como su consecuente y la constante T (*verum*) como su antecedente:

$$Od(A) =_{df} T > OA \quad (\text{obligación incondicional derrotable})$$

El operador O de esta definición es el operador de la obligación del sistema estándar de lógica deóntica, es decir el sistema caracterizado por las leyes y la regla siguientes:

- A1. $\vdash O(T)$
 A2. $\vdash O(A \wedge B) \equiv O(A) \wedge O(B)$

$A3. \vdash O(\perp)$

$RI. \text{ Si } \vdash A \equiv B, \text{ entonces } \vdash O(A) \equiv O(B)$

El operador Od , como Alchourrón precisa, también satisface los principios del sistema estándar de lógica deóntica. Por lo tanto, se sigue de las contrapartes de $A2$ y de $A3$ para el operador Od que en el sistema de Alchourrón no podemos tener obligaciones *prima facie* que estén en conflicto, es decir que la siguiente es una tesis de AD :

$\vdash \neg(Od(A) \wedge Od(A))$

Sin embargo, Ross establece claramente que, aunque las obligaciones efectivas no pueden estar en conflicto, podemos tener obligaciones *prima facie* que estén en conflicto:

It is the overlooking of the distinction between obligations and responsibilities, between actual obligatoriness and the tendency to be obligatory, that leads to the apparent problem of conflict of duties, and it is by drawing the distinction that we solve the problem, or rather show it to be non-existent. For while an act may well be *prima facie* obligatory in respect of one character and *prima facie* forbidden in virtue of another, it becomes obligatory or forbidden only in virtue of the totality of its ethically relevant characteristics. [Ross, 1939, p.86]

Este punto también puede ser considerado desde un ángulo semántico. El sistema formal de Alchourrón se puede describir, usando las palabras de Brian Chellas, como uno en el cual “asociado a cada mundo posible (...) hay una sola clase de alternativas deónticas”; pero, cuando nos estamos ocupando de deberes *prima facie* nos enfrentamos a una “pluralidad de fundamentos para las obligaciones” [Chellas, 1974]. Por lo tanto, la caracterización de Alchourrón de los deberes *prima facie* en términos de la noción de obligación del sistema estándar de lógica deóntica no parece ser en absoluto adecuada, si estamos buscando una formalización de la noción de Ross. De hecho, es incompatible con algunas características básicas de esta noción.

Los deberes *prima facie* formalizados por Alchourrón se pueden caracterizar como obligaciones que valen en circunstancias normales, pero que pueden ser derrotadas en circunstancias anormales. Por otra parte, la noción de deber *prima facie* de Ross intenta reflejar el hecho de que evaluamos un acto según una pluralidad de principios morales, algunos de los cuales pueden recomendarlo mientras que otros pueden desaconsejar-

lo. Por lo tanto, un conflicto de obligaciones *prima facie* no sólo no es imposible según la teoría de Ross sino que tampoco es infrecuente. Este hecho ha sido reconocido en la literatura y así Asher y Bonevac [Asher & Bonevac, 1996] han presentado una lógica deóntica no-monótona que pretende reflejar esta diferencia entre las obligaciones efectivas y las obligaciones *prima facie*.

El rechazo de *A2* como principio válido para las obligaciones *prima facie* (en el sentido de Ross) parece ser un paso necesario en el camino hacia una lógica adecuada para esta clase de deberes. Esto está de acuerdo con la teoría de Ross: el hecho de que podemos estar bajo dos obligaciones *prima facie* contradictorias no significa que estemos bajo obligación *prima facie* de realizar un acto autocontradictorio y, por lo tanto, lógicamente imposible. Además, dado que nos estamos ocupando aquí de deberes *prima facie*, los conflictos de obligaciones que esta lógica permitiría no son conflictos genuinos: no nos enfrentan a auténticos dilemas morales, es decir situaciones en las cuales es imposible para el agente hacer lo que tiene el deber de hacer.

Otro punto en que la teoría de Alchourrón difiere de la de Ross es el de la relación entre los deberes *prima facie* y los deberes efectivos. En la teoría de Ross las obligaciones efectivas son también obligaciones *prima facie*:

[...] we are *not* obliged to do that which is only *prima facie* obligatory. We are only bound to do that act whose *prima facie* obligatoriness in those respects in which it is *prima facie* obligatory most outweighs its *prima facie* disobligatoriness in those respects in which it is *prima facie* disobligatory. [Ross, 1939, p. 85]

Sin embargo, Alchourrón precisa que los principios siguientes son inválidos en su sistema *AD*:

- |- $O(A) \supset Od(A)$
- |- $Od(A) \supset O(A)$

Por lo tanto, en *AD* los deberes *prima facie* son lógicamente independientes de los deberes efectivos. Esto es, como debe ser, dado la interpretación de Alchourrón de las obligaciones *prima facie*. Pero, si deseamos formalizar la noción de Ross de deber *prima facie* en un sistema de lógica deóntica, tal sistema debe incluir necesariamente el primero de esos dos principios.

Referencias

- Alchourrón, C. (1993), "Philosophical Foundations of Deontic Logic and the Logic of Defeasible Conditionals", en Meyer, J.J. and Wieringa R.J. (eds.), *Deontic Logics in Computer Science: Normative System Specification*, 43-84, Wiley & Sons, Cichester.
- (1994), "Defeasible Logic: Demarcation and Affinities", en Crocco, G., Fariñas del Cerro, L. y Herzig A. (eds), *Conditionals: from Philosophy to Computer Science*.67-102, Clarendon Press, Oxford.
- (1996). "Detachment and defeasibility in deontic logic", en *Studia Logica* 51, 5-18.
- Alchourrón, C., Gardenfors, P. & Makinson, D. (1985), "On the Logic of Theory Change: Partial Meet Contraction and Revision Functions", en *Journal of Symbolic Logic* 50: 510-530.
- Asher, N. & Bonevac, D. (1996), "Prima Facie Obligation", en *Studia Logica* 57: 19-45.
- Becher, V., Fermé, E., Lazzer, S., Oller, C., Palau, G. & Rodríguez, R. (1999), "Some observations on Carlos Alchourrón's theory of defeasible conditionals", en McNamara P. and Prakken, H. (eds.), *Norms, Logics and Information Systems. New studies on Deontic Logic and Computer Science*, 219-230. IOS Press, Amsterdam/Tokio/Washington DC.
- Chellas, B. (1974), "Conditional Obligation", en Stendlund, S. (ed.), *Logical Theory and Semantic Analysis*, 23-33. Reidel, Dordrecht.
- Hansson, B. (1969), "An analysis of some deontic logics", en *Nous* 3: 373-398.
- Loui, R.P. (1997), "Alchourrón and von Wright on Conflict among Norms", en Nute, D. (ed.), *Defeasible Deontic Logic*, 345-351. Kluwer, Dordrecht.
- Ross, D. (1939), *Foundations of Ethics*, Clarendon Press, Oxford.
- von Wright, G.H. (1951), *A Treatise on Induction and Probability*, Routledge & Kegan Paul, London.