EPISTEMOLOGÍA E HISTORIA DE LA CIENCIA

SELECCIÓN DE TRABAJOS DE LAS IX JORNADAS

VOLUMEN 5 (1999), Nº 5

Eduardo Sota
Luis Urtubey
Editores

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

Sistemas de argumentación rebatible y lógica polivalente[1]

*Gustavo Adrián Bodanza**

1. Introducción

Este trabajo tiene un doble propósito: primero, mostrar que la argumentación rebatible puede brindar una semántica clara para algunas lógicas polivalentes, como las que consideran algún valor intermedio no clasificable como distinguido ("emparentado" con la verdad) ni como antidistinguido ("emparentado" con la falsedad); y segundo, mostrar cómo puede operarse lógicamente con el conocimiento tentativo producido (justificado) por un sistema argumentativo. El primer propósito es de importancia lógica. El segundo es de importancia para la Inteligencia Artificial —el campo en el que se han desarrollado los sistemas argumentativos—, ya que el sistema que se presenta sugiere implementaciones computacionales que mejorarían la eficiencia de los programas basados en argumentación rebatible.

El plan se organiza como sigue. En la sección 2 se introducen los conceptos básicos de un sistema argumentativo. En la sección 3 se definen cinco posibles valores que puede tomar una proposición de acuerdo a su estado de justificación en un sistema de argumentación rebatible: V (contextualmente verdadera), F (contextualmente falsa), J (justificada), I (injustificada) y D (desconocido). Luego, en la sección 4 se construye un lenguaje lógico definiendo a partir de esos valores un conjunto de conectivos. Este lenguaje permitirá construir una lógica pentavalente con determinadas características (como, bajo determinadas condiciones, no tener tautologías y hacer válidas todas las reglas de inferencia de la lógica clásica bivalente). Finalmente, en la sección 5 se propone un modo de expresar axiomáticamente en esa lógica el sistema de creencias representado en el sistema argumentativo del cual se toman los valores. Se demuestra que tal sistema axiomático es completo respecto de las proposiciones justificadas en el sistema argumentativo. Por último, se destaca la importancia de este desarrollo desde el punto de vista de la implementación de este tipo de sistemas en programas de Inteligencia Artificial.

2. Sistemas argumentativos

2.1. Lenguaje para un sistema de argumentación

Sea L un lenguaje proposicional formado por: a) variables proposicionales: p, q, r, etc., b) conectivos: $\neg$, $\vee$, $\wedge$ y $\rightarrow$, y c) signos de puntuación: (,). Las fbfs de este lenguaje serán las definidas habitualmente en lógica. Sea '$\Rightarrow$' un símbolo metalingüístico, llamado 'regla rebatible' o 'regla *default*', que conecta dos fórmulas de L, y cuya interpretación informal es ''⟨fórmula de la izquierda⟩ es una buena razón para sostener ⟨fórmula de la derecha⟩''. La introducción de estas reglas permitirá derivar conclusiones tentativas, es decir, conclusiones que podrían ser rechazadas si entrasen en contradicción con otras conclusiones (luego se hablará del proceso por el cual se decide cuándo se efectúan tales rechazos). Que una fórmula φ se derive tentativamente de un conjunto de fórmulas F mediante un conjunto de reglas rebatibles D, se simbolizará con '$F \cup D \mid\!\sim \varphi$'.

* Universidad Nacional del Sur.

En lo que sigue llamaremos 'reglas' tanto a las fórmulas de L cuyo conectivo principal sea $\rightarrow$, como a las reglas rebatibles, especificando de cuál tipo se habla sólo cuando fuere necesario.

2.2. Sistemas de argumentos: argumentos, argumentos concluyentes, argumentos rebatibles.

Los sistemas argumentativos aparecen en el marco de la Inteligencia Artificial a través del estudio del razonamiento rebatible. Los primeros trabajos corresponden a Poole (1985), Lin y Shoham (1987) y Loui (1987), y se continúan con Simari y Loui (1992), Prakken (1993), Dung (1995), Kowalski y Toni (1996) y otros. Las ideas seminales más citadas de la argumentación son las producidas en los estudios lógico-filosóficos de Toulmin (1958) y Rescher (1977). El fin de un (así llamado) *sistema argumentativo* es captar formalmente el carácter del razonamiento rebatible. Supondremos para la faz puramente deductiva del razonamiento el mecanismo de la teoría de pruebas clásica. Por esta razón nos ocuparemos sólo de caracterizar la argumentación rebatible, presuponiendo simplemente que no puede haber conflictos entre argumentos válidos, en el sentido de las lógicas consistentes (resguardaremos esta idea requiriendo la consistencia del conjunto de axiomas del sistema).

DEFINICIÓN 1

Un *sistema argumentativo* es una terna $\langle K, D, Args\rangle$, donde K es un conjunto finito y consistente de fbfs de L (axiomas) llamado 'contexto de creencias' o simplemente 'contexto', D es un conjunto finito de reglas rebatibles y $Args \subseteq 2^D$.

Razones de diseño hacen que distintos sistemas argumentativos consideren definiciones disímiles de 'argumento rebatible'. Para nuestros fines será suficiente hablar de argumentos rebatibles en forma abstracta, ya que sólo nos interesa aquí estudiar el rol de éstos en la valoración veritativa de sentencias en una lógica polivalente. Para esto será importante hacer hincapié en la interacción entre argumentos, pero no tanto en el carácter de sus estructuras internas.

Consideraremos que un argumento (en general) es un conjunto de reglas rebatibles que, aplicadas sobre el conocimiento en K (es decir, las premisas del argumento), y con ayuda de ciertas reglas de inferencia especificadas en el sistema,[2] permiten establecer una prueba para alguna sentencia de L. Si un argumento es vacío, es decir, con la sola aplicación de las reglas de inferencia sobre elementos de K se puede obtener una conclusión, diremos que se trata de una prueba *deductiva* o *concluyente*. Si, en cambio, el argumento no es vacío (i.e., es un conjunto que contiene al menos una regla del tipo '$\Rightarrow$'), diremos que la prueba es un *argumento rebatible*. Todos los elementos del conjunto inicial de creencias (contexto) son soportados por argumentos vacíos y hay, por lo tanto, pruebas concluyentes para ellos. Indicaremos las pruebas concluyentes con el símbolo '$\vdash$', como es habitual en lógica clásica, ubicando el conjunto de premisas a su izquierda y la conclusión a su derecha.

Los requisitos mínimos que pretendemos que cumpla un argumento rebatible se expresan en la siguiente

DEFINICIÓN 2

Dado un sistema argumentativo $SA=\langle K, D, Args\rangle$, $T\in Args$, y llamamos a T *argumento rebatible*, si y sólo si para alguna fórmula atómica $\alpha \in L$:

1 $T \cup K \mid\!\sim \alpha$ (α se deriva rebatiblemente de T junto con K)

2 no es el caso que $T \cup K \mid\!\sim \bot$ (T junto con K no produce inconsistencias)

3 $\forall T' \subset T$, no es el caso que $T' \cup K \mid\!\sim \alpha$ (T es mínimo)

EJEMPLO 1

(Pruebas concluyentes y argumentos rebatibles). En un sistema argumentativo, sea $K=\{p, p \rightarrow r, q \rightarrow \neg r\}$ y sea $D=\{p \Rightarrow q, r \Rightarrow s, s \Rightarrow t\}$. Entonces todos los elementos de K tienen pruebas concluyentes, y así también r (siempre que el modus ponens sea una regla de inferencia del sistema). Además, $\{r \Rightarrow s, s \Rightarrow t\}$ es un argumento rebatible para t, y $\{p \Rightarrow q\}$ no es un argumento rebatible (viola la condición 2).

2.3. Solución de inconsistencias: interacción entre argumentos

El sentido común suele plantearnos situaciones en las que se nos presentan buenas razones para creer algo, y otras buenas razones para creer lo contrario. Por supuesto, estas "buenas razones" no son concluyentes, a menos que nuestro sistema de creencias colapse por completo (siempre hablamos de un sistema totalmente consistente). Si nuestras premisas son consistentes, la solución de inconsistencias a nivel de las conclusiones sólo es posible si disponemos de criterios para decidir cuáles de esas "buenas razones" son preferibles a las otras.

Un sistema argumentativo nos permite representar las buenas razones a través de los argumentos rebatibles. Pero si además queremos resolver las inconsistencias que argumentos y contraargumentos[3] pudieran provocar, debemos incluir en el sistema alguna relación de preferencia entre esos argumentos. Tal relación de preferencia debe ser racionalmente adecuada; por ejemplo, debería privilegiar a las pruebas concluyentes respecto de los argumentos rebatibles.[4] La relación de preferencia permitirá especificar cuándo un argumento es rebatido por otro, de acuerdo al orden que ésta defina sobre el conjunto de todos los argumentos del sistema. Ahora bien, podría ocurrir que, por cuestiones formales, el orden establecido por las preferencias permitiera la aparición de ciclos en las secuencias de rebatibilidad (e.g., el argumento X rebate al argumento Y, éste rebate a Z, y Z rebate a X). De modo que, en estos casos, se requiere además la regulación de un proceso dialéctico que, a través de un conjunto de reglas marco, determine cuáles son los argumentos que quedan sin rebatir en última instancia (por ejemplo, reglas que impidan que un argumento sea atacado dos veces de la misma forma en el proceso dialéctico).

Vamos a considerar, entonces, un sistema argumentativo en el que se declaren justificados a todos los argumentos que no resulten rebatidos por ningún contraargumento en la última instancia del proceso dialéctico definido en el sistema, y de acuerdo a las reglas marco que en él se establezcan. A los argumentos justificados por ese proceso los llamaremos *p-justificados*. De este modo, la argumentación nos permitirá evaluar a las fórmulas del lenguaje L de acuerdo a su condición de "p-justificabilidad" en un sistema como el considerado. Una vez que definamos con precisión los valores que pueden tomar las fórmulas, utilizaremos a éstos como semántica para una lógica polivalente que construiremos.

3. Definición de valores de acuerdo al estado de justificación de las proposiciones

Como hemos visto, la argumentación tiene pleno sentido en un diálogo, es decir, cuando la aceptabilidad de una proposición se dirime a través de un proceso dialéctico de argumentación y contraargumentación. Recordemos haber asumido que una proposición o sentencia está *p-justificada* en el contexto de un determinado cuerpo de creencias cuando hay un argumento rebatible que soporta a esa proposición que no es rebatido por ningún contraargumento en la instancia final de un diálogo, de acuerdo a las reglas y a las condiciones de rebatibilidad del mismo. Recordemos también que al cuerpo de creencias que forma el contexto de un diálogo lo suponemos consistente y a sus elementos irrebatibles, así como todo lo que se derive lógicamente de ellos (esto brinda un requisito mínimo de racionalidad).

De acuerdo a esto, definimos cinco valores posibles para las proposiciones o sentencias de L en el marco de un sistema argumentativo:

- *verdadero —en el contexto—* (V): σ es verdadero (en el contexto) ssi hay un argumento concluyente para σ.
- *falso —en el contexto—* (F): σ es falso (en el contexto) ssi hay un contraargumento concluyente para σ.
- *justificado* (J): σ tiene valor justificado ssi σ no es V ni F, pero es soportado por un argumento p-justificado.
- *injustificado* (I): σ tiene valor injustificado ssi σ no es V ni F, pero hay un contraargumento para σ que está p-justificado.
- *desconocido* (D): el valor de σ es desconocido ssi σ no es V ni F y no tiene argumentos ni contraargumentos p-justificados (es decir, si no puede determinarse contextualmente su valor).

Nótese que V y F no son lo mismo que la verdad y la falsedad tarskianas, respectivamente. En una interpretación bivalente toda proposición es verdadera o falsa en sentido tarskiano, pero puede haber proposiciones que no sean ni V ni F (contextualmente).

De acuerdo a esta semántica introducimos una nueva negación como sigue:

φ	$\neg_5\varphi$
V	F
J	I
D	D
I	J
F	V

Si consideramos como valores distinguidos a V y J, como antidistinguidos a F e I, y hacemos corresponder a D un valor intermedio, nuestra negación '$\neg_5$' resulta análoga a la negación trivalente de Kleene (cf., e.g., Epstein (1990), pp. 250-251). Por otra parte, tomando sólo los valores V y F, esta negación se comporta igual a la negación bivalente clásica.

La disyunción y la conjunción que pueden tomar cinco valores son definidas por las siguientes matrices:

$\varphi\vee_5\psi$	V	J	D	I	F
V	V	V	V	V	V
J	V	J	J	J	J
D	V	J	D	D	D
I	V	J	D	I	I
F	V	J	D	I	F

$\varphi\wedge_5\psi$	V	J	D	I	F
V	V	J	D	I	F
J	J	J	D	I	F
D	D	D	D	I	F
I	I	I	I	I	F
F	F	F	F	F	F

A partir de la negación y la disyunción o la conjunción definidas, podemos definir también una implicación pentavalente $(\neg_5\varphi\vee_5\psi) =_{df} (\varphi\rightarrow_5\psi) =_{df} \neg_5(\varphi\wedge_5\neg\psi)$, cuya matriz es la que sigue:

$\varphi\rightarrow_5\psi$	V	J	D	I	F
V	V	J	D	I	F
J	V	J	D	I	I
D	V	J	D	D	D
I	V	J	J	J	J
F	V	V	V	V	V

Como ocurre con la negación, estos operadores son también análogos a sus respectivos en el sistema de Kleene.

Dado que hemos dado al valor J a un argumento p-justificado (según un proceso determinado) para la sentencia que se considere, podemos interpretar que el argumento que justifica a la implicación $\rightarrow_5$ es el mismo que justifica a su consecuente, o bien el que justifica la negación de su antecedente.

En un sistema lógico con estos operadores, al igual que en la lógica de Kleene, no hay tautologías (i.e., ninguna fórmula tiene valor distinguido para cualquier asignación de valores a sus variables).

4. Lógica pentavalente *L5*: inferencias

Una cuestión interesante es qué inferencias son válidas en una lógica *L5* construida con los conectivos $\neg_5$, $\vee_5$, $\wedge_5$ y $\rightarrow_5$. Consideremos que en una lógica polivalente un razonamiento es válido siempre que al tomar valores distinguidos las premisas, también toma un valor distinguido la conclusión. Entonces, tomando como valores distinguidos *Dist* ={V, J, D}, hay reglas de inferencia del CP clásico que no son válidas en la lógica que postulamos. Por ejemplo, el modus ponens: si φ tiene valor D y ψ tiene valor F, entonces $\varphi\rightarrow_5\psi$ tiene valor D; por lo tanto, de premisas con valor distinguido, en este caso φ y $\varphi\rightarrow_5\psi$, se obtendría la conclusión ψ con valor F.[5]

Por otra parte, si excluimos a D del conjunto de valores distinguidos, como dijimos más arriba, el sistema queda sin tautologías. En cambio, todas las reglas de inferencia de la lógica clásica pasan a ser válidas, puesto que haciendo corresponder a cada valor distinguido el valor '1' y a cada valor antidistinguido el valor '0' en las tablas de los conectivos, las matrices se identifican con las de sus respectivos bivalentes clásicos. Por lo tanto, cada vez que las premisas de esas reglas de inferencia tomen valores distinguidos, también tomará un valor así la conclusión.

Según el modo en que hemos definido los valores, es claro que para nuestros propósitos debemos tomar como distinguidos a V y J, y excluir a D. Así, en *L5* no habrá tautologías, pero serán válidas las mismas reglas de inferencia que en el CP clásico. Simbolizaremos con '$\vdash_L$' y con '$\vdash_{L5}$' a las relaciones de consecuencia lógica en *L* y en *L5*, respectivamente.

5. Construcción de un sistema axiomático en *L5* a partir de un sistema argumentativo

En esta sección mostraremos cómo construir un sistema axiomático con el lenguaje de *L5* de modo que puedan derivarse en él todas las proposiciones verdaderas o que están p-justificadas en un sistema argumentativo *SA*. Para esto definiremos inductivamente una función para expresar en el lenguaje de *L5* todas las fórmulas de *L*; esta función será utilizada luego para expresar como axiomas las creencias expresadas en *SA*.

DEFINICIÓN 3

Sea t: $L \rightarrow L5$ una función inyectiva tal que

1 para cada letra proposicional p en L, $t(p)$ es una letra proposicional en $L5$;

2 para cualesquiera fórmulas α, β en L:

$$t(\neg\alpha) =_{df} \neg_5 t(\alpha)$$
$$t(\alpha\vee\beta) =_{df} t(\alpha)\vee_5 t(\beta)$$
$$t(\alpha\wedge\beta) =_{df} t(\alpha)\wedge_5 t(\beta)$$
$$t(\alpha\rightarrow\beta) =_{df} t(\alpha)\rightarrow_5 t(\beta)$$

Antes de continuar necesitaremos de una noción auxiliar más: la de 'activador' de un argumento rebatible. La idea es identificar la mínima información que se requiere en el contexto de un sistema argumentativo para que un argumento rebatible derive su conclusión. Esto involucra considerar la semántica del contexto, para lo cual será suficiente la noción de *modelo* en lógica clásica.

DEFINICIÓN 4

Sea T un argumento rebatible para φ en un sistema argumentativo $SA=\langle K, D, Args\rangle$; y sea $f(T)$ el conjunto de todas las fbf's de L que ocurren en las reglas rebatibles que forman T. Llamamos *activador* de T a cualquier subconjunto $m\subseteq f(T)$ mínimo tal que

1 algún modelo de K es un modelo de m;

2 $(m \cup T) \mid\!\sim \varphi$.

DEFINICIÓN 5

Sea $SA = \langle K, D, Args\rangle$ un sistema argumentativo. Llamaremos 'lógica pentavalente asociada a *SA*', y lo denotaremos con 'SA_{L5}', al sistema axiomático construido en *L5* a partir de *SA* del siguiente modo:

a) Conjunto de axiomas:

1 para cualquier fórmula $\varphi\in L$, si $\varphi\in K$ entonces $t(\varphi)$ se introduce como axioma en SA_{L5};

2 si un argumento T para una proposición ψ está p-justificado en SA, entonces, siendo $ACT(T)=\{A_1, ..., A_n\}$ el conjunto de activadores de T, se introducen como axiomas en SA_{L5}

una serie de n condicionales $t(\varphi_1)\rightarrow_5 t(\psi)$, ..., $t(\varphi_n)\rightarrow_5 t(\psi)$, donde cada φ_i es la conjunción de todas las fórmulas del activador A_i de T.

3 no hay en SA_{L5} otros axiomas que no sean los introducidos por los ítems 1 ó 2.

b) Reglas de inferencia: las válidas en la lógica clásica.

EJEMPLO 2

Sea SA un sistema argumentativo cuyo contexto de creencias es $K=\{r, r\rightarrow p\}$ y su conjunto de reglas rebatibles es $D=\{p\Rightarrow q, r\Rightarrow\neg q\}$. Sea el argumento $\{r\Rightarrow\neg q\}$ para q el único argumento justificado de SA. Ahora sea SA_{L5} el sistema axiomático lógico pentavalente asociado a SA. Entonces identificamos este sistema con el conjunto de axiomas $SA_{L5} = \{t(r), t(r)\rightarrow_5 t(p), t(r)\rightarrow_5\neg_5 t(q)\}$.

La idea de expresar a los argumentos rebatibles justificados en $L5$ como condicionales se basa en la conveniencia que esto representaría en caso de expandir el sistema argumentativo con la adición de nuevos argumentos. Si, en ese caso, ciertos argumentos antes justificados dejaran de estarlo, para expresar esos cambios en el sistema axiomático bastaría con remover los condicionales correspondientes del conjunto de axiomas y, eventualmente, reemplazarlos por otros que representen a los nuevos argumentos justificados.

Los siguientes resultados[6] muestran que el sistema axiomático resultante mediante el procedimiento descripto es *completo* y *correcto* respecto de las fórmulas p-justificadas en el sistema argumentativo al que se asocia, cuando la lógica L subyacente a éste es clásica.

LEMA 1

Si $K\vdash_L \varphi$ ó φ está p-justificado en SA entonces $SA_{L5} \vdash_{L5} t(\varphi)$.

LEMA 2

Si $SA_{L5} \vdash_{L5} \alpha$ entonces, para algún $\varphi \in L$ tal que $\{t(\varphi)\}\vdash_{L5} \alpha$ en $L5$, $K\vdash_L\varphi$ ó φ está p-justificado en SA.

TEOREMA 1 (completitud fuerte)

Para cualquier sistema axiomático SA_{L5} asociado a un sistema argumentativo SA, se cumple $SA_{L5}\vdash \varphi$ si y sólo si $K\vdash_L\varphi$ ó φ está p-justificado en SA.

6. Conclusiones

La construcción de un sistema axiomático del modo visto aquí permite expresar el conocimiento rebatible de un sistema argumentativo, y puede ser de gran utilidad para sistemas de IA donde el conocimiento representado permanece estático. Las implementaciones de sistemas argumentativos han mostrado en general problemas de eficiencia computacional debido a la complejidad de los procesos de búsqueda de justificaciones. Si un sistema argumentativo está destinado a no actualizar su base de conocimiento, entonces cada argumento que se declare justificado podría ser transformado en axioma de un sistema del tipo de SA_{L5}. De este modo, tal sistema axiomático (expresado en algún lenguaje declarativo como PROLOG) constituiría el primer recurso de consulta a la base de conocimiento, aprovechando las ventajas computacionales que la programación en lógica deductiva presenta respecto de los complejos procesos de los sistemas argumentativos. Asimismo, el sistema permitiría encontrar consecuencias deductivas del conocimiento tentativo justificado sin necesidad de activar los procesos dialécticos cada vez que se realiza una nueva consulta.

Notas

[1] Una versión previa de este trabajo será publicada en *Cuadernos del Sur-Filosofía* (Ediuns) como nota de avance con el título "Interpretación de un sistema lógico pentavalente a partir de un sistema de argumentación rebatible".

[2] Puesto que en el sistema las reglas pueden ser del lenguaje o del metalenguaje, se requiere que las reglas de inferencia adicionales operen en ambos niveles. Cf., e.g., el sistema de Simari y Loui (1992), donde puede utilizarse modus ponens e instanciación tanto en el lenguaje como en el metalenguaje.

[3] La cuestión de cuáles argumentos se consideran contraargumentos es relativa al diseño de cada sistema, razón por la cual consideraremos preestablecida a esta relación.

[4] En el sistema que aquí presentamos esto mismo se hace prohibiendo directamente la formación de argumentos rebatibles que contradigan consecuencias lógicas de la base de creencias (cf. la DEFINICIÓN 2, cláusula 2).

[5] Nótese que, sin embargo, la fórmula $(\varphi\wedge_5(\varphi\rightarrow_5\psi))\rightarrow_5\psi$ toma valor distinguido (D) bajo esos mismos valores.

[6] Las demostraciones no se han incluido por razones de espacio y pueden ser solicitadas al autor.

Referencias

1 Dung, P. M. (1995); "On the acceptability of arguments and its fundamental role in nonmonotonic reasoning, logic programming and *n*-person games"; *Artificial Intelligence* 77, 321-357.

2 Epstein, R. (1990); *The semantic foundations of logic. Volume 1: Propositional Logics*, Kluwer Academic Publishers, Dordrecht.

3 Kowalski, R. y F. Toni (1996); "Abstract argumentation"; *Artificial Intelligence and Law* 4 (3-4), 275-296.

4 Loui, R. (1987); "Defeat Among Arguments: A System of Defeasible Inference"; *Computational Intelligence* 2, 100-106.

5 Poole, D. (1985); "On The Comparison of Theories: Preferring The Most Specific Explanation"; *Proc. of the Ninth IJCAI*, Los Altos (144-147).

6 Prakken, H. (1993); *Logical tools for modeling legal arguments*, Ph.D Thesis, Vrije Universiteit, Amsterdam.

7 Rescher, N. (1977); *Dialectics: a controversy oriented approach to the theory of knowledge*, State University of New York Press, Albany.

8 Simari, G y R. Loui (1992); "A mathematical treatment of defesible reasoning", *Artificial Intelligence* 53, 125-157

9 Toulmin, S. (1958) *The uses of argument*, Cambridge University Press, London, New York.