

EPISTEMOLOGÍA E HISTORIA DE LA CIENCIA

SELECCIÓN DE TRABAJOS DE LAS XXII JORNADAS

VOLUMEN 18 (2012)

Luis Salvatico
Maximiliano Bozzoli
Luciana Presenti

Editores



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



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



Aceptabilidad relativa para marcos argumentativos abstractos

Claudio Andrés Alessio *

Introducción

Los sistemas argumentativos son formalismos que modelan información rebatible y potencialmente contradictoria. En general pueden ser caracterizados en un proceso de tres etapas. Primero se genera un marco argumentativo a partir de una base de conocimiento, i.e. se construye un conjunto de argumentos y se establecen relaciones de derrota a partir de la identificación de los argumentos que son mutuamente excluyentes junto a algún(os) criterio(s) de preferencia. Luego se determina el conjunto de argumentos aceptados (que permiten extender la base de datos inicial) previo cálculo de aceptabilidad a partir del marco definido en el paso anterior. Finalmente se identifica el conjunto de conclusiones justificadas, en general, son las conclusiones de los argumentos aceptables.

La noción central, como es evidente, en tales sistemas es la de argumento aceptable. Intuitivamente, un argumento será considerado aceptable cuando, para cualquier argumento que lo derrota existe algún argumento que lo defiende. Pueden existir varios criterios para establecer qué argumento será considerado aceptable. Tales criterios son denominados *semánticas*. Las mismas establecen diversas condiciones con respecto a diversos niveles de credulidad o escepticismo que un conjunto de argumentos debe satisfacer.

La mayoría de los estudios en relación a la aceptabilidad de argumentos se realizan a partir de lo que se conoce como *Marco Argumentativo Estándar* (MAE) propuesto en (2). MAE es una teoría de argumentación formal que se encuentra definida a partir de dos nociones primitivas: *argumento* y *derrota*. Nada se supone acerca de la estructura interna de los argumentos, aunque es claro que son entidades que pueden derrotar o ser derrotadas por otras entidades del mismo tipo. La relación de derrota, por otro lado, es simplemente una relación binaria entre pares de argumentos.

Un conjunto de argumentos aceptables en MAE es identificado a partir de las llamadas *semánticas* (preferida, estable, básica y completa) que tienen en cuenta las relaciones de derrota para asignar el estado de argumento aceptable. En general, dos condiciones elementales debe satisfacer un conjunto para garantizar que los argumentos que pertenecen a él pueden constituir *extensiones* de MAE: defender a cada argumento que pertenece a él y ser consistente, i.e. no hay pares de argumentos, en tal conjunto, que verifican la relación de derrota. Además de la defensa conjunta y la coherencia, las *semánticas* exigen otras condiciones a un conjunto. Por ejemplo, la *semántica* preferida exige que sea maximal con respecto a esas propiedades. La *semántica* estable requiere que un conjunto sea coherente y derrote a cada argumento que no pertenezca a él. La *semántica* completa demanda además de la coherencia y la autodefensa, que el conjunto incluya cada argumento defendido por él. Finalmente, la *semántica* básica (*grounded*) exige que el conjunto sea construido a partir de argumentos aceptables con respecto al conjunto vacío.

El nivel de abstracción de MAE permite que sea un modelo interesante para estudiar propiedades (no sólo de sistemas basados en argumento) de diferentes sistemas formales que modelan información rebatible y potencialmente contradictoria. La *semántica* estable de MAE captura la *semántica* de los principales enfoques para razonamiento no monotónico

* UCCuyo – CONICET, claudioalessio@uccuyo.edu.ar

tales como las extensiones de la Lógica Default, expansiones estables de Lógica Autoepistémica, modelos estables de Programación Lógica y la semántica crédula de Redes de Herencia. Por otro lado, la semántica básica permite definir la Lógica Rebatible de Pollock y las semánticas bien comportadas de Programación Lógica,

A pesar de su gran potencial es posible identificar ejemplos que parecen poner en duda algunos de los aspectos de MAE. En particular (3) presenta el siguiente ejemplo que aparentemente sustenta esta afirmación. Con tal ejemplo es posible observar que la aceptabilidad es afectada por el tipo de derrota y defensa que un argumento recibe.

Ejemplo 1: Supóngase que por lo general los abogados son adinerados, pero cierta subclase de abogados, los defensores públicos, tienden a no serlo. Imagínese que existe una ciudad llamada Brentwood, que contiene un amplio número de casas muy caras pero donde existe una pequeña cantidad de propiedades para rentar a precio muy bajo, en consecuencia es posible afirmar que por lo general, los residentes de Brentwood, tienden a ser adinerados, sin embargo, una subclase de los habitantes de Brentwood, los que rentan, tienden a no serlo. Supóngase que Ana es una defensora pública que renta en Brentwood y por lo tanto, una abogada residente de Brentwood. A partir de tal información es posible construir los siguientes argumentos:

A: Ana en tanto defensora pública es abogada, y dado que por lo general los abogados son adinerados, Ana es adinerada.

B: Ana en tanto que renta en Brentwood y dado que por lo general quienes rentan en Brentwood no son adinerados, Ana no es adinerada.

C: Ana en tanto que renta en Brentwood es residente de Brentwood y dado que por lo general los residentes de Brentwood son adinerados, Ana es adinerada.

D: Ana en tanto defensora pública, y dado que por lo general los defensores públicos no son adinerados, Ana no es adinerada.

Si las relaciones de derrota son interpretadas como es usual (conflicto más preferencia) es claro que el argumento D derrota estrictamente a A al igual que B con respecto a C. Por otro lado, los argumentos A,B y D,C se derrotan mutuamente. A partir de estas relaciones es posible determinar en MAE que $\{B,D\}$ por un lado y $\{C,A\}$ por otro son conjuntos coherentes que se autodefenden. De modo que crédulamente ambos conjuntos serán extensiones de MAE (o ninguno desde el punto de vista escéptico) contrario a la intuición, dado que es esperable que sólo $\{B,D\}$ sea considerado como un conjunto de argumentos aceptables.

Un problema similar fue identificado en (1) en el contexto de marcos argumentativos basados en preferencia. Según (1), las semánticas de MAE cuando son aplicadas a toma de decisiones pueden llevar a resultado contraintuitivos y fallar en la obtención de las conclusiones correctas. En particular el problema aparece porque la semántica estable considera, en el ejemplo 1, como igualmente buenos a ambos conjuntos a pesar de que para cualquier argumento del segundo $\{C,A\}$, el primero contiene argumentos estrictamente preferidos $\{B,D\}$.

Aunque es cierto que pueda cuestionarse el hecho de que el ejemplo se encuentra correctamente interpretado (Cfr 6), permite ilustrar un problema que puede generalizarse de

la siguiente manera: dadas dos extensiones E y E' de MAE, cada argumento de E' es derrotado asimétricamente por al menos un argumento en E . En términos de preferencia, siguiendo a (1) el mismo problema puede definirse como aquel que aparece cuando para cada argumento de E' , existe al menos un argumento de E tal que es preferido estrictamente. Parecería que el origen de tal problema se encuentra en el hecho de que el tipo de derrota, a saber, si es simétrica o asimétrica, afectará la aceptabilidad de un argumento, contrario a lo que se define en MAE. Tal situación motiva la exploración de nuevas alternativas que tengan en cuenta tales casos y den un resultado diferente a los que se obtiene en las semánticas tradicionales de MAE.

El mejor resultado que podría obtenerse sería la definición de una nueva semántica a fin de no perder las conexiones con otros formalismos definidas a partir de las semánticas tradicionales. Sin embargo no se ha podido definir correctamente tal mecanismo aunque parecería posible su identificación a partir de la semántica completa. En el ejemplo, esta semántica da como extensiones a los siguientes: $\{ \}$, $\{B\}$, $\{D\}$, $\{B,D\}$, $\{A,C\}$. $\{B,D\}$ es una especie de extensión "super-completa" donde cada subconjunto es una extensión completa. Por otro lado en $\{A,C\}$ ni $\{A\}$, ni $\{C\}$ son extensiones completas. Por tanto una posible solución podría explorarse a partir de aquí.

En su lugar se propone una modificación de la noción de aceptabilidad, denominada aceptabilidad relativa, y una proposición que fundamenta posibles conexiones entre un nuevo marco definido a partir de tal noción y MAE. Sin embargo la propuesta adolece de una dificultad que parece grave: Si se definiera un marco argumentativo a partir de la aceptabilidad relativa, las conexiones con la semántica estable se perderían. Esto hace pensar que es necesaria la identificación de una semántica como anteriormente se ha señalado.

El trabajo se organiza como sigue: en la sección 2 se presentan las nociones elementales de un marco argumentativo y se muestran diversos ejemplos generalizados a partir del ejemplo 1. En la sección 3 se presentan dos nociones de aceptabilidad una más fuerte denominada s-aceptabilidad y una más débil llamada r-aceptabilidad. Se evalúa la primera y se identifica un ejemplo clásico en el que el resultado que arrojaría sería indeseable. Posteriormente se presenta la noción de r-aceptabilidad destacando los resultados que obtiene y las dificultades que aparecerían en caso de definir un marco argumentativo a partir de la misma.

Marcos argumentativos

Los marcos argumentativos fueron propuestos en (2). Tales marcos constituyen un modelo de argumentación abstracto definido a partir de dos nociones primitivas abstractas "argumento" y "derrota" (o "ataque"). No existen supuestos sobre la estructura interna o partes de los argumentos y la derrota es simplemente una relación entre pares de argumentos.

A partir de los componentes enunciados, se puede determinar cuándo un argumento es considerado aceptable. La noción de aceptabilidad permite la definición de diferentes tipos de extensiones de un marco argumentativo, i.e. conjuntos de argumentos que conjuntamente se auto-defienden, de acuerdo con un criterio crédulo o escéptico.

Definición 1. Un marco argumentativo estandar es un par $MAE = \langle AR, \rightarrow \rangle$, donde AR es un conjunto de entidades llamadas argumentos y $\rightarrow \subseteq AR \times AR$ es una relación de derrota entre pares de argumentos. Un argumento A se dice *aceptable* con respecto a un subconjunto S de

argumentos de AR , si para todo argumento B tal que $B \rightarrow A$, $\exists C \in S$ tal que $C \rightarrow B$. Un conjunto de argumentos S se dice *admisibles* si cada $A \in S$ es aceptable con respecto a S , y es *libre de conflicto*, i.e. las relaciones de derrota no se dan para algún par de argumentos que pertenecen a S . Una *extensión preferida* es el conjunto admisible maximal de argumentos de MA . Una *extensión completa* de MA es algún subconjunto de argumentos libre de conflicto que es el punto fijo de $\Phi(\bullet)$, donde $\Phi(S) = \{A \mid A \text{ aceptable con respecto a } S\}$, mientras que la *extensión básica (grounded)* es la menor extensión completa (c.r. a \subseteq). Por otra parte, una *extensión estable* es un conjunto de argumentos S libre de conflicto que derrota a todos los argumentos que no pertenecen a S .

Ejemplo 2: (ejemplo 1 revisitado) Sea $\langle \{A, B, C, D\}, \{(A \rightleftharpoons B), (B \rightarrow C), (C \rightleftharpoons D), (D \rightarrow A)\} \rangle$ un marco argumentativo. Las extensiones preferida y estable son: $\{A, C\}$ y $\{B, D\}$ mientras que la extensión básica es $\{\}$. Este ejemplo puede ser ilustrado en la figura 1. Cada nodo representa a un argumento. Las flechas de una dirección denotan derrotas asimétricas, y las flechas de ida y vuelta a las simétricas

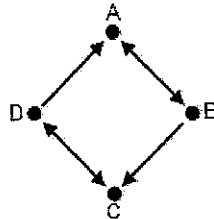


Fig 1

El anterior ejemplo es una traducción formal del ejemplo 1. Si se acepta tal modelación, MAE concluye que hay dos extensiones crédulas. Una para Ana es adinerada $\{A, C\}$ y otra para lo contrario $\{B, D\}$. Escépticamente ninguna de las conclusiones es creída dado que la menor extensión completa es $\{\}$. Sin embargo, intuitivamente es esperable que los argumentos aceptables sean únicamente C y D.

Aunque puede ser discutido el hecho de que el ejemplo 2 es la expresión formal del ejemplo 1, permite ilustrar una familia de ejemplos que presentan una característica que parece problemática y que al parecer no se trata sólo de un problema de representación. El problema se presenta cuando dadas dos extensiones E y E' de MAE, cada argumento de E' es derrotado asimétricamente por al menos un argumento de E . Las siguientes figuras ilustran varios de estos casos.

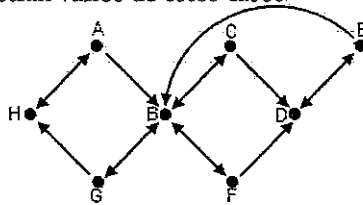


Fig. 2

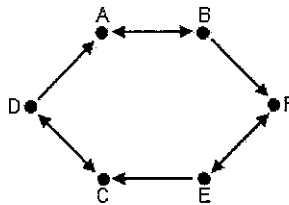


Fig. 4

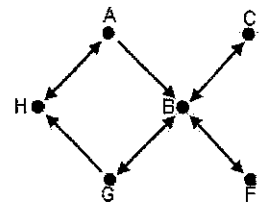


Fig. 3

En la figura 2, hay dos extensiones de MAE. $E = \{G, A, C, F, E\}$ y $E' = \{H, B, D\}$ donde cada argumento de E' es derrotado estrictamente por al menos un argumento de E . En la figura 3 ocurre lo mismo entre $\{D, B, E\}$ y $\{A, F, C\}$. La figura 4 es interesante porque muestra que el problema sigue apareciendo a pesar de que no se dé el hecho de que cada argumento de E derrote asimétricamente a por lo menos un argumento de E' , lo que sí ocurre en las figuras 1-3. En términos de preferencia el problema parece ser más claro aún, dado que aunque E y E' sean conjuntos de argumentos aceptables, E' es un conjunto de argumentos tal que para cada argumento de E' existe al menos un argumento en E que es preferido estrictamente. De modo que si se piensa a ambas extensiones en función de toma de decisiones, parecería que el conjunto E es preferible al conjunto E' y el agente no debería tomar la decisión de manera arbitraria, como es usual en casos de múltiples extensiones. En consecuencia, si una de las dos extensiones es preferible, parecería que la otra no debería ser una extensión.

Si bien es cierto que el problema parece frecuente, los ejemplos no alcanzan para invalidar el proceso de construcción y evaluación de argumentos de MAE tal como lo propone (3). Por el contrario, la aparición de contraejemplos debe verse como una buena ocasión para explorar nuevas alternativas en la modelación de tales sistemas. Por ello, en la sección siguiente se propone una vía de escape a la cuestión considerada. Esta se hará a partir de una noción de aceptabilidad que establece condiciones de defensa relativas al tipo de derrotas que afecta a un argumento. Aunque la propuesta resuelve el problema, los resultados no son óptimos.

Aceptabilidad relativa

El planteo central del presente trabajo consiste en identificar una estrategia que permita que los marcos argumentativos tengan un comportamiento adecuado frente a ejemplos similares a los de las figuras 1-4. Recientemente se han propuesto diversas nociones de aceptabilidad (Cfr 4-5) que permiten explorar mecanismos para considerar y evaluar la aceptabilidad de los argumentos. Una estrategia similar puede seguirse aquí. Retomando el problema es posible identificar a la derrota asimétrica como la clave para que se seleccionen adecuadamente a los argumentos. Una primera propuesta que considera este rol es la siguiente.

Definición 2. Un argumento $A \in AR$ será *s-aceptable* c.r. a un conjunto $S \subseteq AR$ si para todo argumento $B \in AR$ tal que B derrota a A , existe por lo menos un argumento $C \in S$ tal que C derrota asimétricamente a B .

Esta noción de aceptabilidad, aunque puede ser una buena candidata para resolver el problema identificado, parece demasiado exigente. Obsérvese lo que sucede en algunos ejemplos, aunque antes es preciso definir al menos la noción de conjunto admisible:

Definición 3: Un conjunto de argumentos libre de conflictos S será *s-admisible* si cada argumento en S es *s-aceptable* c.r. a S .

Ejemplo 3. Sea $\langle \{A, B, C, D\} \{ (A \Rightarrow B), (B \rightarrow C), (C \Rightarrow D), (D \rightarrow A) \} \rangle$ un marco argumentativo. Sólo el conjunto $\{B, D\}$ es *s-admisible*.

En los siguientes, por razones de brevedad se señalará el conjunto s-admisibles maximal c.r. a \subseteq

Ejemplo 4: Sea $\langle \{A,B,C,D,E,F,G,H\} \{ (A \rightarrow B), (B \rightleftharpoons C), (C \rightarrow D), (D \rightleftharpoons E), (E \rightarrow B), (B \rightleftharpoons G), (G \rightarrow H), (H \rightleftharpoons A) \} \rangle$ un marco argumentativo. El conjunto $\{A,C,E,F,G\}$ es s-admisibles maximal.

Ejemplo 5: Sea $\langle \{A,B,C,D,E,F\} \{ (A \rightleftharpoons B), (B \rightarrow F), (F \rightleftharpoons E), (E \rightarrow C), (C \rightleftharpoons D), (D \rightarrow A) \} \rangle$ un marco argumentativo $\{D,B,E\}$ es s-admisibles maximal.

Ejemplo 6: Sea $\langle \{A,B,C,F,G,H\} \{ (A \rightarrow B), (B \rightleftharpoons G), (B \rightleftharpoons C), (B \rightleftharpoons F), (G \rightarrow H), (H \rightleftharpoons A) \} \rangle$ un marco argumentativo $\{A,G,C,F\}$ es s-admisibles maximal.

El siguiente ejemplo es el marco argumentativo del clásico diamante de Nixon, como se verá la definición 2 hace que ninguno de los argumentos en cuestión sea aceptable.

Ejemplo 7: Sea $\langle \{A,B\} \{ (A \rightleftharpoons B) \} \rangle$ un marco argumentativo. El conjunto $\{ \}$ es s-admisibles maximal.

De modo que la definición 2 parece demasiado exigente para modelar adecuadamente razonamiento default o por lo menos la versión crédula. Ello sucede porque tal definición anula la autodefensa. Sin embargo se pueden relajar las exigencias y obtener los mismos resultados que con s-aceptabilidad de la siguiente manera:

Definición 4: Un argumento $A \in AR$ será *r-aceptable* c.r. a un conjunto $S \subseteq AR$ si para todo argumento $B \in AR$ tal que:

1. Si B derrota asimétricamente a A, existe por lo menos un argumento $C \in S$ tal que C derrota asimétricamente a B, o
- ii Si B derrota simétricamente a A, existe por lo menos un argumento $C \in S$ tal que C derrota a B (simétrica o asimétricamente)

Definición 5: Un conjunto de argumentos libre de conflictos S será *r-admisibles* si cada argumento en S es r-aceptable c.r. a S

La diferencia con la definición 2 radica en el hecho que S debe defender estrictamente sólo a los argumentos derrotados estrictamente y en consecuencia permite la autodefensa de modo que esta noción puede dar cuenta de lo esperado para el caso del diamante de Nixon y al mismo tiempo brindar los mismos resultados que s-admisibilidad en los otros ejemplos

Las relaciones con MAE pueden realizarse a partir de la siguiente proposición:

Proposición 1. Si un conjunto S es r-admisibles, entonces S es admisible.

Prueba. Sea S un conjunto r -admisibles. Es claro que si un argumento A es r -aceptable con respecto a S entonces A es aceptable con respecto a S , de modo que si cada argumento en S es r -aceptable entonces cada argumento en S es aceptable, en consecuencia, si cada argumento en S es aceptable, S también es admisible.

La noción de r -aceptabilidad propone una defensa relativa al tipo de derrota que afecte a un argumento. A partir de la misma es posible construir un marco argumentativo alternativo a MAE que opere a partir de ella. La proposición 1 permitirá establecer conexiones con MAE. Sin embargo, una de las dificultades que tendrá la construcción de tal marco es la definición de la semántica estable. ¿Una extensión será r -estable cuando derrote asimétricamente a cada argumento que no pertenece a la extensión? Tal idea parece demasiado extremista. El ejemplo 7, en ese caso, no tendría extensión r -estable. Por otro lado, si la semántica estable se conserva igual que en MAE, el nuevo marco invalidaría las relaciones entre extensiones estables y preferidas ya que cada extensión estable no necesariamente sería r -preferida (i.e. r -admisibles maximal). En el ejemplo 2 mientras que $\{B,D\}$ y $\{A,C\}$ serían estables en sentido estándar $\{A,C\}$ no es r -preferida. En consecuencia, la tarea consistirá en explorar y valorar una definición de una extensión estable menos exigente que la r -estable o establecer algún puente entre r -preferidas y estables en el sentido estándar.

Si la búsqueda anterior no conduce a algún resultado alentador, parecería interesante explotar la idea esbozada en la introducción sobre la identificación de una semántica que defina como extensiones a aquellos conjuntos que verifican la propiedad de ser una extensión completa y a su vez cada subconjunto propio es una extensión completa. En principio cada argumento autodefendido o aceptable con respecto al conjunto vacío parecerían satisfacer esta propiedad, en consecuencia tal semántica debiera seleccionar como extensión al conjunto coherente que satisfaga la condición de que cada argumento que pertenece al conjunto sea aceptable con respecto a sí mismo. Tal semántica permitiría obtener los mismos resultados que r -aceptabilidad pero al mismo tiempo conservar la unidad de MAE.

Conclusión

En el presente trabajo se ha pretendido sentar las bases para que los marcos argumentativos eviten el comportamiento extraño indicado por (3) mediante la extensión de la noción de aceptabilidad. Los resultados parecen favorables y de hecho se ha mostrado el comportamiento adecuado frente a los ejemplos citados y además se ha demostrado una propiedad entre el conjunto r -admisibles y el conjunto admisible de MAE. Sin embargo, la semántica estable que puede definirse a partir de la r -aceptabilidad parece ser demasiado exigente o problemática en caso de conservar la estándar. Esto supone una dificultad por el hecho de que las conexiones entre MAE y otros formalismos se definen a partir de esta semántica. Por ello, el trabajo futuro consistirá en explorar y valorar los resultados que puedan obtenerse en esa línea pero simultáneamente investigar posibles vías de solución al mismo problema pero a partir de la definición de una semántica que someramente ha sido expresada, y que puede ser identificada en la semántica completa como aquella extensión completa tal que cada subconjunto propio es a su vez una extensión completa.

Bibliografía

- DIMOPOULOS, Y.; MORAITIS, P., AMGOUD L. Extending Argumentation to Make Good Decisions. en *Proc. of ADT'2009*. Pp. 225-236, 2009.
- DUNG, P.M. On the Acceptability of Arguments and its Fundamental Role in Nonmonotonic Reasoning, Logic Programming and N-Person Games. *Artificial Intelligence*, 77: 321–358, 1995
- HORTY, J. Argument Construction and Reinstatement in Logics For Defeasible Reasoning. *Artificial Intelligence and Law*, 9: 1-28, 2001.
- MARTINEZ, D.; GARCÍA, A.; SIMARI, G. On Defense Strength of Blocking Defeaters in Admissible Sets. en *Proc. of KSEM'2007*. Pp 140-152, 2007.
- MARTINEZ, D., GARCÍA, A.; SIMARI, G. Strong and Weak forms of Abstract Argument Defense. en *Proc. of COMMA'2008* pp 216—227, 2008.
- PRAKKEN, H. Intuitions and the Modelling of Defeasible Reasoning: Some Case Studies. en *Proc. of the 9th International Workshop on NMR'2002*. Pp: 91–102, 2002.