Argumentación Rebatible Con Bases Disyuntivas

Gustavo A. Bodanza Guillermo R. Simari
Grupo de Investigación en Inteligencia Artificial (G.I.I.A.)
Dpto. de Ciencias de la Computación
Universidad Nacional del Sur
Alem 1253, Bahía Blanca (8.000)
Fax: (54)(91)553933
Tel.: (54)(91)20776(ext.357 | 354)
e-mail{ccbodanz | grs}@arcriba.edu.ar

Palabras clave: *razonamiento rebatible, sistema argumentativo, base de creencias disyuntiva.*

Abstract

Este trabajo presenta un sistema de razonamiento rebatible basado en argumentos, que permite justificar una sentencia a partir de una creencia disyuntiva cuando cada una de las creencias en disyunción presupone una buena razón para sostener esa sentencia. El formalismo introducido es una extensión del sistema definido por Simari y Loui en *A Mathematical Treatment of Defeasible Reasoning*, en provecho de su potencial teorético e implementacional.

Argumentación Rebatible Con Bases Disyuntivas

1 Introducción

Presentamos un sistema de razonamiento basado en argumentos, desarrollado a partir de la estructura introducida por Simari y Loui en *A Mathematical Treatment of Defeasible Reasoning* [SL92] (o $MTDR$). Nuestro propósito es ampliar la capacidad de razonamiento de este sistema a fin de poder formular argumentos basados en creencias disyuntivas.

$MTDR$ presenta un marco de razonamiento rebatible en el cual puede representarse el conocimiento de un agente racional. A partir de la información disonible, el sistema puede construir un argumento aceptable *prima facie* para una hipótesis H, y del mismo modo pude construir argumentos que la refuten. Luego, mediante un mecanismo de comparación entre argumentos, el sistema puede decidir si la hipótesis H está o no *justificada* respecto de la información contenida en su base de creencias.

El sistema $MTDR$ posée una estructura teorética robusta, que brinda soluciones intuitivamente adecuadas para los casos que en la literatura afín se presentan como problemáticos. Consideraciones acerca de la posibilidad de razonar rebatiblemente a partir de una base conteniendo creencias disyuntivas, nos inducen a desarrollar nuevos conceptos, que definiremos a partir de los introducidos por Simari y Loui, aprovechando su potencial teorético e implementacional.

Otros interesantes desarrollos de $MTDR$ han sido y están siendo producidos: una formalización del proceso dialéctico de justificación [SCG94], y un sistema de mantenimiento y revisión de bases de argumentos [GCS93], son el resultado de sucesivas investigaciones de García *et al.*; Augusto *et al.*[AS94] extienden $MTDR$ permitiendo el razonamiento a partir de sentencias cuyos valores de verdad dependen del tiempo. Nuestra investigación se agrega a éstas con compatibilidad teorética, dejando abierta la posibilidad de su integración en un amplio sistema que las contenga.

El trabajo se estructura como sigue. En la sección 2. presentamos el sistema $MTDR$ y sus propiedades respecto de realizar inferencias a partir de creencias disyuntivas. En la sección 3. introducimos el marco de nuestro sistema con su caracterización lógica. En la sección 4. veremos las ideas para su implementación, brindando en 4.1, introductoriamente, las correspondientes a $MTDR$. En la sección 5. puntualizaremos algunas cualidades de nuestro marco.

2 El sistema $MTDR$

A continuación presentamos brevemente el marco teorético de $MTDR$, el cual sustentará las definiciones de nuestro marco. En la sección 2.1 brindamos el aspecto lógico del sistema, y en la sección 2.2 analizamos sus limitaciones respecto del comportamiento que pretenderemos del nuestro.

2.1 Características Lógicas de $MTDR$

Las creencias de un agente a son representadas en un sistema formal $\mathbb{L}$. El lenguaje de $\mathbb{L}$ está compuesto por un lenguaje de primer orden $\mathcal{L}$ más la relación metalingüística '$\succ\!\!-$' 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. Las reglas de inferencia consideradas son modus ponens y generalización. El comportamiento de las reglas rebatibles respecto de estas reglas de inferencia es el mismo que el de las implicaciones materiales.

Un conjunto consistente $\mathcal{K} \subset \mathcal{L}$ se llama *contexto* y puede ser particionado en dos subconjuntos $\mathcal{K}_N$ y $\mathcal{K}_C$ de creencias *generales* (necesarias) y *particulares* (contingentes), respectivamente.

Las creencias de a en $\mathbb{L}$ están representadas por el par $(\mathcal{K}, \Delta)$, llamado *Estructura Lógica Rebatible*, donde Δ es un conjunto de reglas rebatibles. $\Delta^{\downarrow}$ es el conjunto de todos los miembros de Δ instanciados.

Definición 2.1 Sea $\Gamma \subseteq \mathcal{K} \cup \Delta^{\downarrow}$. Un literal instanciado h es una *consecuencia rebatible* de Γ, en símbolos $\Gamma \mid\!\sim h$, si y sólo si existe una secuencia finita $A_1, \ldots, A_n$ tal que $A_n = h$ y para $1 \leq i \leq n$, se verifica $A_i \in \Gamma$, o A_i es una consecuencia directa de los elementos precedentes en la secuencia por la aplicación de alguna de las reglas de inferencia consideradas.

Definición 2.2 Dado un contexto $\mathcal{K}$, un conjunto Δ de reglas rebatibles, y un literal instanciado h de $\mathcal{L}$, decimos que un subconjunto T de $\Delta^{\downarrow}$ es una *estructura argumentativa* para h en el contexto $\mathcal{K}$,en símbolos $\langle T, h\rangle_{\mathcal{K}}$, si y sólo si

1. $\mathcal{K} \cup T \mid\!\sim h$
2. $\mathcal{K} \cup T \not\mid\!\sim \perp$
3. $\nexists T' \subset T, \mathcal{K} \cup T' \mid\!\sim h.$

Dada una estructura argumentativa $\langle T, h\rangle_{\mathcal{K}}$, decimos que T es un *argumento* para h. Un *subargumento* de $\langle T, h\rangle_{\mathcal{K}}$ es un argumento $\langle S, j\rangle_{\mathcal{K}}$ tal que $S \subseteq T$.

Ejemplo 2.1 Sea $\mathcal{K} = \{p, q, q \supset r\}$ el contexto, y sea $\Delta = \{q \succ\!\!- u, p \succ\!\!- s, s \wedge r \succ\!\!- t\}$ el conjunto de reglas rebatibles. Entonces $\langle\{p \succ\!\!- s, s \wedge r \succ\!\!- t\}, t\rangle_{\mathcal{K}}$ es una estructura argumentativa para t; y $\langle\{p \succ\!\!- s\}, s\rangle_{\mathcal{K}}$ es un subargumento de ésta.

Definición 2.3 Dos estructuras argumentativas $\langle T_1, h_1\rangle_{\mathcal{K}}$ y $\langle T_2, h_2\rangle_{\mathcal{K}}$ *discrepan*, denotado por $\langle T_1, h_1\rangle_{\mathcal{K}} \bowtie_{\mathcal{K}} \langle T_2, h_2\rangle_{\mathcal{K}}$, si y sólo si $\mathcal{K} \cup \{h_1, h_2\} \vdash \perp$.

Definición 2.4 Dadas dos estructuras argumentativas $\langle T_1, h_1\rangle_{\mathcal{K}}$ y $\langle T_2, h_2\rangle_{\mathcal{K}}$, decimos que $\langle T_1, h_1\rangle_{\mathcal{K}}$ *contraargumenta* $\langle T_2, h_2\rangle_{\mathcal{K}}$, en símbolos $\langle T_1, h_1\rangle_{\mathcal{K}} \overset{h}{\otimes\!\!\rightarrow} \langle T_2, h_2\rangle_{\mathcal{K}}$, si y sólo si existe un subargumento $\langle T, h\rangle_{\mathcal{K}}$ de $\langle T_2, h_2\rangle_{\mathcal{K}}$ tal que $\langle T_1, h_1\rangle_{\mathcal{K}} \bowtie_{\mathcal{K}} \langle T, h\rangle_{\mathcal{K}}$, *i.e.*, $\mathcal{K} \cup \{h_1, h\} \vdash \perp$.

Definición 2.5 Dadas dos estructuras argumentativas $\langle T_1, h_1\rangle_{\mathcal{K}}$ y $\langle T_2, h_2\rangle_{\mathcal{K}}$, decimos que $\langle T_1, h_1\rangle_{\mathcal{K}}$ es *estrictamente más específica* que $\langle T_2, h_2\rangle_{\mathcal{K}}$, en símbolos $\langle T_1, h_1\rangle_{\mathcal{K}} \succ_{\mathbf{spec}} \langle T_2, h_2\rangle_{\mathcal{K}}$, si y sólo si

1. para todo literal instanciado e tal que $\{e\} \cup \mathcal{K} \cup T_1 \mid\!\sim h_1$, también $\{e\} \cup \mathcal{K} \cup T_2 \mid\!\sim h_2$, y
2. existe un literal instanciado e tal que
   (a) $\{e\} \cup \mathcal{K} \cup T_2 \mid\!\sim h_2$
   (b) $\{e\} \cup \mathcal{K} \cup T_1 \not\mid\!\sim h_1$
   (c) $\{e\} \cup \mathcal{K} \not\vdash h_2$.

Definición 2.6 Dadas dos estructuras argumentativas $\langle T_1, h_1\rangle_{\mathcal{K}}$ y $\langle T_2, h_2\rangle_{\mathcal{K}}$, decimos que $\langle T_1, h_1\rangle_{\mathcal{K}}$ *rebate* a $\langle T_2, h_2\rangle_{\mathcal{K}}$ en el literal h, en símbolos $\langle T_1, h_1\rangle_{\mathcal{K}} \gg_{\mathbf{def}} \langle T_2, h_2\rangle_{\mathcal{K}}$, si y sólo si existe un subargumento $\langle T, h\rangle_{\mathcal{K}}$ de $\langle T_2, h_2\rangle_{\mathcal{K}}$ tal que

1. $\langle T_1, h_1\rangle_{\mathcal{K}}$ contraargumenta a $\langle T_2, h_2\rangle_{\mathcal{K}}$ en h, y
2. $\langle T_1, h_1\rangle_{\mathcal{K}}$ es estrictamente más específico que $\langle T, h\rangle_{\mathcal{K}}$.

Ejemplo 2.2 $\langle\{p \wedge q \succ\!\!- r\}, r\rangle_{\mathcal{K}} \gg_{\mathbf{def}} \langle\{p \succ\!\!- s, s \succ\!\!- \neg r\}, \neg r\rangle_{\mathcal{K}}$, puesto que

1. $\langle\{p \wedge q \succ\!\!- r\}, r\rangle_{\mathcal{K}} \overset{r}{\otimes\!\!\rightarrow} \langle\{p \succ\!\!- s, s \succ\!\!- \neg r\}, \neg r\rangle_{\mathcal{K}}$, y
2. $\langle\{p \wedge q \succ\!\!- r\}, r\rangle_{\mathcal{K}} \succ_{\mathbf{spec}} \langle\{p \succ\!\!- s, s \succ\!\!- \neg r\}, \neg r\rangle_{\mathcal{K}}$.

2.2 Propiedades de *MTDR*

Hemos visto los principales conceptos introducidos en *MTDR*, el tipo de inferencias que permite realizar, y cómo se construyen. Nuestro interés es ahora analizar tal sistema en procura de hallar las razones por las cuales las propiedades que buscamos no se cumplen, y de encontrar las vetas por donde introducir los arreglos necesarios de manera de no alterar con éstos las intuiciones originales de Simari y Loui.

El tipo de razonamiento que intentamos representar es el siguiente: si creemos que A es una buena razón para sostener H, a la vez que creemos que B es también una buena razón para esa hipótesis, entonces deberíamos sostener H cuando creyéramos en la verdad de $A \vee B$, aún si no pudiéramos decidir entre la verdad de A y la verdad de B. Veamos que en *MTDR* las premisas de este razonamiento estarían simbolizadas por '$A \succ\!\!- H$' y '$B \succ\!\!- H$', y la creencia en $A \vee B$ sería introducida en el contexto $\mathcal{K}$ como una creencia particular.

Dado que la interpretación de la relación ' $\succ\!\!-$ ' es incierta, su comportamiento no es del todo comparable al de la implicación material: resulta apropiada como premisa de modus ponens, pero no, por ejemplo, como premisa de una contraposición. para ver ésto, sustitúyase 'A' por 'x es una persona' y 'H' por 'x no es diabético', y por contraposición obtendríamos que los diabéticos suelen no ser personas[*] la premisa resulta razonable pero no la conclusión. Por esta razón *MTDR* no permite utilizar las reglas rebatibles como premisas de las reglas de inferencia aplicables a la imlicación material, salvo modus ponens y generalización. De modo que un argumento como

$$\langle\{A\succ\!\!-H, B\succ\!\!-H\}, H\rangle_{\mathcal{K}} \tag{1}$$

no se adecua a la definición de estructura argumentativa introducida en $MTDR$, por no darse el caso que $\{A\succ\!\!-H, B\succ\!\!-H\} \cup \{A \vee B\} \mid\!\sim H$.

La pregunta de rigor es si el tipo de razonamiento que buscamos es intuitivamente aceptable dada la naturaleza de $\succ\!\!-$, o si, por el contrario, puede conducir a conclusiones insensatas como ocurre con la contraposición. La respuesta es que si el modus ponens es aceptable, nuestro razonamiento también lo es, ya que no es otra cosa que la aplicación alternativa de un modus ponens $(A, A\succ\!\!-H \mid\!\sim H)$ u otro $(B, B\succ\!\!-H \mid\!\sim H)$, según la verdad de la disyunción venga dada por uno u otro de sus componentes. El problema se reduce entonces a si se aceptan o no creencias disyuntivas en la base de creencias. Y si se aceptan, es apropiado considerar como válidos argumentos como (1).

La idea acerca de los arreglos necesarios para nuestros fines es la siguiente: dado que $MTDR$ tiene una estructura teorética sólida y presenta solvencia frente a los casos analizados en la literatura afín, no realizaremos modificaciones internas en él, sino definiremos un sistema que lo contenga en forma íntegra y tal como fuera definido originalmente. Así, no introduciremos, por ejemplo, una meta-meta-regla que permita inferir '$\alpha \vee \beta\succ\!\!-\gamma$' a partir de '$\alpha\succ\!\!-\gamma$' y '$\beta\succ\!\!-\gamma$' como podría sugerirse en una primera aproximación, pues tal modificación supondría complicaciones tanto lógicas como implementacionales que desvirtuarían la intuición original de los autores.

3 Marco para razonar rebatiblemente a partir de creencias disyuntivas

Retomando nuestro ejemplo, supongamos que tenemos el conjunto de reglas rebatibles $\Delta = \{A\succ\!\!-H, B\succ\!\!-H\}$ y el contexto $\mathcal{K} = \{A \vee B\}$. Según está definida en $MTDR$, la relación de consecuencia rebatible '$\mid\!\sim$ ' nos permitiría formar un argumento para H sólo si A, B o ambas fueran elementos de $\mathcal{K}$. De modo que en nuestro contexto no hay argumento posible para H. Entonces tenemos dos posibilidades: 1) modificamos la definición de '$\mid\!\sim$ ' de modo que $\mathcal{K} \cup \Delta \mid\!\sim H$ sea verdadero, o 2) definimos una nueva relación de consecuencia rebatible a partir de los elementos ya existentes en $MTDR$. Lo primero no lo haremos por las razones expuestas en el parágrafo anterior. Veamos, pues, como hacemos lo segundo.

Intentaremos definir la relación buscada a partir de la relación ya definida $\mid\!\sim$. Supongamos que nuestro cuerpo de creencias es $\mathcal{K} = \{A \vee B\}$. Entonces creemos en la verdad de

[*]ejemplo debido a Delgrande et al. [DSJ94]

A o en la verdad de B o de ambas, pero no podemos decidir cuál de estas tres posibilidades nos hacen creer en la verdad de la disyunción. Pero como sabemos que al menos una entre A y B es verdadera, si sostenemos las hipótesis ' $A \succ\!\!- H$' y '$B \succ\!\!- H$' entonces podremos inferir H a partir de argumentos alternativos: en caso que A sea verdadera sabemos que $\{A\} \cup \{A \succ\!\!- H\} \mathrel{|\!\sim} H$, y en caso que B lo sea tendremos que $\{B\} \cup \{B \succ\!\!- H\} \mathrel{|\!\sim} H$. De modo que la idea es introducir una relación $\mathrel{|\!\sim}^{d}$ tal que $X \mathrel{|\!\sim}^{d} P$ exprese:X contiene una familia de conjuntos de reglas, cada uno de los cuales permite inferir P junto con algún modelo de la base de creencias en cuestión. Como los modelos que hacen verdadera a una disyunción $A \vee B$ son $\{A\}$, $\{B\}$ y $\{A, B\}$ entonces cada uno de estos conjuntos unido al conjunto de reglas $\{A \succ\!\!- H, B \succ\!\!- H\}$ permiten inferir rebatiblemente H[†] La siguiente definición formaliza esta idea:

Definición 3.1 Sea $\mathcal{K}$ un conjunto de creencias, y Δ un conjunto de reglas rebatibles. Entonces la sentencia β es una *d-consecuencia rebatible* de $\mathcal{K} \cup \Delta$, en símbolos $\mathcal{K} \cup \Delta \mathrel{|\!\sim}^{d} \beta$, si y sólo si se cumple alguna de las siguientes condiciones:

- $\mathcal{K} \cup \Delta \mathrel{|\!\sim} \beta$
- $\exists \alpha \in \mathcal{K}, \alpha = (\alpha_1 \vee \ldots \vee \alpha_n)$ tal que $\{\alpha_1\} \cup \Delta \mathrel{|\!\sim} \beta$ y ...y $\{\alpha_n\} \cup \Delta \mathrel{|\!\sim} \beta$.

Como es claro, toda consecuencia rebatible de un conjunto $\mathcal{K}$ es tambíen una d-consecuencia rebatible (pero no, obviamente, lo converso), esto es, $\mathrel{|\!\sim}_{\mathcal{K}} \subseteq \mathrel{|\!\sim}^{d}_{\mathcal{K}}$.

Veamos ahora cómo introducir la noción de argumento en este marco. De forma similar a *MTDR*, que define 'estructura argumentativa' a partir de $\mathrel{|\!\sim}$, nosotros definimos 'd-estructura argumentativa' a partir de $\mathrel{|\!\sim}^{d}$.

Definición 3.2 Dado un contexto $\mathcal{K}$ y un conjunto de reglas rebatibles Δ decimos que un subconjunto T de $\Delta^{\downarrow}$ es un *d-argumento* para h en el contexto $\mathcal{K}$, denotado por $\langle T, h\rangle^{d}_{\mathcal{K}}$, si y sólo si

1. $\mathcal{K} \cup T \mathrel{|\!\sim}^{d} h$
2. no es el caso que $\mathcal{K} \cup T \mathrel{|\!\sim}^{d} \bot$
3. $\nexists T' \subset T, \mathcal{K} \cup T' \mathrel{|\!\sim}^{d} h$ (T es minimal).

A partir de esta definición y de la definición de estructura argumentativa en $MTDR$ se puede mostrar que si $\langle T, h\rangle_{\mathcal{K}}$ es una estructura argumentativa, entonces también es una d-estructura argumentativa (basta recordar que $\mathrel{|\!\sim}_{\mathcal{K}} \subseteq \mathrel{|\!\sim}^{d}_{\mathcal{K}}$). En adelante llamaremos *argumentos standard* a las estructuras argumentativas válidas en $MTDR$. Las relaciones de discrepancia, contraargumentación, especificidad y rebatimiento se introducen en este marco sustituyendo en sus respectivas definiciones -dadas en $MTDR$- cada ocurrencia de un argumento standard por su correspondiente d-argumento, y cada ocurrencia de '$\mathrel{|\!\sim}$ ' por '$\mathrel{|\!\sim}^{d}$ '.

[†]Nótese que si el conjunto de reglas fuera $\Delta = \{A \succ\!\!- H\}$ habría un modelo de $\mathcal{K}$, a saber $\{B\}$, tal que $\{B\} \cup \Delta \not\mathrel{|\!\sim} H$, y por lo tanto no sería el caso que $\mathcal{K} \cup \Delta \mathrel{|\!\sim}^{d} H$.

Ejemplo 3.1 (*d-estructura argumentativa*). Sean $\mathcal{K} = \{e, f, a \vee b\}$ y $\Delta = \{a \succ\!\!- \neg h, b \succ\!\!- \neg h, c \succ\!\!- h, e \wedge f \succ\!\!- h, e \succ\!\!- c\}$. Entonces $\langle\{a \succ\!\!- \neg h, b \succ\!\!- \neg h\}, \neg h\rangle^d_{\mathcal{K}}$ es una d-estructura argumentativa para h; $\langle\{e \wedge f \succ\!\!- h\}, h\rangle^d_{\mathcal{K}}$ y $\langle\{e \succ\!\!- c, c \succ\!\!- h\}, h\rangle^d_{\mathcal{K}}$ son d-estructuras argumentativas para h; $\langle\{e \wedge f \succ\!\!- h, e \succ\!\!- c, c \succ\!\!- h\}, h\rangle^d_{\mathcal{K}}$ *no* es una d-estructura argumentativa válida (por no ser minimal); y ninguna d-estructura argumentativa para h es más específica que cualquiera d-estructura argumentativa para $\neg h$ y viceversa.

Ejemplo 3.2
(*especificidad*). Sean $\mathcal{K} = \{a \vee b, a \vee b \vee c\}$ y $\Delta = \{a \succ\!\!- h, b \succ\!\!- h, c \succ\!\!- d, d \wedge h \succ\!\!- f\}$. Entonces $\langle\{a \succ\!\!- h, b \succ\!\!- h\}, h\rangle^d_{\mathcal{K}} \succ_{\mathbf{spec}} \langle\{a \succ\!\!- h, b \succ\!\!- h, c \succ\!\!- d, d \wedge h \succ\!\!- f\}, f\rangle^d_{\mathcal{K}}$.

Ejemplo 3.3 (*rebatimiento*). Sean $\mathcal{K} = \{a \vee b, a \vee b \vee c, c \wedge d\}$ y $\Delta = \{a \succ\!\!- h, b \succ\!\!- h, c \succ\!\!- d, d \wedge h \succ\!\!- f, c \wedge d \succ\!\!- \neg h\}$. Entonces $\langle\{c \wedge d \succ\!\!- \neg h\}, \neg h\rangle^d_{\mathcal{K}_1} \gg_{\mathbf{def}} \langle\{a \succ\!\!- h, b \succ\!\!- h, c \succ\!\!- h\}, h\rangle^d_{\mathcal{K}_1}$, pero no es el caso que $\langle\{c \wedge d \succ\!\!- \neg h\}, \neg h\rangle^d_{\mathcal{K}_1} \gg_{\mathbf{def}} \langle\{a \succ\!\!- h, b \succ\!\!- h\}, h\rangle^d_{\mathcal{K}_1}$.

Como resultado de las nociones definidas, sentencias que en $MTDR$ no cuentan con soporte argumentativo pueden encontrar sustentación en este marco más general. Esto a su vez tiene una serie de implicaciones. Primero, en el mismo contexto, cualquier justificación dada por $MTDR$ es una d-estructura argumentativa en nuestro sistema. Segundo, una d-estructura argumentativa (que no sea un argumento en $MTDR$) puede rebatir a una justificación dada por $MTDR$ sólo si la misma disyunción que activa a la primera activa también a la segunda. Cabe aclarar que esto último puede ocurrir sólo en extraños casos como el siguiente:

Ejemplo 3.4 Sea $\mathcal{K} = \{a \vee b\}$ y $\Delta = \{a \vee b \vee c \succ\!\!- h, a \succ\!\!- \neg h, b \succ\!\!- \neg h\}$. Entonces $\langle\{a \succ\!\!- \neg h, b \succ\!\!- \neg h\}, \neg h\rangle^d_{\mathcal{K}} \gg_{\mathbf{def}} \langle\{a \vee b \vee c \succ\!\!- h\}, h\rangle^d_{\mathcal{K}}$. Nótese que el argumento refutador no es un argumento standard mientras que el refutado sí lo es. Por lo tanto, dada la consulta h? en este contexto, la respuesta de $MTDR$ es *sí* y la de nuestro sistema es *no*. Lo que hace más extraño al caso es que ninguna de las dos respuestas resulta intuitiva. Esto puede solucionarse (en ambos sistemas) sustituyendo cada regla de la forma '$A_1 \vee \ldots \vee A_n \succ\!\!- B$' por un conjunto de reglas $\{A_1 \succ\!\!- B, \ldots, A_n \succ\!\!- B\}$. De este modo, ninguno de los dos sistemas podría decidir una respuesta: en el caso de $MTDR$ porque no podría formar argumentos ni para h ni para $\neg h$; y en el caso de nuestro sistema, porque el único argumento que podría formar para h (*viz.*, $\langle\{a \succ\!\!- h, b \succ\!\!- h\}, h\rangle^d_{\mathcal{K}}$) y el único argumento que podría formar para $\neg h$ (*viz.*, $\langle\{a \succ\!\!- \neg h, b \succ\!\!- \neg h\}, \neg h\rangle^d_{\mathcal{K}}$) se bloquearían entre sí, sin ser ninguno de ellos preferible por especificidad.

4 Respecto de la implementación del sistema

4.1 La implementación computacional de $MTDR$: el sistema jf

Así como hemos introducido nuestro marco teorético como una extensión de las nociones definidas previamente en $MTDR$, también la implementación práctica de nuestro sistema está basada en una extensión del lenguaje utilizado por Simari y Loui en su sistema **jf**. El sistema **jf** está definido en el marco del lenguaje Prolog de programación lógica. Brindamos a continuación sus ideas básicas (supondremos la familiaridad del lector con los rudimentos de la programación lógica).

Definición 4.1 Una *cláusula definida* es una cláusula de la forma

$$B \Leftarrow A_1, \ldots, A_n$$

con sólo un átomo como consecuente. El consecuente B es llamado *cabeza* y el antecedente $A_1, \ldots, A_n$ el *cuerpo* de la cláusula.

Definición 4.2 Una *meta definida* es una cláusula de la forma

$$\Leftarrow A_1, \ldots, A_n$$

donde cada A_i se denomina *submeta*. Una cláusula *unitaria* es una cláusula de la forma

$$B \Leftarrow true$$

a la que llamamos *hecho* (*fact*).

Definición 4.3 Una cláusula *de Horn* es una cláusula definida o una cláusula unitaria.

Simari y Loui extienden el lenguaje introduciendo las dos entidades siguientes:

Definición 4.4 Una cláusula *rebatible* es una cláusula de la forma

$$B \multimap A_1, \ldots, A_n$$

con sólo un átomo como consecuente.

Definición 4.5 La relación *neg* es un prefijo formando parte del nombre de un átomo (no es un operador).

La relación *neg* permite representar (la creencia en) hechos negativos. Los hechos negativos se relacionan con los afirmativos del modo usual. El sistema **jf** reconoce que $neg\ negA = A$, y trata al átomo $negA$ del mismo modo que a cualquier otro átomo; *i.e.*, $negA$ se considera una consecuencia de un conjunto de cláusulas definidas y rebatibles R si y sólo si $negA$ es derivable de R via un número finito de aplicaciones de modus ponens (de ningún modo esta relación se conecta con el operador prefijo `not` definido en Prolog). Un átomo $negA$ puede aparecer en cualquier parte de una cláusula definida o rebatible.

Finalmente, una base de conocimiento es definida como sigue:

Definición 4.6 Una *base de conocimiento* $\mathbb{K}$ es un conjunto finito de cláusulas definidas y rebatibles que pueden contener átomos afectados por la relación *neg*. $\mathbb{K}$ es el equivalente a lo que fue llamado Estructura Lógica Rebatible, representando en él al conjunto $\mathcal{K}$ por medio de cláusulas definidas, y al conjunto Δ por medio de cláusulas rebatibles.

El aspecto procedimental de **jf** es el siguiente: su input es una base de conocimiento $\mathbb{K}$ y una consulta específica $Q?$. Q es una instancia específica de un átomo, pudiendo llevar el prefijo *neg*. El sistema intenta encontrar, buscando en $\mathbb{K}$, un argumento no rebatible para Q. Si encuentra esa justificación, el output del sistema será una estructura argumentativa para Q, más todos los posibles rebatidores. En este caso la respuesta será afirmativa. Si para cada argumento encontrado para Q se encuentra al menos un rebatidor entonces la respuesta es negativa; y es indefinida si no se encuentra un argumento para Q, o si para cada argumento encontrado hay un contraargumento, entre los cuales no hay preferencia por especificidad.

Los argumentos para Q son construídos por *backward chaining* sobre $\mathbb{K}$. La reducción de una meta G es el reemplazo de G por una instancia específica de una cláusula (definida o rebatible) cuya cabeza es idéntica a G. El encadenamiento finaliza cuando G es soportado por una cláusula unitaria $G \Leftarrow true$. Los rebatidores para un argumento son buscados del mismo modo a partir de la negación de los átomos que ocurren en el argumento original. Finalmente el sistema compara el argumento con sus contraargumentos por especificidad, utilizando las sentencias que activan al primero (*i.e.*, las sentencias que junto con las reglas del argumento permiten inferir Q mediante un número finito de aplicaciones de modus ponens) y las sentencias que activan a los segundos. (Para más detalles del sistema remitimos a Simari y Loui[SL92]).

4.2 Implementación de la teoría introducida: el sistema jf^d

Para la implementación de nuestro marco teorético necesitaremos extender el lenguaje utilizado para definir **jf**. Así como la sintaxis del lenguaje Prolog resulta apropiada como base para definir **jf**, para el nuestro resultará apropiada la sintaxis del lenguaje de programación lógica disyuntiva [‡]

Definición 4.7 Una cláusula *disyuntiva* es una fórmula de la forma

$$A_1 \vee \ldots \vee A_n$$

donde cada A_i es un literal, y todas las variables que ocurren en ellos se consideran universalmente cuantificadas. Una instancia específica de una cláusula disyuntiva está formada con literales instanciados. A fines de incrementar el poder expresivo del lenguaje resulta apropiado permitir a los A_i ser conjunciones de literales.

Definición 4.8 Una cláusula instanciada $A = A_1 \vee \ldots \vee A_n$ es una *subcláusula* de una cláusula instanciada $B = B_1 \vee \ldots \vee B_m$ si para cada A_i, $1 \leq i \leq n$, hay un B_j tal que $A_i = B_j$. La cláusula A se llama *subcláusula propia* de B si es una subcláusula de B y hay un B_j, $1 \leq j \leq m$, tal que para todo A_i, $B_j \neq A_i$.

[‡]Quienes no estén familiarizados con este lenguaje, pueden consultar la obra de Lobo *et al.*[LMR92], de donde tomamos las definiciones fundamentales que a continuación mostramos.

Definición 4.9 Una *aserción disyuntiva* es una cláusula indefinida (*i.e.*,una cláusula con más de un literal en la cabeza) sin cuerpo. Nosotros escribiremos las aserciones disyuntivas utilizando, como en **jf**, el átomo especial *true*:

$$A_1 \vee \ldots \vee A_n \Leftarrow true$$

Las aserciones disyuntivas representan hechos o creencias disyuntivas. Estas son las únicas cláusulas conteniendo disyunciones en su cabeza que permitiremos en nuestro sistema, en primer lugar, para no alterar el espíritu de **jf**, y en segundo lugar, para evitar computaciones expensivas e innecesarias. Con estos elementos definidos podremos representar una base de creencias conteniendo creencias disyuntivas.

Definición 4.10 Una *base de creencias* $\mathbb{K}^d$ es un conjunto finito de cláusulas definidas, cláusulas rebatibles y aserciones disyuntivas.

El conjunto $\mathcal{K}$ de una Estructura Lógica Rebatible será representado en $\mathbb{K}^d$ mediante cláusulas definidas y aserciones disyuntivas, mientras Δ será representado utilizando cláusulas rebatibles. Se prohibirá (a modo de *navaja de Ockham*) que $\mathbb{K}^d$ contenga explícitamente aserciones disyuntivas si contiene otras que sean subcláusulas de éstas[§]

A continuación exponemos las características procedimentales de $\mathbf{jf}^d$.

4.3 Búsqueda de justificaciones

El mérito de $\mathbf{jf}^d$ es encontrar ciertos argumentos que **jf** no puede encontrar, a saber, aquellos que son activados por sentencias disyuntivas. Un argumento tal puede ser, *e.g.*, $\langle\{a \succ\!\!- c, b \succ\!\!- c\}, c\rangle^d_{\mathcal{K}}$ donde $\mathcal{K} = \{a \vee b\}$. Pero supongamos que a y b fueran miembros de $\mathcal{K}$: en este caso tendríamos dos argumentos que serían más específicos que el anterior: $\langle\{a \succ\!\!- c\}, c\rangle^d_{\mathcal{K}}$ y $\langle\{b \succ\!\!- c\}, c\rangle^d_{\mathcal{K}}$. Ambos son argumentos que **jf** puede encontrar. Entonces es conveniente utilizar a **jf** como procedimiento interno del sistema $\mathbf{jf}^d$. Así, haremos que el proceso de búsqueda de un argumento no standard para una consulta Q? comience sólo en caso de no hallar un argumento standard para Q mediante **jf**.

El input de $\mathbf{jf}^d$ será una base de creencias $\mathbb{K}^d$ y una consulta Q? que consistirá de un átomo instanciado. Esta consulta invocará al procedimiento **jf** , que buscará en $\mathbb{K}^d$ mediante *backward chaining* un argumento para Q, reduciendo metas (como fue descripto antes) hasta encontrar una cláusula de la forma $G_i \Leftarrow true$. Si de esta forma, mediante **jf**, no puede encontrarse un argumento para Q, entonces comenzará la búsqueda de un argumento soportado por alguna cláusula disyuntiva.

El proceso de esta búsqueda se divide en dos partes: la primera es idéntica al funcionamiento de **jf**, sólo que en lugar de detener el proceso de reducción de metas al encontrar una cláusula de la forma $G_i \Leftarrow true$, lo hará cuando encuentre una aserción disyuntiva $A_1 \vee \ldots \vee A_n \Leftarrow true$ tal que algún A_i sea idéntico a G_i. Esto puede hacerse reemplazando la meta G_i por la orden `subclau(`$\{G_i\}$`,`X`)?` que obligue al sistema a buscar una aserción disyuntiva $X \Leftarrow true$ que contenga como subcláusula a G_i, y continuar el proceso de encadenamiento si ésta no es hallada.

[§] *i.e.*, cláusulas disyuntivas que sean consecuencias lógicas de otras cláusulas disyuntivas pertenecientes a $\mathbb{K}^d$.

La segunda parte se activa cuando una aserción disyuntiva $C = C_1 \vee \ldots \vee C_n$ es encontrada. En este caso, habrá que verificar si existe para cada C_j $(1 \leq j \leq n)$ una cadena de reglas que la conecte con G_i, o alguna de las metas reducidas antes de G_i. Esto garantizará que el argumento resultante sea activado por cualquier C_j que haga verdadera a C.

Dada la prohibición de que $\mathbb{K}^d$ contenga una aserción disyuntiva si contiene alguna otra que es subcláusula de ésta, la búsqueda no será en ningún caso ociosa.[¶] Si una búsqueda exhaustiva no da en hallar una d-estructura argumentativa para Q, entonces la respuesta quedará indecisa. Si, en cambio, una d-estructura argumentativa para Q es encontrada, entonces todo el proceso recomenzará, esta vez para hallar un rebatidor del argumento encontrado. Finalmente, los argumentos y contraargumentos para Q son testeados por especificidad a fin de decidir la respuesta.

Los siguientes ejemplos ilustran el funcionamiento del sistema $\mathbf{jf}^d$.

Ejemplo 4.1 Sea $\mathbb{K}^d = \{e \multimap\!\!< c \wedge d, d \multimap\!\!< a, d \multimap\!\!< b, c \Leftarrow true, a \vee b \Leftarrow true\}$ y e? una consulta. El proceso comienza reduciendo la meta e, que remite a la reducción de las submetas c y d. La submeta c es reducida al encontrar $c \Leftarrow true$. Luego el sistema buscará exhaustivamente la reducción de d, sin encontrar ninguna cláusula unitaria donde detenerse luego de intentar reducir a y reducir b. Por lo tanto comenzará la búsqueda de una d-estructura argumentativa para d, escogiendo no-determinísticamente entre a y b. Supongamos que comience por a. Entoces se activará la orden `subclau({a},X)`, mediante la cual encontrará $\{X{=}a \vee b \Leftarrow true\}$. Luego verificará que existe una conexión entre d y b (la otra componente de la disyunción), al hallar $d \multimap\!\!< b$. Con esto, finalmente, habrá hallado el argumento $\{e \multimap\!\!< c \wedge d, d \multimap\!\!< a, d \multimap\!\!< b, c \Leftarrow true, a \vee b \Leftarrow true\}$.

Ejemplo 4.2 Sea $\mathbb{K}^d = \{d \multimap\!\!< a, d \multimap\!\!< b, neg\ d \multimap\!\!< c, b \Leftarrow true, c \Leftarrow true, a \vee b \Leftarrow true\}$ y d? una consulta. Entonces $\mathbf{jf}^d$ encontrará primero el argumento $\{d \multimap\!\!< a, d \multimap\!\!< b, a \vee b \Leftarrow true\}$, y luego el contraargumento $\{neg\ d \multimap\!\!< c, c \Leftarrow true\}$. Este último es más específico que el primero (debido a la presencia de '$b \Leftarrow c$' en $\mathbb{K}^d$), por lo cual la respuesta a la consulta será negativa.

5 Consideraciones finales

Hemos desarrollado un marco que amplía la capacidad de razonamiento del sistema $MTDR$, al permitir la construcción de argumentos sustentados por premisas disyuntivas. Todos los arreglos realizados sobre ese sistema son externos, de modo que sus ideas conceptuales, así como su robustez, quedan inalteradas.

En su aspecto formal, el conjunto de argumentos válidos en nuestro sistema contiene al conjunto de todos los argumentos válidos en $MTDR$; y en la diferencia entre ambos radica la posibilidad de que justificaciones dadas por uno no sean tales en el otro. Respecto de su implementación, esto también representa una ventaja: toda justificación encontrada por **jf** puede considerarse en $\mathbf{jf}^d$ como fundamento para una respuesta provisoria.

[¶]cf. ejemplo 3.2: la disyunción $a \vee b \vee c$ no puede ser explícitamente representada en $\mathbb{K}^d$; por lo tanto, el argumento para f (menos específico que el argumento para h) no puede ser construído por $\mathbf{jf}^d$.

Por último, y atendiendo a un interés más general, el presente trabajo puede ser visto como un experimento de combinación entre un sistema de razonamiento rebatible y la programación lógica disyuntiva.

Referencias

[AS94] Augusto, J.C. y G. Simari (1994);"Un Sistema Argumentativo con Referencias a Momentos de Tiempo"; *Proc. de la 23 JAIIO*, Septiembre 1994 (81-92).

[DSJ94] Delgrande, J., S. Torsten y W.K. Jackson (1994);"Alternative Approaches to Default Logic"; *Artificial Intelligence* 70 (167-237).

[GCS93] García, A., C. Chesñevar y G. Simari (1993); "Making Argument Systems Computationally Attractive"; *Proceedings of the XIII International Conference of the Chilean Society for Computer Science*, October 1993.

[LMR92] Lobo, J., J. Minker y A. Rajasekar (1992); *Foundations of Disjunctive Logic Programing*; The MIT Press, Cambridge MA.

[SCG94] Simari G., C. Chesñevar y A. García,(1994); "The Role of Dialectics in Defeasible Argumentation"; *Proceedings of the XIV International Conference of the Chilean Society for Computer Science*, October 1994.

[SL92] Simari, G. y R. Loui (1992);"A Mathematical Treatment of Defeasible Argumentation"; *Artificial Intelligence* 53 (125-157).