Sobre la Relación de Preferencia entre Argumentos

Laura A. Cecchi

Depto. de Informática y Estadística - Fa.E.A.
Universidad Nacional del Comahue
e-mail:lcecchi@uncoma.edu.ar

Guillermo R. Simari

Depto. de Ciencias de la Computación
Universidad Nacional del Sur
e-mail:grs@cs.uns.edu.ar

Palabras Claves: Argumentos - Razonamiento Rebatible - Especificidad Generalizada

Resumen

Los diferentes modelos de sistemas argumentativos son formalismos de representación y razonamiento no monotónico que permiten caracterizar una de los principales habilidades de la inteligencia humana: la argumentación.

Al definir un framework argumentativo se exige un criterio de preferencia entre argumentos, sin restringir la relación que formaliza su comportamiento. Esto motivó el estudio dicha relación, desde el punto de vista matemático, a fin de caracterizarla.

El propósito de este trabajo es presentar los resultados obtenidos en dicho análisis, emitiendo juicio sobre la aptitud de cada propiedad.

Asimismo, se consideraron dos criterios particulares: la especificidad, implementada a través de la especificidad generalizada y la equi-especificidad [9, 6], y las relaciones basadas en prioridades entre reglas. En ambos casos, se los estudió sin restringir el conjunto de argumentos a aquellos en conflicto.

1 Introducción

Los diferentes modelos de sistemas argumentativos son formalismos de representación y razonamiento no monotónico que permiten caracterizar una de los principales habilidades de la inteligencia humana: la argumentación.

Dung en [5] refleja con la vieja frase *"el que ríe último, ríe mejor"* el modo en que los humanos realizamos el proceso argumentativo. Aquel argumento que no tiene contraargumentos es el que gana el debate. En este proceso se requiere de un mecanismo que nos permita determinar si un argumento que está en *conflicto* con otro es mejor que este último.

La especificidad[7] que ha sido utilizada en diferentes sistemas de razonamiento, como las redes semánticas [5, 10, 11], la lógica default [7] y los sistemas argumentativos [9], es uno de los mecanismos que nos permite elegir entre caminos que compiten. Si bien existen varias definiciones de especificidad, todas operan sobre el mismo principio: *información más específica, es probable que sea más exacta.*

Un criterio alternativo para decidir entre diferentes argumentos que compiten es el basado en prioridades entre reglas. Determinar si un argumento es mejor que otro involucra un análisis del conjunto de las reglas rebatibles que lo componen.

Al definir un framework argumentativo se exige un criterio de preferencia entre argumentos, sin restringir la relación que formaliza su comportamiento. Esto motivó el estudio de dicha relación, desde el punto de vista matemático, a fin de caracterizarla. El propósito de este trabajo es presentar los resultados obtenidos en dicho análisis, emitiendo juicio sobre los beneficios y las desventajas de que se cumpla cada propiedad. Asimismo, se consideran dos criterios particulares: la especificidad, implementada a través de la especificidad generalizada y la equi-especificidad [9, 6], y las relaciones basadas en prioridades entre reglas.

La organización de este trabajo es la siguiente. En la sección 2, se define una relación de preferencia, describiendo dos casos particulares: argumentos preferidos y argumentos incomparables. A partir de la formalización de dichas nociones se estudian las diferentes propiedades matemáticas, analizando la importancia de que se verfiquen. En la sección 3, se describen el criterio de especificidad generalizada y la equiespecificidad y se los caracteriza en función de las propiedades detalladas en la sección anterior. Análogamente, en la sección 4, se analiza el criterio de preferencia basado en prioridades entre reglas. Todos los resultados expuestos en las secciones 3 y 4 están justificados a través de demostraciones o contraejemplos, según corresponda. Finalmente, en la sección 5, se presentan nuestras conclusiones y se mencionan algunos posibles trabajos de investigación futuros.

2 Relación de preferencia entre argumentos

En los *sistemas de argumentación*, que constituyen una formalización del razonamiento no monótono, una creencia será explicada después de que el agente realice un análisis dialéctico, teniendo en cuenta las explicaciones tentativas que pueda construir para tal creencia. Las explicaciones que soportan a la creencia se denominan *argumentos* y son el fundamento del razonamiento rebatible.

La construcción de los argumentos es monotónica, i.e., un argumento permanece inalterable aún si agregamos nuevas premisas. La no monotonicidad no está fundamentada en términos de un argumento simple, sino en términos de interacción entre argumentos conflictivos. Cada vez que necesitamos determinar si una creencia está garantizada por un sistema argumentativo,

y, por lo tanto, es aceptada por el agente, deberemos realizar el análisis dialéctico en el que los argumentos a favor y en contra de esa creencia sean considerados. La pérdida de la monotonicidad se reflejará al descartar un argumento por ser derrotado por otro.

El análisis dialéctico formaliza la interacción entre argumentos en conflicto, así, necesitamos la noción de derrota entre argumentos. Informalmente, un argumento derrota a otro que lo ataca, ya sea en su conclusión final o en un resultado intermedio, si es mejor según algún criterio de preferencia.

Las definiciones de frameworks para sistemas argumentativos incluyen un parámetro que se corresponde con la relación binaria de derrrota entre argumentos enfrentados.

Definición 2.1. (Framework Argumentativo)[5]
Un framework argumentativo es un par $\langle Args, derrotador \rangle$ donde $Args$ es el conjunto de argumentos y $derrotador$[1] es una relación binaria entre $Args$. ■

La relación *derrotador* está definida entre argumentos que se atacan. Sin embargo, nuestro interés estará centrado en el análisis de la relación binaria de preferencia entre argumentos, aún si éstos no están en conflicto.

Definición 2.2. (Relación de Preferencia $\mathcal{R}$)
Sea `Args` el conjunto de todos los argumentos que se pueden obtener a partir de un sistema argumentativo. Una *relación de preferencia* $\mathcal{R} \subseteq \texttt{Args} \times \texttt{Args}$ es cualquier relación binaria definida sobre `Args`. ■

Así, si un argumento $\mathcal{A}_1$ es mejor o igual que otro $\mathcal{A}_2$, lo notaremos $\mathcal{A}_1 \ \mathcal{R} \ \mathcal{A}_2$ o bien como par ordenado $(\mathcal{A}_1, \mathcal{A}_2)$.

Dados dos argumentos deberemos considerar dos casos especiales, alguno de ellos es mejor o bien son incomparables. Diferenciamos dos relaciones binarias, basadas en la relación de preferencia $\mathcal{R}$, que formalizan estas ideas.

Definición 2.3. (Relaciones Binarias $\succ$ y $\asymp$)[3]
Sean `Args` el conjunto de todos los argumentos que se pueden obtener a partir de un sistema argumentativo y $\{\mathcal{A}_1, \ \mathcal{A}_2\} \subseteq Args$. Entonces, escribiremos,

1. $\mathcal{A}_1 \succ \mathcal{A}_2$ si $\mathcal{A}_1 \mathcal{R} \mathcal{A}_2$ y $\mathcal{A}_2 \not\mathcal{R} \mathcal{A}_1$. En este caso diremos que $\mathcal{A}_1$ es estrictamente preferido por sobre $\mathcal{A}_2$.

2. $\mathcal{A}_1 \ \asymp \ \mathcal{A}_2$ si

   (a) $\mathcal{A}_1 \mathcal{R} \mathcal{A}_2$ y $\mathcal{A}_2 \mathcal{R} \mathcal{A}_1$.

   (b) $\mathcal{A}_1$ y $\mathcal{A}_2$ no son comparables por $\mathcal{R}$, i.e., los pares $(\mathcal{A}_1, \mathcal{A}_2)$ y $(\mathcal{A}_2, \mathcal{A}_1)$ no pertenecen a la relación $\mathcal{R}$.

■

Notaremos $\mathcal{R}_{\mathrm{P}}$, $\succ_{\mathrm{P}}$ y $\asymp_{\mathrm{P}}$ cuando querramos asociar las relaciones a un criterio de preferencia determinado[3].

Si bien tenemos como meta que al comparar un par de argumentos podamos decidir cuál es el mejor, pretenderemos, asimismo, que la elección refleje los nexos existentes entre los objetos del mundo. Por esta razón, seleccionamos y analizamos el siguiente conjunto de propiedades que podría cumplir una relación de preferencia, emitiendo juicio sobre su aptitud.

[1] El artículo [5] denomina a la relación *derrotador* "ataque". Sin embargo, creemos, al igual que muchos autores [8], que *derrotador* es más adecuada.

Reflexiva: $\mathcal{A}_1 \mathcal{R}_{\mathtt{P}} \mathcal{A}_1$, i.e., $\mathcal{A}_1$ es mejor que $\mathcal{A}_1$ bajo el criterio $\mathtt{P}$.

Ningún argumento puede ser mejor que sí mismo, por lo tanto, nuestro parecer es que una relación de preferencia no sólo no debería ser reflexiva, sino que debería ser **irreflexiva**.

Simétrica: $\mathcal{A}_1 \mathcal{R}_{\mathtt{P}} \mathcal{A}_2$ entonces $\mathcal{A}_2 \mathcal{R}_{\mathtt{P}} \mathcal{A}_1$.

Esta propiedad no es deseable, ya que en caso de cumplirla no podríamos preferir a ninguno de los argumentos comparados. Una relación de preferencia que cumple esta propiedad no es adecuada para nuestro propósito, ya que para todo par de argumentos $\mathcal{A}_1$ y $\mathcal{A}_2$, $\mathcal{A}_1 \asymp \mathcal{A}_2$.

Asimétrica: $\mathcal{A}_1 \mathcal{R}_{\mathtt{P}} \mathcal{A}_2$ entonces $\mathcal{A}_2 \not\mathcal{R}_{\mathtt{P}} \mathcal{A}_1$.

Deberíamos tener predilección sobre las relaciones de preferencia que cumplan esta propiedad. La asimetría trae como consecuencia que el conjunto de pares de argumentos que cumplen $\mathcal{A}_1 \asymp \mathcal{A}_2$ sean sólo aquellos que no están en relación $\mathcal{R}_{\mathtt{P}}$.

Antisimétrica: $\mathcal{A}_1 \mathcal{R}_{\mathtt{P}} \mathcal{A}_2$ y $\mathcal{A}_2 \mathcal{R}_{\mathtt{P}} \mathcal{A}_1$ entonces $\mathcal{A}_2 = \mathcal{A}_1$.

Al igual que la propiedad asimétrica, esta propiedad restringe el conjunto de pares ordenados que cumplen $\mathcal{A}_1 \asymp \mathcal{A}_2$. Así, priviligearemos las relaciones que la verifican.

Transitiva: $\mathcal{A}_1 \mathcal{R}_{\mathtt{P}} \mathcal{A}_2$ y $\mathcal{A}_2 \mathcal{R}_{\mathtt{P}} \mathcal{A}_3$ entonces $\mathcal{A}_1 \mathcal{R}_{\mathtt{P}} \mathcal{A}_3$.

Si bien esta propiedad parece deseable, muchas veces los argumentos $\mathcal{A}_1$ y $\mathcal{A}_3$ no tienen afinidad en el tema que tratan, aunque ambos tengan una conexión con $\mathcal{A}_2$. Nuestro objetivo es poder elegir entre pares de argumentos, pero sin forzar un nexo inexistente en el mundo real.

Euclidiana: $\mathcal{A}_1 \mathcal{R}_{\mathtt{P}} \mathcal{A}_2$ y $\mathcal{A}_1 \mathcal{R}_{\mathtt{P}} \mathcal{A}_3$ entonces $\mathcal{A}_2 \mathcal{R}_{\mathtt{P}} \mathcal{A}_3$.

En [4], J. Delgrande presenta una noción más debilitada de esta propiedad, que denomina *Forward Connected*:

$$\mathcal{A}_1 \mathcal{R}_{\mathtt{P}} \mathcal{A}_2 \text{ y } \mathcal{A}_1 \mathcal{R}_{\mathtt{P}} \mathcal{A}_3 \text{ entonces } \mathcal{A}_2 \mathcal{R}_{\mathtt{P}} \mathcal{A}_3 \text{ o } \mathcal{A}_3 \mathcal{R}_{\mathtt{P}} \mathcal{A}_2$$

Esta propiedad establece un modo diferente de conectar argumentos; si una argumento está en relación con dos argumentos,entonces estos últimos deben tener alguna relación. Al igual que en el caso de la propiedad transitiva, esta propiedad podría forzar la conexión entre argumentos, posiblemente, dispares. Sin embago, creemos que en el mundo real es más probable que se encuentren situaciones que verifiquen dicha condición.

Una relación de preferencia que cumple la propiedad simétrica no será de gran utilidad, ya que haría que todos los elementos en relación fuesen incomparables y, por lo tanto, la elección sería nula. En consecuencia, restringiremos las relaciones a aquellas que *no son de equivalencia*.

En mucha literatura sobre sistemas argumentativos, se exige que la relación de preferencia sea un orden[13] o un orden estricto[1]. Nuestro juicio es que la relación de preferencia debería ser restringida a aquellas relaciones binarias que cumplen las propiedades irreflexiva y asimétrica. Esto ayudaría a resolver algunos conflictos que se presentan en el proceso dialéctico.

En las secciones que siguen analizaremos dos clases de relaciones de preferencia: la especificidad generalizada y las basadas en prioridades para las reglas.

3 Especificidad Generalizada

La especificidad [7, 9, 6] es uno de los mecanismos más estudiados que permite decidir entre caminos que compiten. Su principal beneficio reside en el hecho de evitar que el usuario tenga que preocuparse por dar un valor cuantitativo a las reglas que permiten concluir un literal. Lamentablemente, ningún mecanismo de preferencia[12, 8] se comporta de la forma adecuada, representando nuestra intuición, en todos los casos. Sin embargo, creemos que la especificidad se adapta a la mayoría de los casos, ya que captura la intuición del mundo real a través de un principio ampliamente utilizado.

La siguiente definición caracteriza el criterio de especificidad generalizada, enmarcándolo en el ambiente de la Programación en Lógica Rebatible[6].

Definición 3.1. (Derivable Rebatiblemente)[6]
Sean $\prec Q_1, \ldots, Q_n$ una meta rebatible y R una regla estricta o rebatible con cabeza Q_i, $1 \leq i \leq n$ y cuerpo $L_1, \ldots, L_m$. Luego la nueva meta rebatible

$$\prec Q_1, \ldots, Q_{i-1}, L_1, \ldots, L_m, Q_{i+1}, \ldots, Q_n$$

es *derivable rebatiblemente.* ■

Definición 3.2. (Derivación Rebatible)[6]
Sean $\mathcal{P}$ un programa rebatible y Q una consulta. Una *derivación rebatible* consiste de una secuencia finita de metas rebatibles $\prec Q = Q_0, Q_1, \ldots, Q_n$, donde cada Q_{i+1} es derivado rebatiblemente a partir de $Q_i, 0 \leq i < n$ y $Q_n = \square$ es la meta vacía. Diremos que Q se deriva rebatiblemente a partir de $\mathcal{P}$. ■

Definición 3.3. (Derivable estrictamente)[6]
Sean $\prec Q_1, \ldots, Q_n$ una meta rebatible y R una regla estricta con cabeza Q_i, $1 \leq i \leq n$ y cuerpo $L_1, \ldots, L_m$. Luego la nueva meta rebatible

$$\prec Q_1, \ldots, Q_{i-1}, L_1, \ldots, L_m, Q_{i+1}, \ldots, Q_n$$

es *derivable estrictamente.* ■

Definición 3.4. (Derivación Estricta)[6]
Sean $\mathcal{P}$ un *plr* y Q una consulta rebatible. Una *derivación estricta* consiste de una secuencia finita de metas rebatibles $\prec Q = Q_0, Q_1, \ldots, Q_n$, donde cada Q_{i+1} es derivado estrictamente a partir de $Q_i, 0 \leq i < n$ y $Q_n = \square$ es la meta vacía. Diremos que Q se deriva estrictamente a partir de $\mathcal{P}$. ■

Definición 3.5. (Especificidad Generalizada)[6]
Sean $\mathcal{P} = (\Pi, \Delta)$ un programa lógico rebatible y Π_G el conjunto de las reglas estrictas de Π, sin incluir los hechos. Sea $\mathcal{F}$ el conjunto de todos los literales que tienen una derivación rebatible a partir de $\mathcal{P}$ ($\mathcal{F}$ será considerado como un conjunto de hechos). Sean $\langle \mathcal{A}_1, h_1 \rangle$ y $\langle \mathcal{A}_2, h_2 \rangle$ dos estructuras de argumentos obtenidos desde $\mathcal{P}$. $\langle \mathcal{A}_1, h_1 \rangle$ es *estrictamente más específico que* $\langle \mathcal{A}_2, h_2 \rangle$, que notaremos $\mathcal{A}_1 \mathcal{R}_{\mathcal{E}} \mathcal{A}_2$ o $(\mathcal{A}_1, \mathcal{A}_2) \in \mathcal{R}_{\mathcal{E}}$, si se cumplen las siguientes condiciones:

1. para todo $H \subseteq \mathcal{F}$: si existe una derivación rebatible para h_1 a partir de $\Pi_G \cup H \cup \mathcal{A}_1$ y no existe una derivación estricta para h_1 a partir de $\Pi_G \cup H$, entonces existe una derivación rebatible para h_2 a partir de $\Pi_G \cup H \cup \mathcal{A}_2$, y

2. existe al menos un conjunto $H' \subseteq \mathcal{F}$ tal que: existe una derivación rebatible para h_2 a partir de $\Pi_G \cup H' \cup \mathcal{A}_2$, no existe una derivación estricta de h_2 a partir de $\Pi_G \cup H'$ y no existe una derivación rebatible de h_1 a partir de $\Pi_G \cup H' \cup \mathcal{A}_1$.

■

La primer condición de esta regla establece cada vez que se activa, de modo no trivial, el argumento $\mathcal{A}_1$ se debe activar el argumento $\mathcal{A}_2$. La segunda condición asegura que si se activa el argumento $\mathcal{A}_2$, en forma no trivial, entonces no debe activarse el argumento $\mathcal{A}_1$.

La especificidad generalizada cumple las propiedades irreflexiva, asimétrica y antisimétrica como se muestra a continuación. Las demostraciones de los lemas se encuentran en el apéndice.

Lema 3.1. La relación de preferencia entre argumentos basada en el criterio de especificidad generalizada es irreflexiva, i.e., para toda estructura de argumento $\langle \mathcal{A}_1, h_1 \rangle$ de un programa lógico rebatible $\mathcal{P}$, $(\mathcal{A}_1, \mathcal{A}_1) \notin \mathcal{R}_{\mathcal{E}}$. ■

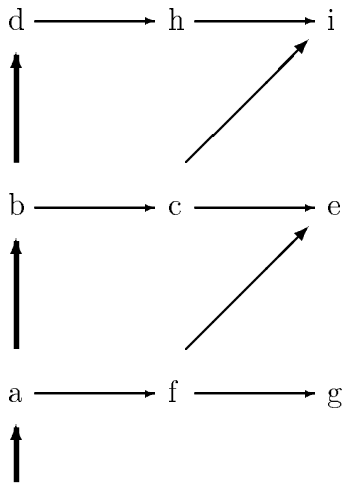
Lema 3.2. La relación de preferencia entre argumentos basada en el criterio de especificidad generalizada es asimétrica, i.e., para todo par de estructuras de argumento $\langle \mathcal{A}_1, h_1 \rangle$, $\langle \mathcal{A}_2, h_2 \rangle$ de un programa lógico rebatible $\mathcal{P}$, si $(\mathcal{A}_1, \mathcal{A}_2) \in \mathcal{R}_{\mathcal{E}}$ entonces $(\mathcal{A}_2, \mathcal{A}_1) \notin \mathcal{R}_{\mathcal{E}}$. ■

Corolario 3.3. La relación de preferencia entre argumentos basada en el criterio de especificidad generalizada es antisimétrica, i.e., para todo par de estructuras de argumento $\langle \mathcal{A}_1, h_1 \rangle$ y $\langle \mathcal{A}_2, h_2 \rangle$ de un programa lógico rebatible $\mathcal{P}$, si $(\mathcal{A}_1, \mathcal{A}_2) \in \mathcal{R}_{\mathcal{E}}$ y $(\mathcal{A}_2, \mathcal{A}_1) \in \mathcal{R}_{\mathcal{E}}$, entonces $\mathcal{A}_1 = \mathcal{A}_2$. ■

La especificidad generalizada no es transitiva. El siguiente es un contraejemplo. Para poder apreciar mejor la situación, se graficó el ejemplo, siguiendo la notación gráfica clásica en Inteligencia Artificial: las flechas más gruesas indican que existe una regla estricta y las flechas finas indican que existe una regla rebatible.

Ejemplo 3.1. Consideremos el siguiente programa lógico rebatible:

$$\Pi = \left\{ \begin{array}{l} \texttt{a.} \\ \texttt{b}\leftarrow \texttt{a.} \\ \texttt{d}\leftarrow\texttt{b.} \\ \texttt{e}\leftarrow \texttt{f.} \\ \texttt{i}\leftarrow \texttt{c.} \end{array} \right\} \qquad \Delta = \left\{ \begin{array}{l} \texttt{f} \prec \texttt{a} \\ \texttt{g} \prec \texttt{f} \\ \texttt{c} \prec \texttt{b} \\ \texttt{e} \prec \texttt{c} \\ \texttt{h} \prec \texttt{d} \\ \texttt{i} \prec \texttt{h} \end{array} \right\}$$



Los argumentos:

$$\mathcal{A}_1 = \{g \prec f, f \prec a\}$$

$$\mathcal{A}_2 = \{e \prec c, c \prec b\}$$

$$\mathcal{A}_3 = \{i \prec h, h \prec d\}$$

cumplen

$$\mathcal{A}_1 \succ_{\mathcal{E}} \mathcal{A}_2 \;\; y \;\; \mathcal{A}_2 \succ_{\mathcal{E}} \mathcal{A}_3$$

pero no ocurre que

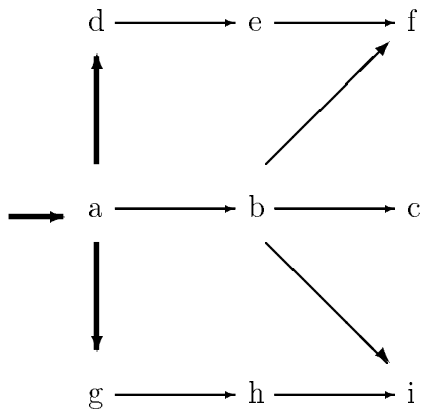
$$\mathcal{A}_1 \succ_{\mathcal{E}} \mathcal{A}_3$$

■

La especificidad generalizada no cumple la propiedad euclideana.

Ejemplo 3.2. Consideremos el siguiente programa lógico rebatible:

$$\Pi = \left\{ \begin{array}{l} \texttt{a.} \\ \texttt{d} \leftarrow \texttt{a.} \\ \texttt{g} \leftarrow \texttt{a.} \\ \texttt{f} \leftarrow \texttt{b.} \\ \texttt{i} \leftarrow \texttt{b.} \end{array} \right\} \qquad \Delta = \left\{ \begin{array}{l} \texttt{b} \prec \texttt{a} \\ \texttt{c} \prec \texttt{b} \\ \texttt{h} \prec \texttt{g} \\ \texttt{i} \prec \texttt{h} \\ \texttt{e} \prec \texttt{d} \\ \texttt{f} \prec \texttt{e} \end{array} \right\}$$



Los argumentos:

$$\mathcal{A}_1 = \{c \prec b, b \prec a\}$$

$$\mathcal{A}_2 = \{f \prec e, e \prec d\}$$

$$\mathcal{A}_3 = \{i \prec h, h \prec g\}$$

cumplen

$$\mathcal{A}_1 \succ_{\mathcal{E}} \mathcal{A}_2 \;\; y \;\; \mathcal{A}_1 \succ_{\mathcal{E}} \mathcal{A}_3$$

pero no ocurre que

$$\mathcal{A}_2 \succ_{\mathcal{E}} \mathcal{A}_3 \text{ ni } \mathcal{A}_3 \succ_{\mathcal{E}} \mathcal{A}_2$$

En el gráfico de la figura 1, se muestra una posible interpretación para la topología en la que se ve claramente que los argumentos $\mathcal{A}_1$ y $\mathcal{A}_2$ tienen alguna conexión, al igual que $\mathcal{A}_1$ y $\mathcal{A}_3$, pero que no es posible preferir entre $\mathcal{A}_2$ y $\mathcal{A}_3$. En forma similar podemos encontrar una interpretación en el mundo real, para la topología que muestra que la especificidad generalizada no es transitiva. ■

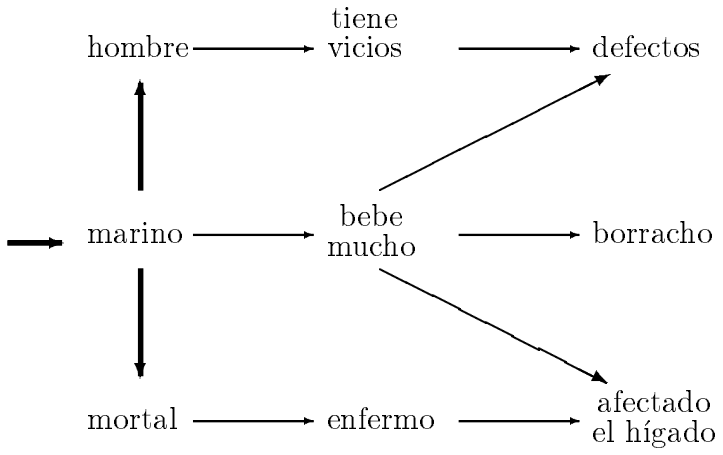


Figura 1: Una posible interpretación para la topología del contraejemplo de la propiedad euclidiana

Una definición que se desprende de la noción de especificidad generalizada es la de equi-especificidad.

Definición 3.6. **(Equi-especificidad)**[6]
Sea $\mathcal{P} = (\Pi, \Delta)$ un programa lógico rebatible, siendo Π el conjunto de las reglas estrictas y Δ el conjunto de reglas rebatibles. Dos argumentos $\langle A_1, h_1\rangle$ y $\langle A_2, h_2\rangle$ de $\mathcal{P}$ son equi-específicos, lo denotaremos $\langle A_1, h_1\rangle \equiv \langle A_2, h_2\rangle$ si y sólo si $A_1 = A_2$ y el literal h_2 tiene una derivación estricta a partir de $\Pi \cup \{h_1\}$ y el literal h_1 tiene una una derivación estricta a partir de $\Pi \cup \{h_2\}$. ■

Se puede probar que la equi-especificidad es una relación de equivalencia[6]. Por supuesto, al aplicar sobre el conjunto cociente la relación de preferencia especificidad generalizada, resurgen todas las características ya estudiadas para la especificidad.

Ejemplo 3.3. Consideremos el ejemplo 3.1. Las clases serán:

$$\begin{array}{lll}
C1 & = \{ & \langle \mathcal{A}_1, g\rangle\} \\
C2 & = \{ & \langle \mathcal{A}_2, e\rangle\} \\
C3 & = \{ & \langle \mathcal{A}_3, i\rangle\} \\
C4 & = \{ & \langle \emptyset, a\rangle, \langle \emptyset, b\rangle, \langle \emptyset, d\rangle\} \\
C5 & = \{ & \langle c \prec b, c\rangle\} \\
C6 & = \{ & \langle c \prec b, i\rangle\} \\
C7 & = \{ & \langle f \prec a, f\rangle\} \\
C8 & = \{ & \langle f \prec a, e\rangle\} \\
C9 & = \{ & \langle h \prec d, h\rangle\}
\end{array}$$

Las clases C_1, C_2 y C_3 muestran que no se cumple la transitividad. ■

Cabe aclarar que hemos analizado la especifidad generalizada desde un punto de vista matemático, sin restringirnos al conjunto de los argumentos sobre los que es más probable que sea aplicada. En otras palabras, no hemos limitado el estudio a aquellos pares de argumentos en conflicto. Bajo este marco general y como consecuencia del desarrollo de esta sección, podemos

concluir que el criterio de preferencia especificidad generalizada no puede ser clasificada como una equivalencia, ni como un orden bajo ninguno de sus subórdenes.

4 Comparación basada en Prioridades entre las Reglas

Algunos sistemas argumentativos definen prioridades entre las reglas[13, 2], con el objeto de decidir entre argumentos que compiten. Bajo este mecanismo, el programador debe encargarse de etiquetar cada regla rebatible con una etiqueta que asigne la importancia de aplicar tal inferencia sobre las otras. Las reglas estrictas no requieren de etiquetas, ya que, por un lado, no es posible preferir entre ellas y, por otro, tienen prioridad absoluta sobre las rebatibles. Con el objeto de comparar dos argumentos, deberemos analizar el conjunto de reglas rebatibles que componen cada uno de ellos.

Definición 4.1. **(Criterio de Preferencia basado en Prioridades entre Reglas)**[6, 2]
Sean $\mathcal{P} = (\Pi, \Delta)$ un programa lógico rebatible, siendo Π el conjunto de las reglas estrictas y Δ el conjunto de reglas rebatibles y $>$ un orden estricto definido explícitamente entre reglas rebatibles. Dados dos argumentos $\langle A_1, h_1\rangle$ y $\langle A_2, h_2\rangle$ de $\mathcal{P}$, diremos que $\langle A_1, h_1\rangle$ es preferido sobre $\langle A_2, h_2\rangle$ si y sólo si:

1. existe al menos una regla $r_a \in A_1$, y una regla $r_b \in A_2$, tal que $r_a > r_b$, y
2. no existe $r'_a \in \mathcal{A}_1$ y $r'_b \in \mathcal{A}_2$, tal que $r'_b > r'_a$

■

En este contexto, los argumentos vacíos tienen peferencia sobre cualquier otro argumento no vacío. Sin embargo, los argumentos vacíos son incomparables entre ellos.

La preferencia de un argumento sobre otro basados en las prioridades entre reglas cumple las propiedades irreflexiva, asimétrica, antisimétrica y transitiva como lo muestran los siguientes lemas (las demostraciones se muestran en el apéndice):

Lema 4.1. La relación de preferencia entre argumentos basada en prioridades entre reglas es irreflexiva, i.e., para toda estructura de argumento $\langle \mathcal{A}_1, h_1\rangle$ de un programa lógico rebatible $\mathcal{P}$, $(\mathcal{A}_1, \mathcal{A}_1) \notin \mathcal{R}_{\mathcal{P}}$. ■

Lema 4.2. La relación de preferencia entre argumentos basada en prioridades entre reglas es asimétrica, i.e., para todo par de estructuras de argumento $\langle \mathcal{A}_1, h_1\rangle$, $\langle \mathcal{A}_2, h_2\rangle$ de un programa lógico rebatible $\mathcal{P}$, si $(\mathcal{A}_1, \mathcal{A}_2) \in \mathcal{R}_{\mathcal{P}}$ entonces $(\mathcal{A}_2, \mathcal{A}_1) \notin \mathcal{R}_{\mathcal{P}}$. ■

Corolario 4.3. La relación de preferencia entre argumentos basada en en prioridades entre reglas es antisimétrica. ■

Lema 4.4. La relación de preferencia entre argumentos basada en prioridades entre reglas es transitiva. ■

Este criterio no cumple la propiedad euclidiana. Esto puede verificarse asumiendo dos argumentos $\mathcal{A}_2$ y $\mathcal{A}_3$ cuyas reglas tengan todas la misma prioridad.

En un marco análogo al que hemos estudiado a la especificidad generalizada, podemos concluir por los lemas anteriores que la relación de preferencia entre argumentos basada en prioridades entre reglas es un orden estricto parcial. Si bien, mostramos que no siempre la propiedad transitiva captura la intuición, el beneficio de este criterio estará en el fundamento matemático que soporta su definición.

5 Conclusiones y Trabajos Futuros

Con el espíritu de caracterizar las relaciones que capturan el comportamiento de un criterio de preferencia, hemos presentado una formalización de dicha relación entre argumentos y analizado un conjunto de propiedades, detallando la importancia de que se cumplan, fundamentalmente, la irreflexiva y la asimétrica. Se mostró, ejemplificando con interpretaciones en el mundo real, que la transitividad y la propiedad euclidiana no siempre capturan la intuición.

Se consideraron dos criterios muy utilizados en el marco de los sistemas argumentativos: la especificidad generalizada y la relación de preferencia basada en prioridades entre reglas. En ambos casos, se los estudió sin restringir el conjunto de argumentos a aquellos en conflicto.

Estos resultados nos motivan a contemplar entre nuestros trabajos futuros la caracterización del comportamiento de las relaciones de preferencia, restringiendo el conjunto de argumentos a aquellos que están en conflicto.

A Apéndice

Lema 3.1: *La relación de preferencia entre argumentos basada en el criterio de especificidad generalizada es irreflexiva, i.e., para toda estructura de argumento $\langle \mathcal{A}_1, h_1 \rangle$ de un programa lógico rebatible $\mathcal{P}$, $(\mathcal{A}_1, \mathcal{A}_1) \notin \mathcal{R}_{\mathcal{E}}$.*

Demostración: Sean $\mathcal{P} = (\Pi, \Delta)$ un programa lógico rebatible y $\langle \mathcal{A}_1, h_1 \rangle$ una estructura de argumento en $\mathcal{P}$. Supongamos que $(\mathcal{A}_1, \mathcal{A}_1) \in \mathcal{R}_{\mathcal{E}}$, luego por definición de especificidad generalizada (condición (2)), existe $H' \subseteq \mathcal{F}$, siendo $\mathcal{F}$ el conjunto de todos los literales que tienen una derivación rebatible a partir de $\mathcal{P}$, tal que $\Pi_G \cup H' \cup \mathcal{A}_1 \mid\!\sim h_1$ y $\Pi_G \cup H' \not\mid\!\sim h_1$ y $\Pi_G \cup H' \cup \mathcal{A}_1 \not\mid\!\sim h_1$. Este absurdo provino de suponer que existe algún argumento $\mathcal{A}_1$ tal que $(\mathcal{A}_1, \mathcal{A}_1) \in \mathcal{R}_{\mathcal{E}}$.
$\therefore$ La especificidad generalizada es irreflexiva. ■

Lema 3.2: *La relación de preferencia entre argumentos basada en el criterio de especificidad generalizada es asimétrica, i.e., para todo par de estructuras de argumento $\langle \mathcal{A}_1, h_1 \rangle$, $\langle \mathcal{A}_2, h_2 \rangle$ de un programa lógico rebatible $\mathcal{P}$, si $(\mathcal{A}_1, \mathcal{A}_2) \in \mathcal{R}_{\mathcal{E}}$ entonces $(\mathcal{A}_2, \mathcal{A}_1) \notin \mathcal{R}_{\mathcal{E}}$.*

Demostración: Sean $\mathcal{P} = (\Pi, \Delta)$ un programa lógico rebatible y $\langle \mathcal{A}_1, h_1 \rangle$ y $\langle \mathcal{A}_2, h_2 \rangle$ estructuras de argumentos en $\mathcal{P}$. Supongamos que $(\mathcal{A}_1, \mathcal{A}_2) \in \mathcal{R}_{\mathcal{E}}$, luego por definición de especificidad generalizada (condición (2)), existe $H' \subseteq \mathcal{F}$, siendo $\mathcal{F}$ el conjunto de todos los literales que tienen una derivación rebatible a partir de $\mathcal{P}$, tal que $\Pi_G \cup H' \cup \mathcal{A}_2 \mid\!\sim h_2$ y $\Pi_G \cup H' \not\mid\!\sim h_2$ y $\Pi_G \cup H' \cup \mathcal{A}_1 \not\mid\!\sim h_1$. Así no se satisface que para todo $H \subseteq \mathcal{F}$: si $\Pi_G \cup H \cup \mathcal{A}_2 \mid\!\sim h_2$ y $\Pi_G \cup H \not\mid\!\sim h_2$ entonces $\Pi_G \cup H \cup \mathcal{A}_1 \mid\!\sim h_1$. Luego no se cumple que $(\mathcal{A}_2, \mathcal{A}_1) \in \mathcal{R}_{\mathcal{E}}$.
$\therefore$ La especificidad generalizada es asimétrica. ■

Lema 4.1: *La relación de preferencia entre argumentos basada en prioridades entre reglas es irreflexiva, i.e., para toda estructura de argumento $\langle \mathcal{A}_1, h_1 \rangle$ de un programa lógico rebatible $\mathcal{P}$, $(\mathcal{A}_1, \mathcal{A}_1) \notin \mathcal{R}_{\mathcal{P}}$.*

Demostración: *Sean $\mathcal{P} = (\Pi, \Delta)$ un programa lógico rebatible y $\langle \mathcal{A}_1, h_1 \rangle$ una estructura de argumento en $\mathcal{P}$. Supongamos que $(\mathcal{A}_1, \mathcal{A}_1) \in \mathcal{R}_{\mathcal{P}}$, luego por definición 4.1 existe al menos una regla $r \in \mathcal{A}_1$ que cumple $r > r$. Este absurdo provino de suponer que existe algún argumento $\mathcal{A}_1$ tal que $(\mathcal{A}_1, \mathcal{A}_1) \in \mathcal{R}_{\mathcal{P}}$.*
$\therefore$ La relación de preferencia entre argumentos basada en prioridades entre reglas es irreflexiva. ■

Lema 4.2: La relación de preferencia entre argumentos basada en prioridades entre reglas es asimétrica.

Demostración: *Sean* $\mathcal{P} = (\Pi, \Delta)$ *un programa lógico rebatible y* $\langle \mathcal{A}_1, h_1 \rangle$ *y* $\langle \mathcal{A}_2, h_2 \rangle$ *estructuras de argumentos en* $\mathcal{P}$. *Supongamos que* $(\mathcal{A}_1, \mathcal{A}_2) \in \mathcal{R}_\mathcal{P}$, *luego por definición 4.1 no existen reglas* $r_a \in \mathcal{A}_1$ *y* $r_b \in \mathcal{A}_2$ *tales que* $r_b > r_a$. *Por lo tanto,* $(\mathcal{A}_2, \mathcal{A}_1) \notin \mathcal{R}_\mathcal{P}$.
$\therefore$ *La relación de preferencia entre argumentos basada en prioridades entre reglas es asimétrica.* ■

Lema 4.4: La relación de preferencia entre argumentos basada en prioridades entre reglas es transitiva.

Demostración: *Sean* $\mathcal{P} = (\Pi, \Delta)$ *un programa lógico rebatible y* $\langle \mathcal{A}_1, h_1 \rangle$, $\langle \mathcal{A}_2, h_2 \rangle$ *y* $\langle \mathcal{A}_3, h_3 \rangle$ *estructuras de argumentos en* $\mathcal{P}$. *Supongamos que* $(\mathcal{A}_1, \mathcal{A}_2) \in \mathcal{R}_\mathcal{P}$ *y* $(\mathcal{A}_2, \mathcal{A}_3) \in \mathcal{R}_\mathcal{P}$, *luego por definición 4.1, existe al menos una regla* $r_a \in \mathcal{A}_1$, $r_b, r'_b \in \mathcal{A}_2$ *y* $r_c \in \mathcal{A}_3$ *tales que* $r_a > r_b$ *y* $r'_b > r_c$. *Sea* r^s_a *la regla con mayor prioridad de* $\mathcal{A}_1$. *Así* $r^s_a > r_b$. *Ya que la segunda condición de la definición 4.1 no permite que exista ninguna regla en* $\mathcal{A}_2$ *con mayor prioridad que una de* $\mathcal{A}_1$, *entonces* $r^s_a > r'_b$ *o tienen la misma prioridad. En ambos casos, se cumple que* $r^s_a > r_c$.
Resta probar que no existen reglas $r'_a \in \mathcal{A}_1$ *y* $r'_c \in \mathcal{A}_3$ *tales que* $r'_c > r'_a$. *Supongamos que existen un par de reglas* r'_a *y* r'_c *que verifican*

$$r'_c > r'_a \tag{1}$$

Por la segunda de la definición 4.1 y sabiendo que $(\mathcal{A}_1, \mathcal{A}_2) \in \mathcal{R}_\mathcal{P}$, *tenemos que para toda regla en* $r_a \in \mathcal{A}_1$ *y toda regla* $r_b \in \mathcal{A}_2$ *o bien* $r_a > r_b$ *o bien son incomparables porque tienen la misma etiqueta. Por lo tanto, conociendo (1), concluimos que* $r'_c > r_b$ *lo cual viola la definición de* $(\mathcal{A}_2, \mathcal{A}_3) \in \mathcal{R}_\mathcal{P}$. *Por lo tanto, no existen tal par de reglas.*
$\therefore$ *La relación de preferencia entre argumentos basada en prioridades entre reglas es transitiva.* ■

Referencias

[1] *G. Antoniou, D. Billington, and M.J. Maher. Normal forms for defeasible logic. In J. Jaffar, editor,* Proceedings International Joint Conference and Symposium on Logic Programming, *pages 160–174. MIT Press, 1998.*

[2] *Grigoris Antoniou, Michael Maher, and David Billington. Defeasible Logic versus logic programming without negation as failure.* Journal of Logic Programming, *(42):47–57, 2000.*

[3] *Carlos Iván Chesñevar.* Formalización de los Procesos de Argumentación Rebatible como Sistemas Deductivos Etiquetados. *PhD thesis, Universidad Nacional del Sur, 2001.*

[4] *James P. Delgrande. A Logic for Representing Default and Prototypical Properties. In* Proc. on Knowledge Representation - IJCAI, *pages 423–429, 1987.*

[5] *Phan M. Dung. On the acceptability of arguments and its fundamental role in nonmonotonic reasoning annd logic programming and n-person games.* Artificial Intelligence, *77:321–357, 1995.*

[6] *Alejandro J. García.* Programación en Lógica Rebatible: Lenguaje, Semántica Operacional y Paralelismo. *PhD thesis, Universidad Nacional del Sur, 2000.*

[7] *David L. Poole. On the Comparasion of Theories: Preferring the Most Specific Explanation. In* Proc. IJCAI, *pages 144–147, 1985.*

[8] *Henry Prakken and Gerard Vreeswijk. Logics for defeasible argumentation. In D. Gabbay, editor,* Handbook of Philosophical Logic. *Kluwer Academic, second edition, 2000.*

[9] *Guillermo R. Simari and R.P. Loui. A mathematical treatment of defeasible reasoning and its implementation.* Artificial Intelligence, *pages 125–157, 1992.*

[10] *Richmond H. Thomason and John F. Horty. Logics for Inheritance Theory. In* Proceedings of the Second Workshop on Nonmonotonic Reasoning - LNAI, *1990.*

[11] *Richmond H. Thomason, John F. Horty, and David S. Touretzky. A calculus for inheritance in monotonic semantic nets. In Z. Ras and M. Zemankova, editors,* Methodologies for intelligent systems - Proceedings of the Second International Symposium on Methodologies for Intelligent Systems, *pages 280–287, North-Holland, Amsterdam, 1987.*

[12] *David S. Touretzky, John F. Horty, and Richmond H. Thomason. A Clash of Intuitions: The Current State of Nonmonotonic Multiple Inheritance Systems. In Morgan Kaufmaun, editor,* Proceedings of IJCAI, *pages 476–482, Los Altos, 1987.*

[13] *Gerard Vreeswijk. Abstract Argumentation Systems. Technical Report IR-289, Vrije Universiteit Amsterdam, 1993.*