

Razonamiento Rebatible Por Casos

Gustavo A. Bodanza

Centro de Investigaciones de Lógica y Filosofía de la Ciencia (C.I.L.F.)

Departamento de Humanidades

Universidad Nacional del Sur

12 de octubre y San Juan, Bahía Blanca (8.000)

Tel.: (54)(91)20776(ext.357)

e-mail:ccbodanz@criba.edu.ar

Palabras clave: *razonamiento rebatible, argumento, razonamiento por casos.*

Resumen

Este trabajo constituye un estudio de las dificultades de la incorporación de información disyuntiva en un sistema conteniendo reglas rebatibles. Se presenta una ampliación del sistema *MTDR* a fin de aumentar la capacidad de éste, permitiendo el razonamiento por casos mediante argumentos.

En un artículo anterior ([BS95]) se brindaba una aproximación al problema introduciendo una relación de inferencia rebatible que incluía la prueba por casos. Tal solución presenta limitaciones en determinados contextos. Este es un enfoque alternativo que da cuenta de esos problemas. La solución aquí viene dada por la definición de nuevos tipos de argumentos especiales para el manejo adecuado de información disyuntiva.

Razonamiento Rebatible por Casos

1 Introducción

Los sistemas de razonamiento rebatible utilizan algún tipo de condicional más débil que la implicación material, llamado *regla rebatible*, que permite formalizar expresiones tales como “si es razonable sostener p entonces es razonable sostener q ” (donde en lugar de ‘razonable’ también puede aparecer otra palabra que no denote una conexión necesaria entre el antecedente y el consecuente).¹ Estos condicionales son adecuados sólo como premisas de razonamientos por *modus ponens*, pero no de otros, como *modus tollens* o *contraposición*.² Esto hace que la operatividad de las reglas rebatibles quede restringida solamente a esa forma de razonamiento. Sin embargo, esta restricción tiene como efecto colateral el no permitir otras formas de razonamiento acordes a la intuición, como lo es el razonamiento por casos, aún cuando éste encuentra fundamento en el propio *modus ponens*. Nuestro propósito es hallar un sistema basado en reglas rebatibles en el que inferencias derivables mediante este tipo de razonamientos sean factibles.

El estudio que presentamos se realiza en el marco de un sistema argumentativo, *i.e.*, un sistema en el que las inferencias encuentran justificación o no, como resultado de la disputa entre argumentos esgrimidos a favor y en contra. Tales argumentos están conformados por reglas rebatibles, y el sistema aporta una relación de preferencia entre ellos y, también, el mecanismo por el cual se lleva a cabo la disputa (las “reglas de diálogo”). Por su sencillez y potencial, tomamos el sistema descrito por Simari y Loui en [SL92] (al cual nos referiremos mediante la sigla *MTDR*), cuyas nociones básicas exponemos en la siguiente sección.

2 Razonamiento rebatible mediante argumentos: el sistema *MTDR*

Presentamos las características formales del sistema argumentativo *MTDR*, de Simari y Loui. Luego de ver las dificultades que éste manifiesta para la representación de disyunciones, introduciremos nuestras sugerencias para el logro de los fines propuestos.

Las creencias de un agente a son representadas en *MTDR* mediante un lenguaje que está compuesto por un lenguaje de primer orden \mathcal{L} , más la relación metalingüística ‘ \succ ’

¹Sistemas clásicos de este tipo son, *e.g.*, los de Reiter [R80], McDermott y Doyle [MD80] y Poole [P85].

²Por ejemplo, las razones que nos permiten concluir que una persona tomada al azar no es diabética (razones, en este caso, estadísticas), no nos permiten concluir que un diabético tomado al azar no es una persona.

entre miembros de \mathcal{L} que contienen variables libres. Los elementos de esta relación se llaman *reglas rebatibles* y tienen la forma ' $\alpha \succ \beta$ ' (donde α y β son fbf's de \mathcal{L}), y se interpretan informalmente como "creer en α es una buena razón para creer en β ". Una instancia de una regla rebatible se obtiene reemplazando todas las variables libres por constantes.

Un conjunto consistente $\mathcal{K} \subset \mathcal{L}$ se llama *contexto* y puede ser particionado en dos subconjuntos \mathcal{K}_G y \mathcal{K}_P de creencias *generales* y *particulares*, respectivamente.

Las creencias de a en *MTDR* están representadas por el par (\mathcal{K}, Δ) , llamado *Estructura Lógica Rebatible*, donde Δ es un conjunto de reglas rebatibles. Δ^\dagger es el conjunto de todos los miembros de Δ instanciados. Se introduce además una relación de inferencia no deductiva, definida como sigue:

Definición 2.1 Sea $\Gamma \subseteq \mathcal{K} \cup \Delta^\dagger$. Un literal instanciado h es una *consecuencia rebatible* de Γ , en símbolos $\Gamma \sim h$, si y sólo si existe una secuencia finita A_1, \dots, A_n tal que $A_n = h$ y para todo $1 \leq i \leq n$, se verifica $A_i \in \Gamma$, o A_i es una consecuencia directa de los elementos precedentes en la secuencia por la aplicación de modus ponens o instanciación. Las reglas rebatibles instanciadas son consideradas del mismo modo que el condicional material para la aplicación de modus ponens.

La relación \sim determina el conjunto de todas las sentencias *sancionadas* en el sistema. Las sentencias *justificadas*, *i.e.*, las que el sistema recomienda, comprenden un subconjunto de éste. Luego veremos que el sistema brinda un criterio para determinar cuáles son las inferencias recomendadas en caso de presentarse inconsistencias. Cuando esto ocurre, la consistencia se resguarda realizando una selección entre la información derivada, eliminando las partes conflictivas que se sustentan en argumentos menos preferidos (más adelante se aclararán estos puntos).

3 Problemas en la combinación de reglas rebatibles con disyunciones

Puesto que la relación de consecuencia rebatible \sim no comprende otras reglas de inferencia que no sean modus ponens e instanciación, excluye ciertos tipos de razonamiento intuitivamente correctos, aún cuando las inferencias procedan de premisas rebatibles. Uno de éstos es el *razonamiento por casos*, cuya validez se vincula con la del modus ponens.

Ejemplo 3.1 Del conjunto $\{\alpha \succ \gamma, \beta \succ \gamma, \alpha \vee \beta\}$ no se deriva rebatiblemente γ , aún cuando $\{\alpha \succ \gamma, \alpha\} \sim \gamma$, y $\{\beta \succ \gamma, \beta\} \sim \gamma$.

En un sistema "escéptico" como *MTDR*, también puede ocurrir que la presencia de disyunciones provoque, contrariamente a lo esperado, ciertos resultados más propios de sistemas "crédulos". Considérese el siguiente ejemplo:

Ejemplo 3.2 (Poole [P91]). Sean $\Delta = \{ave(x) \succ \neg vuela(x), emu(x) \succ \neg \neg vuela(x)\}$, y $\mathcal{K} = \{emu(x) \rightarrow ave(x), ave(polly) \wedge ave(tweety), emu(polly) \vee emu(tweety)\}$. Entonces son consecuencias rebatibles *vuela(polly)* y *vuela(tweety)*, y no lo son $\neg vuela(polly)$ ni $\neg vuela(tweety)$. Las sentencias inferidas son demasiado fuertes, mientras debería inferirse que uno de los dos, Polly o Tweety, no vuela.

En este ejemplo vemos que mientras pueden formarse argumentos para la inferencia de $vuela(polly)$ y $vuela(tweety)$, no es posible hallar argumentos para $\neg vuela(polly)$ ni para $\neg vuela(tweety)$. Resumidamente, un argumento o *estructura argumentativa* (o, como también los llamaremos, *argumento standard*) para h en el contexto \mathcal{K} es un par $\langle T, h \rangle_{\mathcal{K}}$, donde $T \subseteq \Delta^{\downarrow}$, tal que de $\mathcal{K} \cup T$ se infiere rebatiblemente h , y no se infiere rebatiblemente \perp . Otro requisito es que T sea mínimo para la inferencia de h (cf. Simari y Loui [SL92]). Al conjunto de todos los argumentos standard que pueden construirse en el sistema se lo denota con ‘AStruc’. Cuando no halla lugar a confusión prescindiremos del subíndice que denota el contexto en la simbolización de argumentos.

La finalidad del sistema es manejar en forma racional un cuerpo de creencias del cual puedan inferirse rebatiblemente otras creencias, pudiendo éstas ser inconsistentes entre sí. Para esto es necesario contar con una relación de preferencia que permita decidir cuáles serán los argumentos preferidos entre todos los que el agente a pueda construir. En *MTDR* esta relación es de *especificidad*. Nosotros aquí nos referiremos en general a una relación abstracta \succ , a la cual supondremos transitiva. Otra relación importante es la de *rebate a...* La idea general es que un argumento rebate a otro si las sentencias soportadas por cada uno son contradictorias entre sí, y el primero es preferido al segundo por \succ , pero no se da lo inverso.

Volvamos ahora al ejemplo 3.2. Si deseamos que en ese contexto el sistema se comporte escépticamente, dado que inevitablemente pueden darse argumentos para $vuela(polly)$ y $vuela(tweety)$, podemos intentar construir algún tipo de argumentación que, sustentada en los casos de la disyunción $emu(polly) \vee emu(tweety)$, permita bloquear las inferencias indeseadas. Hacer esto mediante la definición de una nueva relación de inferencia que se agregue a \vdash sería complicar el diseño de un sistema en principio simple. De modo que el propósito será utilizar sólo la relación \vdash , y buscar una nueva noción de *argumento por casos*.

4 Argumentos rebatibles y razonamiento por casos

El razonamiento correcto en el contexto del ejemplo 3.2 consiste en suponer qué podría inferirse si cada caso de la disyunción en cuestión perteneciera al contexto de conocimiento. Por otra parte, es importante notar que los argumentos construidos en base a los casos $emu(polly)$ y $emu(tweety)$ pueden ser utilizados para *interferir* inferencias ($vuela(polly)$, $vuela(tweety)$), pero no para *soportar* inferencias ($\neg vuela(polly)$, $\neg vuela(tweety)$).

De modo que, primero, definiremos formalmente la noción de *argumento por caso* y, luego, cuando formalicemos el proceso de construcción de todas las justificaciones del sistema, restringiremos su uso como interferentes.

Definición 4.1 Sean \mathcal{K} un contexto, Δ un conjunto de reglas rebatibles, h un literal instanciado de \mathcal{L} , y T un subconjunto de Δ^{\downarrow} . Entonces son *c-argumentos* todos y sólo aquellos que cumplen alguna de las siguientes cláusulas:

1. Si $\langle T, h \rangle_{\mathcal{K}}$ es un argumento standard, entonces $\langle T, h \rangle_{\mathcal{K}}$ es un c-argumento.
2. Si existe una sentencia $e = (e_1 \vee \dots \vee e_n)$ en \mathcal{K} tal que para algún e_i , $1 \leq i \leq n$, $\langle T, h \rangle_{\mathcal{K} \cup \{e_i\}}$ es un c-argumento para alguna sentencia h , entonces $\langle T, h \rangle_{\mathcal{K}}$ es un c-argumento.

3. Si $\langle T_1, h_1 \rangle_{\mathcal{K}}, \dots, \langle T_{n-1}, h_{n-1} \rangle_{\mathcal{K}}$ son c-argumentos, y $T_n \subseteq \Delta^\dagger$ es tal que $\langle T_n, h \rangle_{\mathcal{K} \cup \{h_1, \dots, h_{n-1}\}}$ es un c-argumento para h , entonces $\langle T, h \rangle_{\mathcal{K}}$, donde $T = \bigcup_{i=1}^n T_i$, es un c-argumento.

Ejemplo 4.1 Remitimos al contexto del ejemplo 3.2, donde, además de contar con argumentos standard para $vuela(polly)$ y $vuela(tweety)$, tenemos las siguientes c-estructuras argumentativas: $\langle \{emu(polly)\} \rangle \neg vuela(polly)$, $\neg vuela(polly)$ y $\langle \{emu(tweety)\} \rangle \neg vuela(tweety)$, $\neg vuela(tweety)$.

Al conjunto de todos los c-argumentos que pueden construirse en el sistema lo denotamos con 'AStruc_C'. Como consta en la definición 4.1, tenemos que $\text{AStruc} \subseteq \text{AStruc}_C$.

Un problema que puede surgir con esta definición es que pueden llegar a construirse c-argumentos no deseados si el conjunto \mathcal{K} no es mínimo respecto de la relación de consecuencia lógica. Esto ocurre en el siguiente ejemplo:

Ejemplo 4.2 Sea $\mathcal{K} = \{a, a \vee b\}$ y $\Delta^\dagger = \{a \succ c, b \succ \neg c\}$. Entonces $\langle \{a \succ c\}, c \rangle$ es un argumento para c , y $\langle \{b \succ \neg c\}, \neg c \rangle$ es un c-argumento para $\neg c$, puesto que $\langle \{b \succ \neg c\}, \neg c \rangle_{\mathcal{K} \cup \{b\}}$ es una estructura argumentativa. La justificación de c se ve bloqueada por este c-argumento.

El c-argumento $\langle \{b \succ \neg c\}, \neg c \rangle$ interfiere al argumento $\langle \{a \succ c\}, c \rangle$ por el hecho de estar $a \vee b$ en \mathcal{K} , mientras esta disyunción es consecuencia lógica de a . Dicha interferencia es irrelevante. Por lo tanto, es conveniente no utilizar este tipo de argumentos en sistemas donde \mathcal{K} contenga sentencias que sean inferencias no relevantes (en el sentido de las *lógicas de la relevancia*) de otras sentencias de \mathcal{K} .

Como es claro, los argumentos de AStruc_C que no pertenecen a AStruc pueden ser utilizados para interferir o bloquear, pero no para justificar la derivación de una creencia, puesto que se apoyan en hechos cuya verdad es incierta. De modo que son insuficientes para representar razonamientos por casos que *soporten* una creencia. Para esto necesitamos introducir un tipo de argumento que, a partir de una disyunción de creencias, justifique una inferencia rebatible.

Definición 4.2 Sean \mathcal{K} un contexto, Δ un conjunto de reglas rebatibles, h un literal instanciado de \mathcal{L} , y T un subconjunto de Δ^\dagger . Entonces son *d-argumentos* todos y sólo aquellos que cumplen alguna de las siguientes cláusulas:

1. Si $\langle T, h \rangle_{\mathcal{K}}$ es un argumento standard, entonces $\langle T, h \rangle_{\mathcal{K}}$ es un d-argumento.
2. Si existe una sentencia $e = (e_1 \vee \dots \vee e_n)$ en \mathcal{K} tal que para cada e_i , $1 \leq i \leq n$, existe $T_i \subseteq \Delta^\dagger$ tal que $\langle T_i, h \rangle_{\mathcal{K} \cup \{e_i\}}$ es un d-argumento para alguna sentencia h , entonces $\langle T, h \rangle_{\mathcal{K}}$, donde $T = \bigcup_{i=1}^n T_i$, es un d-argumento.
3. Si $\langle T_1, h_1 \rangle_{\mathcal{K}}, \dots, \langle T_{n-1}, h_{n-1} \rangle_{\mathcal{K}}$ son d-argumentos, y $T_n \subseteq \Delta^\dagger$ es tal que $\langle T_n, h \rangle_{\mathcal{K} \cup \{h_1, \dots, h_{n-1}\}}$ es un d-argumento para h , entonces $\langle T, h \rangle_{\mathcal{K}}$, donde $T = \bigcup_{i=1}^n T_i$, es un d-argumento.

Ejemplo 4.3 Sean $\mathcal{K} = \{a \vee b, c, d \vee e\}$ y $\Delta = \{\langle a \rangle \neg h, \langle b \rangle \neg h, \langle c \rangle h, \langle e \rangle h\}$. Entonces $\langle \{a \rangle \neg h, \langle b \rangle \neg h \rangle, \neg h \rangle$ es un d-argumento para h ; $\langle \{e \rangle h \rangle, h \rangle$ es un c-argumento para h ; y $\langle \{c \rangle h \rangle, h \rangle$ es una estructura argumentativa standard para h .

Al conjunto de todos los d-argumentos que pueden construirse en el sistema lo denotamos con 'AStruc_D'. Como consta en la definición 4.2, tenemos que $\text{AStruc} \subseteq \text{AStruc}_D$. También es importante el siguiente resultado:

Lema 4.1 $\text{AStruc}_D \subseteq \text{AStruc}_C$.

Prueba. Esto es una consecuencia directa de las definiciones 4.1 y 4.2. Sea $\langle A, h \rangle \in \text{AStruc}_D$. Si $\langle A, h \rangle$ es un argumento standard, entonces por la cláusula (1) de la definición 4.1, $\langle A, h \rangle$ es un c-argumento. Si no, si cumple la cláusula (2) de la def. 4.2, se sigue inmediatamente que cumple la cláusula (2) de la def. 4.1; y si cumple la cláusula (3) de la def. 4.2, entonces existe $A' \subseteq A$ tal que $\langle A', h \rangle$ cumple con alguna de las cláusulas (1) o (2) de la def. 4.1, y, luego, $\langle A, h \rangle$ cumple con la cláusula (3) de la def. 4.1. \square

Este lema tiene implicancias interesantes desde el punto de vista de la implementación, ya que el proceso de búsqueda de un c-argumento constituye una subrutina en la búsqueda de un d-argumento. El resultado también aporta sencillez, como veremos al final de esta sección, al proceso de búsqueda de creencias justificadas.

Los d-argumentos pueden justificar la derivación rebatible de nuevas creencias, así como interferir las derivaciones de otras. El problema que podría surgir con los c-argumentos si \mathcal{K} contuviera disyunciones inferibles no relevantemente de otras creencias en \mathcal{K} , no puede darse con los que a su vez son d-argumentos, ya que para cualquier d-argumento construido en base a una disyunción $e_1 \vee \dots \vee e_n$, existe otro d-argumento construido en base a $e_i \vee \dots \vee e_j$, para cualesquiera $1 \leq i \leq j \leq n$. A continuación definiremos el procedimiento para construir el conjunto de todos los argumentos justificados en un sistema con c-argumentos y d-argumentos.

En *MTDR*, Simari y Loui dan una definición inductiva que caracteriza el proceso por el cual los argumentos pueden rebatir y ser rebatidos por otros. En ese sistema se considera, por supuesto, sólo el conjunto AStruc de argumentos standard. Aquí describiremos el papel de los argumentos de AStruc_C y AStruc_D, introduciéndolos en dicha definición.³

Definición 4.3 Un argumento está activo como argumento de soporte o de interferencia en distintos niveles según las siguientes cláusulas:

1. Todos los argumentos de AStruc_D son argumentos de soporte de nivel-0, y todos los argumentos de AStruc_C son argumentos de interferencia de nivel-0.
2. Un argumento $\langle A_1, h_1 \rangle \in \text{AStruc}_D$ es un argumento de soporte de nivel-($n+1$) ssi no existe ningún argumento de interferencia de nivel- n $\langle A_2, h_2 \rangle$ que lo contraargumente en algún h .
3. Un argumento $\langle A_1, h_1 \rangle \in \text{AStruc}_C$ es un argumento de interferencia de nivel-($n+1$) ssi no existe ningún argumento de interferencia de nivel- n $\langle A_2, h_2 \rangle$ que lo rebata.

³Esta definición se apoya en el lema 4.1.

Definición 4.4 Un argumento $\langle A, h \rangle \in \mathbf{AStruc}_C$ justifica h si y sólo si existe un m tal que, para todo $n \geq m$, $\langle A, h \rangle$ es un argumento de soporte para h en el nivel n . Decimos que h está justificado en $\Omega \subseteq \mathbf{AStruc}_C$ si existe un argumento $\langle A, h \rangle \in \Omega$ que justifica a h .

5 Algunas características del sistema descripto

En este punto es necesario que nos detengamos a hacer algunas observaciones. Una cuestión a tener en cuenta es que la relación \vdash , en este sistema, es puramente instrumental. En todos los sistemas basados en reglas, una relación de este tipo permite sancionar un conjunto de sentencias que contiene (o al menos es igual) al conjunto de sentencias que son efectivamente justificadas, según las herramientas con que cuentan (e.g., en el caso de *MTDR*, la definición de *sentencia justificada* y sus definiciones auxiliares (cf. [SL92], p. 141).⁴ La intuición que subyace a esto es que en todo argumento (en el sentido más general del término) la conclusión debe estar conectada con las premisas mediante alguna relación de inferencia. Esto no ocurre en el sistema que hemos descripto aquí, al menos, directamente. Como es fácil notar, para cualquier argumento $\langle A, h \rangle \in \mathbf{AStruc}_C$ que no sea un argumento standard, no se da que $\mathcal{K} \cup A \vdash h$. En consecuencia, tenemos que el sistema puede justificar más sentencias que las que la relación \vdash permite derivar. Sin embargo, el conjunto de sentencias derivadas no debe necesariamente estar determinado por tal relación de inferencia: podemos ver en ella, simplemente, una herramienta auxiliar utilizada para la operación ¿Cuáles serán, entonces, las inferencias derivadas por el sistema? Serán todas aquellas sentencias para las cuales existe un argumento en \mathbf{AStruc}_D . Entonces tendremos, como en el resto de los sistemas, que el conjunto de sentencias justificadas en el sistema es un subconjunto del conjunto de sentencias que el sistema deriva.

6 Conclusión

En la representación de razonamiento rebatible, la disyunción plantea problemas fundamentalmente de tipo sintáctico, lo cual hace que en muchos sistemas quede excluida de sus formalismos. Este trabajo ha intentado mostrar un tratamiento viable en el marco formal de un sistema argumentativo.

Como resultados positivos del estudio se pueden resaltar: (a) la ampliación de la capacidad inferencial del sistema *MTDR*, al incorporar la posibilidad de razonar por casos; (b) la obtención de respuestas del sistema más racionales, hallando justificaciones para inferencias intuitivas que el sistema original no derivaba, y bloqueando otras que surgían insensatamente; (c) la caracterización clara de los tipos de argumentos que pueden formarse a partir de información disyuntiva y sus respectivos roles en la sustentación y bloqueo de inferencias; (d) la adaptación sencilla de esos argumentos al proceso formal de búsqueda de justificaciones descripto para *MTDR*.

No está dentro del alcance de este trabajo el estudio de los aspectos implementacionales del problema, cuestión que dejamos pendiente. De todos modos, como posible veta de

⁴Esto también ocurre en el sistema descripto en Bodanza y Simari [BS95] para el razonamiento por casos con argumentos. El sistema justifica un conjunto de sentencias que está incluido en el de sentencias sancionadas por la relación \vdash^d .

investigación señalamos la explotación de las herramientas de la programación lógica disyuntiva⁵ y los procesos ya sugeridos en [BS95].

Referencias

- [B95] Bodanza, G. (1995); "Razonamiento Rebatible: Dos Tendencias y Sus Perspectivas"; *Actas de las VI Jornadas de Epistemología e Historia de la Ciencia*, Córdoba.
- [BS95] Bodanza, G. y G. Simari (1995); "Argumentación Rebatible con Bases Disyuntivas"; *Actas del I Congreso Argentino de Ciencias de la Computación y II Workshop sobre Aspectos Teóricos de la Inteligencia Artificial*, Bahía Blanca (313-326).
- [LMR92] Lobo, J., J. Minker y A. Rajasekar (1992); *Foundations of Disjunctive Logic Programming*; The MIT Press, Cambridge MA.
- [L87] Loui, R. (1987); "Defeat Among Arguments: A System of Defeasible Inference"; *Computational Intelligence* 3 (3).
- [M80] McCarthy, J. (1980); "Circumscription -A Form of Non Monotonic Reasoning"; *Artificial Intelligence* 13 (27-39).
- [MD80] McDermott, D. y J. Doyle (1980); "Non-monotonic Logic I"; *Artificial Intelligence* 13 (41-72).
- [PK87] Pollock, J. (1987); "Defeasible Reasoning"; *Cognitive Science* 11 (481-518).
- [P85] Poole, D. (1985); "On The Comparison of Theories: Preferring The Most Specific Explanation"; *Proc. of the Ninth IJCAI*, Los Altos (144-147).
- [P91] Poole, D. (1991); "The Effect of Knowledge on Belief: Conditioning, Specificity and The Lottery Paradox in Default Reasoning"; *Artificial Intelligence* 49 (281-307).
- [R80] Reiter, R. (1980); "A Logic For Default Reasoning"; *Artificial Intelligence* 13 (81-132).
- [SL92] Simari, G. y R. Loui (1992); "A Mathematical Treatment of Defeasible Argumentation"; *Artificial Intelligence* 53 (125-157).

⁵cf. la obra de Lobo, Minker y Rajasekar [LMR92].