EPISTEMOLOGÍA E HISTORIA DE LA CIENCIA

SELECCIÓN DE TRABAJOS DE LAS XV JORNADAS

VOLUMEN 11 (2005)

TOMO I

Horacio Faas
Aarón Saal
Marisa Velasco
Editores

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

Sobre la credulidad en la aceptación de argumentos

*Gustavo A. Bodanza**

1. Introducción

Un marco argumentativo es una idealización de un debate en el que se tienen en cuenta sólo los argumentos participantes y la relación entre éstos dada cuando, según las reglas que determine el debate, un argumento ataca a otro. La finalidad de tales modelos es investigar qué argumentos deberían resultar ganadores en un debate dado, una vez establecidos los pares de argumentos que están en relación de ataque. A los conjuntos de argumentos ganadores se los llama 'extensiones', y se los puede definir de distintas maneras según distintos criterios. Según algunos, todo marco argumentativo debería tener a lo sumo una extensión. Una extensión obtenida según este criterio será una extensión *escéptica*. Si, por el contrario, en algunos marcos argumentativos se puede obtener más de una extensión, éstas se consideran *crédulas*. La interpretación que puede darse a esto es que una extensión escéptica contiene sólo los argumentos indudablemente defendibles en el marco dado, mientras las extensiones crédulas contienen argumentos que son dudosamente defendibles (porque tienen atacantes), pero también son dudosamente indefendibles (porque a sus vez sus atacantes son dudosamente defendibles). Ahora bien, algunas extensiones crédulas son definidas de tal modo que en ciertos marcos argumentativos no logran dar como ganadores algunos argumentos que son de esperar que un agente crédulo acepte, o sea, dan un resultado más "escéptico" de lo esperado. Nos referimos en particular a las extensiones *preferidas* de P.M. Dung (1995). En este trabajo ensayamos una noción de extensión alternativa, que llamamos 'tolerante', intentando obtener el comportamiento crédulo esperado. Para esto analizaremos algunos de los ejemplos conflictivos, mostraremos cómo las extensiones tolerantes funcionan del modo esperado en éstos, y demostraremos la existencia de tales extensiones para todo marco argumentativo.

2. La teoría argumentativa abstracta de Dung

Puesto que el trabajo estará enteramente basado en la teoría argumentativa dada por Dung (*op. cit.*), la formalización de las nociones de las que partimos serán las definidas por este autor.

Definición 1

(Dung). Un *marco argumentativo* es un par $AF = \langle AR, ataca\rangle$, donde AR es un conjunto de entidades llamadas *argumentos*, y *ataca* es una relación binaria sobre AR, es decir, $ataca \subseteq AR \times AR$. En adelante, cuando hablemos de $AF = \langle AR, ataca\rangle$ nos referiremos a un marco argumentativo genérico, salvo que se aclare otra cosa.

La relación *ataca* es arbitraria y no se supone, *prima facie*, que cumpla ninguna propiedad especial. Observemos sólo que tal relación no puede ser transitiva, ya

* Universidad Nacional del Sur CONICET

ccbodanz@criba.edu.ar

que si tenemos una cadena de argumentos A_1, A_2, A_3, tales que $(A_1, A_2), (A_2, A_3) \in$ *ataca*, entonces intuitivamente no podemos decir que esto implique A_1 *ataca* A_3, antes bien, por el contrario, parece intuitivo que A_1 "defiende" A_3. Esto no impide, de todos modos, que en algunos marcos argumentativos se den a la vez $(A_1, A_2), (A_2, A_3), (A_1, A_3) \in$ *ataca*. Si bien se han estudiado en profundidad mecanismos para decidir qué argumentos resultarían ganadores dada una relación de rebatimiento determinada, sólo se han hecho algunos estudios sobre las propiedades que debería cumplir una relación así (*e. g.* Pollock (1990, 1994)).

La teoría de Dung gira en torno a dos nociones básicas: la de "aceptabilidad" de un argumento y la de "admisibilidad" de un conjunto de argumentos. Éstas permitirán luego representar distintas ideas sobre cuáles se consideran los argumentos ganadores de un marco, definiendo distintos tipos de "extensiones". A partir de entonces podremos desarrollar nuestra crítica y la subsiguiente propuesta.

Definición 2

(Dung, 1995) Un argumento A es *aceptable* con respecto a un conjunto de argumentos S, si y sólo si $\forall B$ (B *ataca* $A \rightarrow \exists C$ ($C \in S$ & C *ataca* B)). Un conjunto de argumentos S es *admisible* si y sólo si:

1. $\forall A\ \forall B\ (A,B \in S \rightarrow \neg(A$ *ataca* $B))$ (S está libre de conflictos); y
2. todo argumento perteneciente a S es aceptable con respecto a S.

Para referirse al conjunto de argumentos que resultan "ganadores" en un marco, Dung usa la palabra *extension* (traduciremos literalmente 'extensión'). Podría haber distintos criterios acerca de cuáles argumentos deben ser los ganadores dado un marco determinado, por lo que nuestro autor introduce distintas nociones de 'extensión'. Enfocaremos sobre la que nuestro autor llama 'extensión preferida' (*preferred extension*), que busca capturar un comportamiento crédulo, en el sentido que, frente a conflictos irresolubles unívocamente entre argumentos, ofrece distintas alternativas excluyentes pero igualmente defendibles.

Definición 3

(Dung, 1995) Un conjunto de argumentos S es una *extensión preferida* de AF si y sólo si S es un conjunto máximamente (c.r. a la inclusión conjuntista) admisible de argumentos de AF.

Ejemplo 1

Sea $AF = \langle AR, ataca \rangle$ tal que $AR = \{A, B\}$ y $ataca = \{(A, B), (B, A)\}$. Entonces $\{A\}$ y $\{B\}$ son ambas extensiones preferidas. Este ejemplo, conocido como "Diamante[1] de Nixon", se puede entender con la siguiente interpretación: A: 'Nixon es pacifista puesto que es cuákero', B: 'Nixon no es pacifista puesto que es republicano'. Las dos extensiones representan sendas posiciones en el debate que, si bien no hay un modo imparcial de determinar una ganadora entre ambas, cualquiera de ellas es razonablemente admisible. Un comportamiento escéptico, en cambio, sería no admitir a ninguna.

3. Vindicación y tolerancia

Las extensiones preferidas capturan bien el comportamiento crédulo en muchos marcos argumentativos, como el del último ejemplo, pero no así en otros.

Ejemplo 2

Sea $AF = \langle\{A, B\}, \{(A, A), (A, B)\}\rangle$. Aquí B es atacado por A, pero A se ataca a sí mismo -podríamos entender a A como un argumento paradójico. La literatura coincide en general en que B debería resultar ganador, ya que el auto-ataque de A hace "impertinente" el ataque sobre B. Nótese, sin embargo, que $\{B\}$ no es una extensión preferida.

Ejemplo 3

Sea $AF = \langle\{A, B, C\}, \{(A, B), (B, C), (C, A)\}\rangle$. La única extensión preferida es $\varnothing$. En este ejemplo podría esperarse que $\{A\}$, $\{B\}$ y $\{C\}$ resulten extensiones crédulas de algún tipo, dada la ronda de ataques que hace imposible determinar un ganador[2]. La intuición se puede capturar imaginando los siguientes argumentos, ubicados en un imaginario contexto de discusión científica renacentista:

C: 'Tycho tiene razón puesto que es evidente que el Sol gira alrededor de la Tierra, luego Copérnico no tiene razón',

B: 'Ptolomeo tiene razón puesto que el Sol gira alrededor de la Tierra y los planetas también, por lo tanto Tycho no tiene razón';

A. 'Copérnico tiene razón pues en realidad la Tierra y los otros planetas giran alrededor del Sol, luego Ptolomeo no tiene razón'.

Si bien estos argumentos parecen muy artificiales (al igual que el debate mismo, donde se presentan estos tres argumentos y sólo estos tres), podemos entender aun así que un astrónomo renacentista no rechazaría los tres sino que optaría por uno, del mismo modo que nadie rechazaría las tres teorías desde las que son enunciados por el hecho de no ser una claramente mejor que las otras, sino que adheriría a alguna de ellas (en este sentido, un científico cualquiera parece ser crédulo antes que escéptico).

Parece claro que el problema que se presenta en estos ejemplos tiene que ver con los ciclos impares de ataques (como vimos más arriba en el "diamante de Nixon", un ciclo par no presenta estos problemas). Vamos a proponer que una solución adecuada puede hallarse usando la misma noción de extensión preferida pero de un modo distinto.

La propuesta tiene dos pasos. en primer lugar construimos una *relación* de preferencia en base a la noción de extensión preferida. Puesto que una extensión preferida representa un subconjunto máximo de argumentos que pueden defenderse entre sí, la idea es confrontar unos subconjuntos de argumentos con otros. vamos a tomar aquellos subconjuntos S que pueden defenderse de cualquier subconjunto S' que a su vez pueda defenderse contra S. Lo que estamos haciendo es relativizar la noción de extensión preferida al submarco argumentativo en el que solamente se dan los argumentos de S y de S'. Sea $AF|_S = \langle S, \mathit{ataca}|_S\rangle$, donde $S \subseteq AR$ y $\mathit{ataca}|_S$ es la restricción de la relación *ataca* al subconjunto S de AR. Entonces,

Definición 4
Para cualesquiera conjuntos de argumentos S, $S' \subseteq AR$, decimos que S es *al menos tan preferido como* S' si y sólo si S es una extensión preferida en $AF\,|_{S \cup S'}$.

Nótese que esta relación es reflexiva pero no necesariamente transitiva.

Definición 5
Un conjunto de argumentos S es una *posición vindicada* si y sólo si para todo conjunto de argumentos S', si S' es al menos tan preferido como S entonces S es al menos tan preferido como S'.

En el Ejemplo 2 podemos ver que $\{B\}$ es una posición vindicada, ya que ni $\varnothing$ ni $\{A\}$ son al menos tan preferidas como $\{B\}$.

Para dar el segundo paso, observemos primero el Ejemplo 3. Ni $\{A\}$ ni $\{B\}$ ni $\{C\}$ son posiciones vindicadas, puesto que $\{A\}$ es tan preferida como $\{B\}$ pero no viceversa, y lo mismo se cumple entre $\{B\}$ y $\{C\}$, y entre $\{C\}$ y $\{A\}$. Ahora bien, podemos ver también el problema desde un punto de vista estratégico. Al considerar si podemos sostener $\{A\}$ veremos que $\{C\}$ es preferible, luego sostendríamos $\{C\}$ en vez de $\{A\}$; pero al considerar $\{C\}$ veremos que $\{B\}$ es preferible, luego sostendríamos $\{B\}$ en vez de $\{C\}$; y al considerar $\{B\}$ veremos que $\{A\}$ es preferible, luego sostendríamos $\{A\}$ en vez de $\{B\}$. Es decir, siguiendo la estrategia de cambiar la postura argumentativa por aquellas que se presenten como mejores volvemos siempre a la misma. Entonces, no carece de sentido que, adoptando un criterio algo laxo, decidamos adoptar indistintamente cualquiera de las posturas que aparecen recurrentemente en el proceso de revisión. Sería como decir "quizás no podamos defender nuestros argumentos sobre una base sólida pero, por alguna razón, nos vemos reconsiderándolos una y otra vez; luego, parecen argumentos razonablemente aceptables". Podemos caracterizar esta actitud como *tolerante*, ya que estamos dando curso a ciertos argumentos que no podemos aprobar estrictamente.

Definición 6
Un conjunto de argumentos $S \subseteq AR$ es una extensión *tolerante* en AF si y sólo si para todo $S' \subseteq AR$, si S' es al menos tan preferido como S, entonces existe una secuencia $S_1, \ldots, S_n \subseteq AR$ tal que $S_1{=}S'$, $S_n{=}S$ y S_{k+1} es al menos tan preferido como S_k, donde $1 \leq k < n$.

Dada esta definición obtenemos que:

Lema 1
Toda posición vindicada es una extensión tolerante.

Prueba. Se hace inmediato si rescribimos la definición de 'posición vindicada' del siguiente modo: para cualquier posición vindicada A y para cualquier subconjunto de argumentos B al menos tan preferido como A existe una secuencia S_1, S_2, tal que $S_1{=}A$, $S_2{=}B$ y S_{k+1} es al menos tan preferido como S_k, donde $1 \leq k < 2$. *QED*.

La noción de extensión tolerante, siendo más general que la de posición vindicada como acabamos de ver, nos permite tratar como esperábamos tanto el Ejemplo 2 como el Ejemplo 3.

Observación 1
La noción de posición vindicada no está definida para todo marco argumentativo.

Ejemplo 4
Ejemplo 3 *revisitado.* Sea $AF = \langle\{A, B, C\}, \{(A, B), (B, C), (C, A)\}\rangle$. Entonces $\{A\}$, $\{B\}$ y $\{C\}$ son todas extensiones tolerantes pero no existe posición vindicada.

Ejemplo 5
Diamante de Nixon revisitado. En $AF = \langle\{A, B\}, \{(A,B), (B,A)\}\rangle$, tanto $\{A\}$ como $\{B\}$ son a la vez posiciones vindicadas, extensiones preferidas y extensiones tolerantes.

Ejemplo 6
Sea $AF = \langle\{A, B, C\}, \{(A, A), (A, B), (B, C)\}\rangle$. Entonces $\{B\}$ es una extensión tolerante. Nótese que $\{B\}$ es al menos tan preferida como $\varnothing$, $\{A\}$, $\{C\}$ y $\{A, C\}$ (además del resto de subconjuntos distintos de $\{B\}$, todos internamente conflictivos) pero no viceversa.

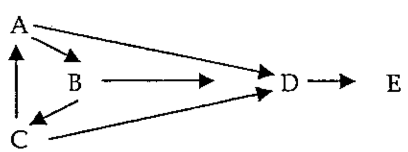
Teorema 1
Todo marco argumentativo finito posee una extensión tolerante.

Prueba. Supongamos lo contrario, entonces para todo subconjunto de argumentos S se cumple que (a) existe un subconjunto S' al menos tan preferido como S, y (b) no existe una secuencia $S_1, \ldots, S_n \subseteq AR$ tal que $S_1=S'$, $S_n=S$ y S_{k+1} es al menos tan preferido como S_k, donde $1 \leq k < n$. Pero como esto es cierto para todo subconjunto S y AF es finito, aun cuando se agoten todos los subconjuntos del marco siguiendo la serie de preferencias, aquel subconjunto que sea al menos tan preferido como el último no repetido debe haber aparecido antes en la serie. De este modo se cierra un ciclo de preferencias, contradiciendo a (b). Luego, todo marco argumentativo posee una extensión tolerante. *QED*.

4. Justificación crédula

El siguiente ejemplo nos permitirá entender la noción que buscamos de *justificación* desde un punto de vista crédulo.

Ejemplo 7
Argumento "flotante". Sea $AF = \langle\{A, B, C, D, E\}, \{(A, B), (B, C), (C, A), (A, D), (B, D), (C, D), (D, E)\}\rangle$. (Ver el gráfico de abajo, donde las flechas indican los ataques.) Entonces $\{A, E\}$, $\{B, E\}$ y $\{C, E\}$ son las extensiones tolerantes puesto que cada subconjunto en la secuencia $\{A, E\}$, $\{B, E\}$, $\{C, E\}$, $\{A, E\}$ es al menos tan preferido como su sucesor y no existe ningún otro subconjunto $S \subseteq AR$ distinto de esos tres que sea al menos tan preferido como cualquiera de ellos.



Siendo que E (el cual "flota" sobre D gracias a A, B y C) pertenece a todas las extensiones tolerantes, decimos que E está *crédulamente justificado* en AF. En general,

Definición 7
Un argumento A está *crédulamente justificado* en AF si y sólo si $A \in S$ para todo conjunto de argumentos S que sea una extensión tolerante de AF.

Es claro que, si bien todo marco argumentativo posee al menos una extensión tolerante, no todos ellos poseen argumentos justificados (cf., Ejemplo 1 y Ejemplo 3).

5. Conclusiones

A través de la noción de 'extensión tolerante' hemos ofrecido una solución al tipo de problemas que se le plantean a la noción de 'extensión preferida' como modelo de comportamiento argumentativo crédulo. También hemos mostrado que todo marco argumentativo posee al menos una extensión tolerante.

Cabe decir que otros autores han considerado soluciones distintas, en algunos casos con resultados similares en cuanto al comportamiento (Baroni-Giacomin (2003)), y en otros directamente se han seguido criterios absolutamente distintos a los nuestros (*e. g.* Jakobovitz-Vermeir (1999)). La discusión de estos enfoques la reservaremos para un trabajo más detallado.

Notas

[1] El término 'diamante' refiere a la forma de cierta representación gráfica de ese marco argumentativo.

[2] Dung, como otros autores, no está de acuerdo en este punto porque, sostiene, un ciclo impar de ataques hace que un mismo argumento sea a la vez atacante y defensor *indirecto* de otro (en el ejemplo, A se defiende indirectamente de C porque ataca al atacante de éste, B, pero por la misma razón también se ataca indirectamente a sí mismo ya que B es su defensor). Aunque esta opinión es muy atendible, no la discutiremos aquí por razones de espacio.

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(2):321–358.

2. Baroni, P. y M. Giacomin (2003). "Solving semantic problems with odd-length cycles in argumentation". En *Proceedings of the 7th European Conference on Symbolic and Quantitative Approaches to Reasoning with Uncertainty (ECSQARU 2003)*, Aalborg, Denmark. LNAI 2711, Springer-Verlag, 440–451

3. Jakobovitz, H. y D. Vermeir (1999). "Dialectic Semantics for Argumentation Frameworks". En *Proc. ICAIL'99*, 53–62. ACM Press.

4. Pollock, J. (1990). *Nomic probability and the foundations of induction*. Oxford University Press, New York-Oxford.

5. Pollock, J (1994) "Justification and defeat". *Artificial Intelligence* 67:377–407