CDD: 160

Argumentos Máximamente Específicos en Argumentación Rebatible

Maximally Specific Arguments in Defeasible Argumentation

Claudio Andrés Alessio

Instituto de Investigaciones Económicas y Sociales del Sur
San Andrés 800 - Altos de Palihue Bahía Blanca
Buenos Aires 8000
Argentina
claudioalessio@uccuyo.edu.ar

Received: 04.03.2016; Revised: 12.07.2016; Accepted: 26.07.2016

DOI: http://dx.doi.org/10.1590/0100-6045.2016.V39N2.CAA

Agradezco a los revisores anónimos que han hecho críticas y sugerencias útiles que permitieron mejorar este trabajo.

Este trabajo ha sido realizado bajo el proyecto PICT 2013-1489 de la Agencia Nacional de Promoción Científica y Tecnológica (ANPCyT), Argentina.

Resumen: *DeLP* (*Defeasible Logic Programming*) is a defeasible argumentation system that captures common sense reasoning features. Examples proposed in the literature show that *DeLP* gets counterintuitive results. We suggest a possible cause of this problem and we propose an approach to neutralize it. The approach is based on the pre-selection criterion of arguments, called maximal specificity. The criterion establishes that an argument will be considered like maximally specific only if, for every argument based on more specific evidence which does not explain the same conclusion, it explains neither the contrary nor something more general.

Palabras clave: Razonamiento Default, Especificidad, Restablecimiento, Derrota.

Abstract: *DeLP* (*Defeasible Logic Programming*) es un sistema de argumentación rebatible que captura razonamientos de sentido común. Algunos ejemplos propuestos en la literatura muestran que el sistema obtiene resultados contraintuitivos. Una posible causa de tales resultados es propuesta y se sugiere una manera de neutralizarla. Esto se hace mediante una preselección de los argumentos en base a un criterio denominado máxima especificidad. El criterio consiste básicamente en lo siguiente, un argumento será considerado como máximamente específicos solo si, para cualquier argumento basado en evidencia más específica que no explique la misma conclusión, este no explica lo contrario ni algo más general.

Keywords: Default Reasoning, Specificity, Reinstatement, Defeat.

1. Introducción

Un razonamiento default (*default reasoning*) es un razonamiento que da buenas pero tentativas razones para la conclusión que sustenta, aunque puede verse retractado frente a información adicional.

Este tipo de razonamientos se corresponde, según Reiter (1980), con el proceso de derivar conclusiones basadas en un patrón de inferencia de la forma: *dada la información 'x' y en ausencia de información contraria, concluir 'y'*. Por ejemplo, si se sabe que Tweety es un ave, y se tiene en cuenta que, en general, las aves vuelan, y en ausencia de información contraria se puede concluir que Tweety vuela, entonces concluya que Tweety vuela.

Como ya se dijo, este patrón de inferencia es sensible al arribo de nueva información porque está sujeto a excepciones. Las excepciones para el ejemplo propuesto pueden ser casos como que el ave en cuestión está enferma, o tiene un ala rota, o está embalsamada, o es un ave no voladora. Los razonamientos default son útiles, aunque inseguros, para obtener información a pesar de la ausencia de conocimiento total sobre el mundo. Permiten completar los *huecos* en el conocimiento, son razonamientos que autorizan a inferir a pesar de la ausencia de conocimiento total sobre el mundo.

Uno de los primeros formalismos que pretendió modelar este tipo de razonamiento fue la Lógica Default (*Default Logic*) propuesta por Reiter (1980). La Lógica Default distingue los hechos prototípicos (*prototypical facts*) de los hechos estrictos (*hard facts*) sobre el mundo, i.e. distingue situaciones como '*por lo general los mamíferos no vuelan*' de '*todos los perros son mamíferos'*. Los hechos prototípicos, pueden ser vistos como reglas de inferencia llamadas reglas defaults. Las reglas defaults permiten extender plausiblemente, aunque no infaliblemente, las conclusiones que pueden obtenerse a partir de las reglas estrictas.

Reiter sostuvo que es posible pensar una teoría sobre la realidad, llamada teoría default (*default theory*), constituida por un conjunto de conocimientos ciertos sobre el mundo y un conjunto de conocimientos rebatibles a partir de la cual pueden obtenerse *extensiones* (*extensions*), i.e. conclusiones plausibles. De una teoría default se pueden obtener extensiones mutuamente excluyentes (*múltiples extensiones*), esto es así porque la información incompleta puede llevar a la aparición de inconsistencias que es preciso manejar de algún modo.

Reiter, inicialmente, sostuvo que las múltiples extensiones de una teoría pueden ser vistas como maneras alternativas de completar el conocimiento sobre el mundo. Por ello tenía sentido que el sistema no determine cuál de las extensiones mutuamente excluyentes sea elegida o preferida. Las extensiones mutuamente excluyentes o inconsistentes eran encapsuladas en conjuntos distintos de *consecuencias lógicas y default de la teoría*.

A pesar de que la idea de Reiter sobre cómo interpretar las múltiples extensiones parece razonable, prontamente fue advertido que bajo ciertas situaciones no es correcto que se obtengan ciertas extensiones, más en particular, Reiter y Criscuolo (1981) evidenciaron que la Lógica Default sancionaba como aceptables conjuntos de extensiones que simplemente eran contraintuitivas.

El problema detectado por Reiter y Criscuolo en la Lógica Default, la obtención de extensiones intuitivamente incorrectas, puede ser ilustrado con el famoso ejemplo de Tweety citado anteriormente. Supóngase que la teoría default cuenta con la siguiente información: '*Por lo general las aves vuelan*' (rd_1), '*por lo general los pingüinos no vuelan*' (rd_2), '*Todos los pingüinos son aves*' (re_1), '*Tweety es pingüino'* (h_1). Teniendo en cuenta h_1 y re_1, es claro que Tweety es ave. Ahora bien, teniendo en cuenta tal consecuencia, en conjunto con la regla rd_1 se obtiene la *extensión rebatible* '*Tweety vuela*'. A partir de la misma información también puede sancionarse *la conclusión rebatible* '*Tweety no vuela*', sustentada en la regla rd_2 y el hecho h_1. Ahora bien, dada la información disponible es claro que '*Tweety es un ave voladora*' no es una conclusión intuitivamente correcta puesto que ya se sabe que Tweety es un pingüino y que por lo tanto se trata de un ave no voladora, de una de esas aves excepcionales. La lógica de Reiter es incapaz de bloquear casos como estos. En tal sistema ambas conclusiones serán aceptadas, i.e. "Tweety vuela", "Tweety no vuela", lo que es claramente incorrecto. Para resolverlo se propusieron varias estrategias que *mutatis mutandis* apelan al principio de especificidad (*specificity*), el principio que establece que una subclase es preferible (*o cancela o bloquea*) a la superclase en caso de conflicto entre ellas. En el ejemplo, un sistema dotado del criterio de especificidad, elegiría aquella conclusión sustentada por la información más específica, en este caso: *Tweety no vuela*.

Diversos autores han pretendido definir formalmente el criterio de especificidad en la literatura. El primer significado preciso de especificidad fue dado por Touretzky (1984). Su definición fue propuesta en el marco de un sistema de herencias con excepciones. La idea consiste en lo siguiente: Si un objeto '*X*' hereda de un objeto '*Y*' la propiedad '*p*' y también hereda de otro objeto '*Z*', la propiedad '*no p*', entonces el objeto '*X*' conservará la propiedad '*p*' si y sólo si existe un camino, i.e. una *prueba*, de '*X*' hacia '*Z*' vía '*Y*' pero no viceversa. '*X*' conservará la propiedad '*p*' si y sólo si '*Y*' es una subclase de '*Z*', porque la subclase es siempre preferida. Retomando el reciente ejemplo, bajo esta perspectiva puede ser interpretado de la siguiente manera: Tweety, a partir del hecho de que es pingüino, hereda la propiedad de que no vuela, mientras que, a partir de que es ave, hereda la propiedad de volar. Dado que '*ser pingüino*' es una subclase de '*las aves*', será siempre preferida la subclase y el sistema de herencia propuesto por Touretzky obtendrá como conclusión "*Tweety no vuela*" y, como es esperable, la conclusión '*Tweety vuela*' no es obtenida en el sistema.

Por otro lado, Poole (1985) propuso un criterio de especificidad para resolver teorías complementarias en el marco del sistema *Theorist* por él propuesto. En tal sistema, las reglas defaults son utilizadas como posibles hipótesis en una teoría que busca explicar o predecir un estado de cosas. Tal estado de cosas es predecible o explicable si se sigue lógicamente de un conjunto consistente de instancias de reglas default junto con hechos observables e instancias de un conjunto de sentencias de primer orden. Podría suceder que un estado de cosas '*h*' se encontrara sustentado por un conjunto '*D*' pero al mismo tiempo, el sistema podría contar con un conjunto '*D**' que sustenta '*h**', donde '*h**' consiste en la negación de '*h*'. Frente a tal situación Poole define el criterio de *teoría más específica* (*most specific theory*) como mecanismo para determinar si alguna de las teorías es mejor. Sucintamente, tal criterio consiste en lo siguiente: Una teoría '*D*' es más específica que '*D**' si cada vez que '*D**' es aplicable, '*D*' también lo es, y existen casos en que '*D*' es aplicable, pero no lo es '*D**'.

Ya en el terreno de la argumentación rebatible, que más adelante será presentado brevemente, Loui (1987) propuso el empleo de la especificidad como un criterio de comparación de argumentos en conflicto. La especificidad, desde el punto de vista de Loui puede expresarse de la siguiente manera. Si A y B son argumentos donde las conclusiones Con(A) y Con(B) de ambos argumentos son inconsistentes, y las razones que sustentan Con(A) permite deducir las razones de Con(B) entonces A es tan específico como B, y si B no es tan especifico como A, entonces A es más específico que B.

En el sistema *MTDR* (Simari y Loui, 1992), la especificidad permite también dirimir argumentos que no pueden ser aceptados conjuntamente. La idea general consiste en establecer que un argumento A será mejor que otro B cuando A se base en mayor evidencia que B o que A emplee menos reglas que B. Intuitivamente, un argumento A será más específico que B cuando todo lo que activa a A, activa también a B, pero algo que activa a B, no activa a A.

Adicionalmente a la manera de implementar especificidad destacada recientemente, hay que incorporar aquellos abordajes que lo hacen mediante una comparación entre las reglas, y donde la especificidad es sancionada en el lenguaje. Sistemas que implementan este abordaje son, por ejemplo, los propuestos por Prakken y Sartor (1996), Simari y García (2004), Dung y Son (2000), o Brewka (1994) entre otros.

Un estudio general de la especificidad en sistemas que modelan razonamiento default puede encontrarse en (Benferhat, 2000) y un estudio sobre los principios que gobiernan el empleo de este criterio en (Moinard, 1990).

La especificidad, en tanto principio aplicado al razonamiento default, permite dirimir situaciones conflictivas. Ahora bien, parece ser que su empleo, que ha permitido resolver problemas en otros terrenos, especialmente en *Lógica Default* o *Theorist*, trae aparejada la aparición de situaciones problemáticas cuando es implementada en los sistemas argumentativos. Antes de señalar el problema se brindará brevemente una introducción de este enfoque.

Los sistemas argumentativos, conocidos también como argumentación rebatible, es un enfoque inaugurado con los trabajos de Pollock (1987) y Loui (1987). En esta área se proponen una serie de formalismos que modelan información tentativa y potencialmente contradictoria mediante la construcción, comparación y evaluación de argumentos a favor y en contra de ciertas afirmaciones. La búsqueda final, en tales sistemas, consiste en la identificación de aquellos argumentos que finalizado el proceso, pueden considerarse como los ganadores, o los que prevalecen frente a las objeciones que han sido planteadas. Las conclusiones de los argumentos vencedores se consideran las extensiones del sistema.

Estos sistemas pueden ser caracterizados a través de un proceso de varias etapas o fases: construcción de argumentos, marco argumentativo, selección de los mejores argumentos, determinación de las conclusiones justificadas.

Los argumentos, en tanto entidades que sustentan conclusiones, se construyen a partir de una base de conocimiento previamente especificada en un lenguaje formal determinado y de las condiciones que un argumento debe satisfacer (*fase 1*).

Una vez establecido el conjunto de argumentos, puede suceder que dos o más de ellos no puedan ser simultáneamente aceptados. Con vistas a determinar qué argumento prevalece se apela a diversas relaciones que entre ellos pueden darse. Una de las relaciones más importantes recibe el nombre de derrota (Dung, 1995). El conjunto de argumentos y las relaciones de derrota que entre ellos se dan originan lo que Dung (1995) ha denominado '*marco argumentativo*' (*fase 2*).

El principal problema para un sistema argumentativo, tal como ya se ha señalado brevemente, consiste en determinar qué argumentos, de todos los construidos, pueden considerarse como los ganadores de la disputa, y cuáles los perdedores, en términos un poco más precisos, cuáles de los argumentos pueden considerarse como justificados y cuáles como injustificados. Para ello, luego de que el marco argumentativo ha sido establecido se procede a seleccionar aquellos argumentos que prevalecen frente a sus rivales, tales argumentos o no están derrotados o si lo están, están defendidos. Estos constituirán las extensiones del sistema (*fase 3*). Es importante señalar que argumentación rebatible se supone que prevalecer ante las objeciones asegura que los argumentos se constituyan en buenas (razonables) razones para las conclusiones que sustentan. Esta fase puede hacerse en base a la satisfacción de condiciones previamente especificadas, denominadas '*semánticas*', que un conjunto de argumentos debe verificar, o mediante algún procedimiento de prueba denominado '*juegos argumentativos*'. En ambos casos diversas exigencias podrán pedirse a los argumentos para que estos califiquen como extensiones del sistema. Tales exigencias estarán regidas por criterios más tolerantes, usualmente bajo teorías crédulas o criterios estrictos bajo teorías escépticas (Dung, 1995).

Los argumentos, en un marco argumentativo pueden verse derrotados, sin embargo, esto no significa que contará finalmente como un argumento injustificado o perdedor porque puede verse defendido por otros y por ello restablecido. El restablecimiento es un principio clave en argumentación rebatible, este puede ser ilustrado con el siguiente ejemplo:

> Ejemplo 1.1
> *a*. Tweety vuela porque se sabe que es un ave.
> *b*. Tweety no vuela porque según la observación realizada por Paul, Tweety es un pingüino.
> *c*. La observación de que Tweety es un pingüino no es confiable porque Paul la hizo en condiciones inadecuadas para las técnicas que empleó.

A partir de la consideración de los argumentos *a*, *b* y *c* es posible creer que Tweety vuela puesto que la única objeción para creer en ello, el argumento *b*, ha sido desacreditada por el argumento *c*.

Además de la consideración del restablecimiento, es importante atender al estado de los subargumentos de un argumento, i.e. a los componentes de un argumento. Con esto se puede decir que un argumento justificado debe verificar una especie de principio de composicionalidad. Esencialmente consiste en que un argumento podrá considerarse como justificado cuando todos sus subargumentos estén también justificados. Por caso, en el ejemplo 1.1, considerando sólo los argumento *a* y *b,* se podría decir que el argumento *a* no verifica esta propiedad puesto que el sustento del argumento está derrotado por la conclusión del argumento b.

Una vez que ya se sabe cuál o cuáles son los argumentos justificados, se procede a seleccionar las conclusiones de tales argumentos, estas serán las creencias o extensiones del sistema (*fase 4*).

Diversos sistemas argumentativos específicos han sido propuestos en la literatura. Chesñevar et al (2000) y, Prakken y Vreeswijk (2002) ofrecen unos interesantes panoramas donde exponen y caracterizan los principales formalismos. A pesar de la gran diversidad de sistemas, uno de los que ha recibido especial atención y sendas aplicaciones y refinamientos es el propuesto por Simari y Loui en (1992). Tal sistema puede ser considerado como un sistema clásico o estándar en argumentación rebatible. Este sistema ha sido denominado en la literatura como *MTDR* en relación a las iniciales del título del artículo original.

MTDR fue propuesto como un sistema que integra el criterio de preferencia por especificidad de Poole (1985) y la justificación por niveles de Pollock (1987) más el desarrollo del sistema propuesto originalmente por Loui (1987). Posteriormente, Chesñevar y otros (1994) proponen una serie de refinamientos al sistema original. Luego, *MTDR* y sus refinamientos constituye la base argumentativa de lo que actualmente se conoce como *DeLP* (*Defeasible Logic Programming*)

DeLP es un enfoque de programación que combina los resultados de programación lógica y argumentación rebatible. *DeLP* es capaz de representar información tentativa y potencialmente contradictoria, gestionarla adecuadamente y obtener conclusiones a partir de ella. Tal como se verá más adelante, *DeLP* considera dos tipos de reglas (*program rules*), las rebatibles, con las que se representa la información tentativa y las estrictas, con la que se modela la información estricta que el sistema posee. A partir de tales reglas, *DeLP* construye pruebas para literales y cuenta con un procedimiento justificatorio (*warrant procedure*) que implementa un análisis dialéctico cuyo fin es seleccionar a aquellos argumentos que prevalecen frente a sus rivales. Los literales sustentados por aquellas pruebas rebatibles, o argumentos, que salen airosos del análisis dialéctico serán considerados las creencias del programa.

En *DeLP* puede suceder que un par de argumentos sustente literales contradictorios. Para hacer frente a esos casos, al igual que *MTDR*, *DeLP* cuenta con un mecanismo de comparación de argumentos denominado *especificidad generalizada* (basada en la propuesta en *MTDR*)*,* pero además, cuenta con un mecanismo de comparación basado en reglas. Esto último capacita a *DeLP* para atender a otros criterios de comparación, y por lo tanto, de definir un amplio abanico de derrotas.

Resumiendo entonces, *DeLP* es un formalismo capaz de modelar razonamiento default y dirimir el conflicto entre ellos en base a la aplicación del criterio de especificidad generalizada. Ahora bien, en el presente trabajo se mostrarán una serie de problemas que afectan al sistema y que requieren atención para su resolución. Específicamente esto se hará con la exposición de una serie de ejemplos que, en principio, debería modelar adecuadamente, pero que no lo hace. En particular los problemas pueden ser caracterizados a grandes rasgos de la siguiente manera:

- *DeLP sancionará a ciertos argumentos como justificados cuando no deberían estarlo.*
- *DeLP sancionará a ciertos argumentos como no justificados cuando deberían estarlo.*

El autor es de la opinión que la causa de tales problemas se debe a que *DeLP* no implementa adecuadamente la idea general de que un razonamiento default es aceptable sólo cuando no existe información más específica contraria disponible. Con vistas a neutralizar la aparición de casos problemáticos causados por dicha cuestión, en el presente trabajo se propondrá un criterio de preselección de argumentos apoyado en especificidad. Esta idea se encuentra basada en lo sugerido en (Bodanza, 2011; Bodanza y Alessio, 2010; 2014; Alessio, 2012; 2015). Como novedad se podrá destacar que se abordan una serie de ejemplos no considerados en tales trabajos que originan problemas semejantes a los discutidos por los autores. Además, aquí se afirmará que todos los casos problemáticos tienen un origen común y por lo tanto, la neutralización de los casos extraños debe ser obtenida por el mismo mecanismo de solución.

Los ejemplos que serán analizados, fueron ampliamente discutidos en la literatura (Loui y Stiefvater, 1992; Prakken y Vreeswijk, 2000; Horty, 2001; Prakken, 2002, Caminada, 2004; Antoniou, 2006; Horty, 2007; Dimopoulos, 2009; Alessio, 2012; Bodanza y Alessio, 2010; Bodanza, 2011; Bodanza y Alessio, 2014; 2015; Horty, 2012; Alessio, 2015). En tales trabajos se han realizado diversos intentos de neutralización de los problemas detectados a partir de los ejemplos considerados.

El artículo se organiza como sigue. En la sección 2 se realizará una presentación general del sistema de argumentación rebatible propuesto por García y Simari (2004). En la sección 3 se presentan una serie de ejemplos que suponen desafíos al sistema propuesto. En la sección 4 se realizan algunos comentarios al respecto y se discuten posibles causas del problema y una manera de neutralizarlos. En la sección 5 se ilustran los resultados basados en las modificaciones propuestas en la sección 4 y se demuestran algunas propiedades. Finalmente se concluye.

2. Un sistema argumentativo para razonamiento default

Simari y Loui (1992) propusieron un sistema basado en argumentos para razonamiento derrotable combinando el criterio de preferencias propuesto por Poole (1985) y la justificación por niveles definida por Pollock (1987). Luego ha recibido sucesivas modificaciones con vistas a evitar algunos inconvenientes que el sistema padecía (Chesñevar et al, 1994). En García (1997; 2000) se propone un formalismo basado en el sistema propuesto por Simari y Loui denominado *DeLP* (*Defeasible Logic Programing*) donde se combina a la programación lógica con la argumentación rebatible a fin de modelar información tentativa y potencialmente contradictoria. Tal formalismo ha sido expuesto sistemáticamente por García y Simari (2004), siguiendo tal artículo, será presentado a continuación.

En la introducción se señaló que los sistemas argumentativos pueden ser caracterizados en un proceso consistente de pasos o etapas: construcción de argumentos, comparación, evaluación y selección de las conclusiones justificadas. Atendiendo a ello, se presentará al sistema *DeLP* en cuatro fases o etapas.

En la primera se expondrá el lenguaje formal, y las estructuras esenciales en *DeLP,* argumento y sub-argumento. Luego, la segunda fase será caracterizada en base a las relaciones de contra-argumentación, especificidad, equi-especificidad y preferencia basada en reglas. La fase de evaluación o justificación será definida a partir de diversos componentes conceptuales tales como: *Línea de argumentación; argumentos de soporte, interferencia y concordantes; línea de argumentación aceptable; árbol dialectico; marcado de un árbol dialéctico*. Tales conceptos permitirán determinar cuándo un literal sostenido por un argumento cualquiera puede ser considerado como *justificado* (*fase 4*).

2.1 Construcción de argumentos y lenguaje formal

En *DeLP,* el lenguaje es definido a partir de tres conjuntos disyuntos: un conjunto de hechos, un conjunto de reglas estrictas y un conjunto de reglas rebatibles. Un literal '*L*' en *DeLP* será un átomo 'A' o su negación '∽A', donde '∽' representa la negación clásica. Los literales no tienen variables. Un hecho es un literal, i.e. un átomo o su negación. Una regla estricta es un par ordenado denotado por '*head*$\leftarrow$*body*', donde '*head*' es un literal y '*body*' es un conjunto finito no vació de literales. Una regla rebatible es un par ordenado '*head*$\prec$*body*' donde '*head*' es un literal y '*body*' es un conjunto finito no vacío de literales. Por la definición de literal, claramente, la negación fuerte puede ser empleada en *head y body*.

Las reglas estrictas representan información no rebatible como '*todos los pingüinos son aves*' mientras que las reglas rebatibles permiten representar información tentativa como '*por lo general las aves vuelan*'. '*Tweety es ave*' es un hecho y se representa con una regla estricta con '*body*' vacío, mientras que una regla rebatible (con '*body*' vacío) representará una presunción. En (Martínez, et al., 2012) se desarrolla un sistema de programación lógica que extiende a *DeLP* dotándolo de la capacidad de trabajar con presunciones denominado *PreDeLP*, pero esta extensión no será considerada en el presente artículo.

Atendiendo al lenguaje apuntado recientemente puede decirse que un programa lógico rebatible, es un conjunto posiblemente infinito de hechos, reglas estrictas y rebatibles. En un programa $\mathcal{P}$ se distinguirá el conjunto Π de hechos y reglas estrictas y el conjunto Δ de reglas rebatibles. De modo que el conocimiento de un agente *a* es representado por un par (Π, Δ) que puede llamarse programa $\mathcal{P}$ o estructura lógica rebatible. Π representa el conocimiento no rebatible de *a*. Δ, por su parte, representa la información tentativa y es un conjunto finito de reglas rebatibles.

Una *derivación rebatible* de un literal L a partir del programa $\mathcal{P}$, notado como $\mathcal{P} \mid\!\sim L$, consiste de una secuencia finita $L_1, L_2, \ldots, L_n \mid\!\sim L$ de literales, y para cada literal L_i en la secuencia se tiene que L_i es un hecho en Π, o existe una regla R_i en $\mathcal{P}$ (estricta o rebatible) con *head* igual a L_i y *body* igual a $B_1, B_2, \ldots, B_k$ y todos los literales de *body* son un elemento L_j de la secuencia aparecida antes de L_i ($j<i$). Por su parte, una *derivación estricta* de un literal L a partir del programa $\mathcal{P}$, notado como $\mathcal{P} \vdash L$, consiste en que L es un hecho o todas las reglas usadas para obtener L es una secuencia de reglas estrictas.

Atendiendo al lenguaje formal considerado y a las relaciones de derivación rebatible y estricta pueden ser introducidos los conceptos de argumento y sub-argumento. Un argumento es en *DeLP* una estructura constituida por un conjunto minimal y consistente de reglas rebatibles instanciadas que sustentan una conclusión determinada. Estos son notados como pares $\langle T, h \rangle$ donde T es un conjunto de reglas rebatibles y h la conclusión de la estructura de argumento. Si $\langle S,j \rangle$ es un argumento y se verifica que $S \subseteq T$, entonces $\langle S,j \rangle$ se considerará como sub-argumento de $\langle T,h \rangle$, a su vez, si $S \subset T$, entonces $\langle S,j \rangle$ es un sub-argumento propio de $\langle T,h \rangle$. Esto se encuentra precisado en la definición 2.1.

Definición 2.1: *Argumento y sub-argumento*

[García y Simari, 2004]

Sea h un literal y $\mathcal{P} = (\Pi, \Delta)$ un *delp*. Se dirá que $\langle T,h \rangle$ es una *estructura de argumento* si T es un conjunto de reglas rebatibles de Δ tal que:

i. Existe una derivación rebatible para h a partir de $\Pi \cup T$

ii. El conjunto $\Pi \cup T$ no es contradictorio

iii. T es minimal, no existe un subconjunto propio T' de T tal que T' satisfaga las condiciones *i. y ii.*

Además se dirá que $\langle S,j \rangle$ es un *sub-argumento* de $\langle T,h \rangle$, notado como $\langle S,j \rangle \sqsubseteq \langle T,h \rangle$ si y sólo si $\langle T,h \rangle$ una estructura de argumento para *h,* y $\langle S,j \rangle$ es una estructura de argumento para j tal que $S \subseteq T$. Si $S \subset T$ se dirá que S es un *sub-argumento propio* de T.

□

La condición *i. y ii.* afirma que un argumento es un conjunto consistente de reglas rebatibles que permite la derivación de un literal *h,* La condición *iii.* asegura que el conjunto T sea mínimo, i.e. que no haya información innecesaria para producir la derivación rebatible de *h.*

2.2 Comparación de argumentos: contra-argumentación y preferencia

Dos o más argumentos pueden sustentar conclusiones contradictorias, de modo que no pueden ser conjuntamente aceptados. La definición general que permite capturar casos de argumentos mutuamente excluyentes es la de contra-argumentación. Se dice que un argumento $\langle T_1,h_1 \rangle$ *contra-argumenta* a otro $\langle T_2,h_2 \rangle$ siempre y cuando exista un sub-argumento $\langle T,h \rangle$ de $\langle T_2,h_2 \rangle$ tal que $\langle T_1,h_1 \rangle$

y $\langle T,h \rangle$ están en desacuerdo. Se dice que un par de argumentos están en desacuerdo, en este caso $\langle T_1,h_1 \rangle$ y $\langle T,h \rangle$, cuando las conclusiones h y h_1 en unión con Π implican contradicción.

Definición 2.2: *Contraargumento*

[García y Simari, 2004]

Se dirá que $\langle T_1, h_1 \rangle$ *contra-argumenta* $\langle T_2, h_2 \rangle$ en el literal h si y sólo si existe un sub-argumento $\langle T, h \rangle$ de $\langle T_2, h_2 \rangle$ tal que h_1 y h están en desacuerdo. Si $\langle T_1, h_1 \rangle$ contra-argumenta a $\langle T_2, h_2 \rangle$ en el literal h, entonces h será llamado *punto de contra-argumentación* y el sub-argumento $\langle T, h \rangle$ será llamado *sub-argumento de desacuerdo*.

□

Un contraargumento $\langle T_1, h_1 \rangle$ de $\langle T_2, h_2 \rangle$ puede atacar directamente la conclusión h_2 o un punto intermedio h. Esto permite distinguir entre ataque directo e indirecto. Siendo un ataque directo cuando se niega la conclusión del otro. Indirecto, cuando se ataca o niega un paso de las premisas del otro argumento.

DeLP cuenta con dos maneras de comparar argumentos a fin de dirimir un desacuerdo. Uno basado en la noción de especificidad y el otro mediante prioridades entre reglas. La noción de especificidad es capturada por la definición de especificidad generalizada que será expuesta a continuación. Intuitivamente, esta definición favorece dos aspectos en los argumentos: prefiere aquel que contenga mayor información o que use menos reglas, i.e. es más directo. De modo que un argumento será considerado mejor que otro si es más preciso y conciso que aquel.

Definición 2.3. Especificidad generalizada

[García y Simari, 2004]

Sea $\mathcal{P} = (\Pi, \Delta)$ un *delp* y sea Π_G el conjunto de todas las reglas estrictas en Π (sin incluir hechos). Sea $\boldsymbol{\mathcal{F}}$ el conjunto de todos los literales para los que existe una derivación rebatible a partir de $\mathcal{P}$ ($\boldsymbol{\mathcal{F}}$ será considerado un conjunto de hechos). Sea $\langle T_1, h_1 \rangle$ y $\langle T_2, h_2 \rangle$ dos estructuras de argumento obtenidas a partir de $\mathcal{P}$. Se dirá que $\langle T_1, h_1\rangle$ es *estrictamente más específico* que $\langle T_2, h_2\rangle$, notado como $\langle T_1, h_1\rangle \succ \langle T_2, h_2\rangle$, si se verifican las siguientes condiciones:

Para todo H $\subseteq \boldsymbol{\mathcal{F}}$: Si $\Pi_G \cup H \cup T_1 \mathrel{|\!\!\sim} h_1$ y $\Pi_G \cup H \mathrel{|\!\!\not\sim} h_1$ entonces $\Pi_G \cup H \cup T_2 \mathrel{|\!\!\sim} h_2$, y

Existe H'$\subseteq \boldsymbol{\mathcal{F}}$ tal que $\Pi_G \cup H' \cup T_2 \mathrel{|\!\!\sim} h_2$ y $\Pi_G \cup H' \mathrel{|\!\!\not\sim} h_2$ y $\Pi_G \cup H' \cup T_1 \mathrel{|\!\!\not\sim} h_1$

□

No es posible tener una derivación rebatible de un literal a partir de un conjunto de reglas sin hechos. Por lo tanto, del conjunto $\Pi_G \cup T_1$ no sería posible obtener una derivación rebatible para h_1, sin embargo, $\Pi_G \cup H \cup T_1$ sí, porque H es un conjunto de literales. Por ello se dice que H *activa* a $\langle T_1, h_1\rangle$ o que H es un *conjunto de activación* para $\langle T_1, h_1\rangle$ en el sentido que H es lo que permite que se disparen las reglas en el argumento en cuestión. Por ello se suele decir que un argumento $\langle T_1, h_1\rangle$ será más específico que $\langle T_2, h_2\rangle$ cuando todo lo que activa a $\langle T_1, h_1\rangle$ activa a $\langle T_2, h_2\rangle$ pero algo de lo que activa a $\langle T_2, h_2\rangle$ no activa a $\langle T_1, h_1\rangle$.

En la definición de especificidad

Definición 2.4. Equi-especificidad

[García y Simari, 2004]

Dos argumentos $\langle T_1, h_1\rangle$y $\langle T_2, h_2\rangle$ son *equi-específicos*, denotado como $\langle T_1, h_1\rangle \equiv \langle T_2, h_2\rangle$, si y sólo si $T_1 = T_2$, y el literal h_1 tiene una derivación estricta a partir de $\Pi \cup \{h_2\}$ y h_2 tiene una derivación estricta a partir de $\Pi \cup \{h_1\}$.

□

Dada la estructura modular de *DeLP*, es posible introducir mecanismos de comparación de argumentos adicionales o que no requieran la noción de especificidad generalizada expuesta anteriormente. Una forma de equipar a *DeLP* con otros criterios de comparación puede hacerse mediante el empleo de prioridades entre reglas.

Definición 2.5. Preferencia basada en reglas

[García y Simari, 2004]

Sea $\mathcal{P} = (\Pi,\Delta)$ un delp y $>$ una relación de preferencia explícitamente definida entre las reglas rebatibles. Dadas dos estructuras de argumento $\langle T_1, h_1\rangle$y $\langle T_2, h_2\rangle$, $\langle T_1, h_1\rangle$ *será preferido* a $\langle T_2, h_2\rangle$ si *i.* Existe al menos una regla $r_a \in T_1$ y una regla $r_b \in T_2$ tal que $r_a > r_b$, y

ii. *no* existe un $r'_b \in T_2$y $r'_a \in T_1$ tal que $r'_b > r'_a$

□

Según García y Simari (2004) los criterios de comparación expuestos permiten obtener otro criterio más sofisticado combinándolos. Por ejemplo, considerando primero especificidad generalizada y si ningún argumento es preferido, usar otras prioridades.

Dada una estructura de argumento $\langle T_1, h_1\rangle$ y un contraargumento $\langle T_2, h_2\rangle$ para $\langle T_1, h_1\rangle$ es posible comparar ambos argumentos y decidir cuál prevalece. Si el contraargumento $\langle T_2, h_2\rangle$ es mejor que $\langle T_1, h_1\rangle$ con respecto al criterio de comparación empleado, entonces $\langle T_2, h_2\rangle$ será denominado un *derrotador propio* (*proper defeater*). Si ningún argumento es mejor, entonces se dirá que $\langle T_1, h_1\rangle$ es un *derrotador por bloqueo* (*blocking defetear*) de $\langle T_2, h_2\rangle$ y viceversa.

Con vistas a expresar la idea de derrota apelando al criterio de preferencia, pero que sea lo suficientemente general como para capturar, ya sea la especificidad generalizada o la prioridad entre reglas, un criterio abstracto de preferencia entre argumentos, notado como $\succ$, será empleado a continuación.

Definición 2.6. (derrotador propio)

[García y Simari, 2004]

Dadas dos estructuras de argumento $\langle T_1, h_1\rangle$ y $\langle T_2, h_2\rangle$, $\langle T_1, h_1\rangle$ es un *derrotador propio* de $\langle T_2, h_2\rangle$ en el literal h si y sólo si existe un subargumento $\langle T, h\rangle$ de $\langle T_2, h_2\rangle$ tal que $\langle T_1, h_1\rangle$ contraargumenta $\langle T_2, h_2\rangle$ en h, y $\langle T_1, h_1\rangle \succ \langle T, h\rangle$.

□

Definición 2.7. (derrotador por bloqueo)

[García y Simari, 2004]

Dadas dos estructuras de argumento $\langle T_1, h_1\rangle$y $\langle T_2, h_2\rangle$, $\langle T_1, h_1\rangle$ es un *derrotador por bloqueo* de $\langle T_2, h_2\rangle$ en el literal h si y sólo si existe un sub-argumento $\langle T, h\rangle$ de $\langle T_2, h_2\rangle$ tal que $\langle T_1, h_1\rangle$ contra-argumenta $\langle T_2, h_2\rangle$ en *h*, y $\langle T_1, h_1\rangle \nsucc \langle T, h\rangle$ y $\langle T, h\rangle \nsucc \langle T_1, h_1\rangle$.

□

Integrando ambas definiciones se obtiene la noción de derrota en *DeLP* de la siguiente manera:

Definición 2.8. Derrota

[García y Simari, 2004]

Dadas dos estructuras de argumento $\langle T_1, h_1\rangle$ y $\langle T_2, h_2\rangle$, $\langle T_1, h_1\rangle$ *derrota* a $\langle T_2, h_2\rangle$ si y sólo si

- $\langle T_1, h_1\rangle$ es un derrotador propio de $\langle T_2, h_2\rangle$; o
- $\langle T_1, h_1\rangle$ es un derrotador por bloqueo de $\langle T_2, h_2\rangle$.

□

2.3. Evaluación de argumentos: literal justificado

El agente modelado en *DeLP* creerá en aquellos literales que se encuentran justificados. Con vistas a determinar las condiciones bajo las cuales un literal puede considerarse como justificado se procede, a continuación, a presentar diversos conceptos.

Un primer paso para determinar la justificación de un argumento consiste en saber cuándo un argumento se considera como no derrotado y cuándo, lo contrario. Ahora bien, en orden a establecer si un argumento $\langle T_0, h_0\rangle$ es no derrotado, todos los derrotadores de $\langle T_0, h_0\rangle$ deben ser considerados. Supóngase que $\langle T_1, h_1\rangle$ es un derrotador de $\langle T_0, h_0\rangle$, pero $\langle T_1, h_1\rangle$ puede a la vez ser derrotado y así con otros. De esta manera, una secuencia de argumentos es creada, donde cada elemento de la secuencia derrota su predecesor, esta idea es expresada en el concepto de línea de argumentación.

Definición 2.9. Línea de argumentación

[García y Simari, 2004]

Sea $\mathcal{P}$ = ((Π, Δ) un *delp* y $\langle T_0, h_0\rangle$ una estructura de argumento obtenida a partir de $\mathcal{P}$. Una línea de argumentación para $\langle T_0, h_0\rangle$ es una secuencia de estructuras de argumento a partir de $\mathcal{P}$, notado como $\Lambda=[\langle T_0, h_0\rangle, \langle T_1, h_1\rangle, \langle T_2, h_2\rangle, \langle T_3, h_3\rangle,\ldots]$,donde cada elemento de la secuencia $\langle T_i, h_i\rangle$, i > 0, es un derrotador de su predecesor $\langle T_{i-1}, h_{i-1}\rangle$

□

Dado que una línea de argumentación puede resultar infinita se requiere la imposición de algunas restricciones. Para introducirlas será necesario proponer los conceptos de soporte e interferencia y concordancia.

Definición 2.10. Argumentos de soporte e interferencia

[García y Simari, 2004]

Sea $\Lambda=[\langle T_0, h_0\rangle, \langle T_1, h_1\rangle, \langle T_2, h_2\rangle, \langle T_3, h_3\rangle,\ldots]$ una línea de argumentación, se dirá que $\Lambda_S=[\langle T_0, h_0\rangle, \langle T_2, h_2\rangle, \langle T_4, h_4\rangle,\ldots]$ es el *conjunto de argumentos de soporte* y $\Lambda_I=[\langle T_1, h_1\rangle, \langle T_3, h_3\rangle, \langle T_5, h_5\rangle,\ldots]$ es el *conjunto de argumentos de interferencia.*

□

Dado un argumento $\langle T_0, h_0\rangle$, pueden existir varios derrotadores para $\langle T_0, h_0\rangle$. Cada uno de ellos generará una línea de argumentación diferente. En tal línea de argumentación, los derrotadores podrán tener más de un derrotador. Esto hace que se generen más líneas de argumentación comenzando en $\langle T_0, h_0\rangle$. Por ello, se necesita un proceso que considere todas las posibles líneas.

Definición 2.11. Argumentos concordantes

[García y Simari, 2004]

Sea $\mathcal{P}$ = ((Π, Δ) un *delp* y $\langle T_1, h_1\rangle$ y $\langle T_2, h_2\rangle$ estructuras de argumento obtenidas a partir de $\mathcal{P}$. Se dirá que $\langle T_1, h_1\rangle$ y $\langle T_2, h_2\rangle$ son *argumentos concordantes* si y sólo si el conjunto $\Pi\cup T_1\cup T_2$ no es contradictorio. Más generalmente un conjunto de estructuras de argumentos $\{\langle T_i, h_i\rangle\}\}_{i=1}^{n}$ es concordante si y sólo si $\Pi\cup\bigcup_{i=1}^{n}T_i$ no es contradictorio.

□

Definición 2.12. Línea de argumentación aceptable

[García y Simari, 2004]

Sea $\Lambda=[\langle T_0, h_0\rangle,\ldots,\langle T_i, h_i\rangle,\ldots,\langle T_n, h_n\rangle]$ una línea de argumentación, se dirá que Λ es *aceptable* si y sólo si:

- Λ es una secuencia finita.
- El conjunto Λ_S, de argumentos de soporte, es concordante y el conjunto Λ_I, de argumentos de interferencia, es concordante.
- Ningún argumento $\langle T_k, h_k\rangle$ en Λ es un sub-argumento de un argumento $\langle T_i, h_i\rangle$ aparecido previamente en Λ (i<k).
- Para cualquier *i*, tal que el argumento $\langle T_i, h_i\rangle$ es un derrotador por bloqueo para $\langle T_{i-1}, h_{i-1}\rangle$, si $\langle T_{i+1}, h_{i+1}\rangle$ existe, entonces $\langle T_{i+1}, h_{i+1}\rangle$ es un derrotador propio para $\langle T_i, h_i\rangle$.

□

La noción de línea de argumentación aceptable permite analizar la posibilidad de creer o no en un literal *h* a partir de un de.l.p. determinado. La idea de fondo es que si un argumento, el que se encuentra bajo análisis, tiene un derrotador, para poder considerarlo como una justificación para la afirmación que sustenta, debe contar con un defensor: *un derrotador del que lo derrota*.

Las condiciones dadas en la definición 2.12 evitan situaciones indeseables en el proceso de defensa: usar como defensa un argumento contradictorio con otras defensas, o usar un argumento como defensa de sí mismo, por ejemplo.

El concepto de línea de argumentación aceptable no alcanza para determinar si un literal está o no justificado porque el argumento que lo sustenta puede tener más de un derrotador, lo que origina líneas adicionales de argumentación, conformándose de ese modo en un árbol en torno al argumento bajo análisis. Este concepto será expresado más adelante.

Adicionalmente, vale la pena destacar unrasgo interesante en la definición 2.12: cambiar alguna de las propiedades de la definición permite obtener un comportamiento distinto en el formalismo.

Retornando al planteo de la sección, i.e. cuándo un literal se encuentra justificado en *DeLP,* se dirá que un *literal 'h' estará justificado* si existe una estructura de argumento $\langle T, h\rangle$ no derrotada. Con vistas a establecer si $\langle T, h\rangle$ es no derrotado, el conjunto de derrotadores para tal argumento debe ser considerado. Dado que para cualquier estructura de argumento que derrote a $\langle T, h\rangle$ puede a su vez ser derrotado y así con otros, más de una línea de argumentación puede aparecer. El conjunto de las líneas de argumentación para un argumento cualquiera podrá tener la estructura de árbol, con tal concepto se podrá tener un acercamiento a la idea de literal justificado.

En la definición 2.13 se expresa formalmente la idea empleando la noción de 'árbol dialéctico'.

Definición 2.13. Árbol Dialéctico

[García y Simari, 2004]

Sea $\mathcal{P}$ =(Π, Δ) un *delp* y $\langle T_0, h_0\rangle$ una estructura de argumento obtenida a partir de $\mathcal{P}$. Un *árbol dialéctico* para $\langle T_0, h_0\rangle$, notado $\boldsymbol{\mathcal{T}}\langle T_0, h_0\rangle$, es definido como sigue:

- La raíz del árbol es etiquetado con $\langle T_0, h_0\rangle$.
- Sea N un nodo no raíz del árbol etiquetado $\langle T_n, h_n\rangle$, y $\Lambda= [\langle T_0, h_0\rangle,\langle T_1, h_1\rangle,\langle T_2, h_2\rangle,\ldots, \langle T_n, h_n\rangle]$ la secuencia de etiqueta de los caminos de la raíz a N.
  Sean $\langle S_1, j_1\rangle,\langle S_2, j_2\rangle,\ldots, \langle S_k, j_k\rangle$ todos los derrotadores para $\langle T_n, h_n\rangle$.
  Para cada derrotador $\langle S_i, j_i\rangle$ ($1 \leq i \leq$ k), tal que, la línea de argumentación $\Lambda' = [\langle T_0, h_0\rangle,\langle T_1, h_1\rangle,\langle T_2, h_2\rangle,\ldots,\langle T_n, h_n\rangle,\langle S_i, j_i\rangle]$ sea aceptable, existe un nodo hijo N_i de N etiquetado con $\langle S_i, j_i\rangle$.
  Si no existe ningún derrotador para $\langle T_n, h_n\rangle$ o no existe un $\langle S_i, j_i\rangle$ tal que Λ' es aceptable entonces N es una hoja.

□

En un árbol dialéctico, cualquier nodo *N*, excepto el nodo raíz, representa un nodo derrotador de su padre y las hojas corresponden a argumentos no derrotados. Cada camino desde la raíz a una hoja corresponde a una línea de argumentación diferente.

En orden a decidir si la raíz de un árbol dialéctico es derrotada, un proceso de marcado es definido en *DeLP*. Los nodos serán marcados recursivamente como "D" (derrotado) o "U" (no derrotado) como sigue:

Definición 2.14. marcado de un árbol dialéctico

[García y Simari, 2004]

Sea $\boldsymbol{\mathcal{T}}\langle T, h\rangle$ un árbol dialéctico para $\langle T, h\rangle$, el correspondiente marcado del árbol dialéctico, notado como $\boldsymbol{\mathcal{T}}^*\langle T, h\rangle$, será obtenido marcando cada nodo en $\boldsymbol{\mathcal{T}}\langle T, h\rangle$ de la siguiente manera:

Todas las hojas en $\boldsymbol{\mathcal{T}}\langle T, h\rangle$ serán marcadas como "U" en $\boldsymbol{\mathcal{T}}^*\langle T, h\rangle$.

Sea $\langle S, j\rangle$ un nodo interno de $\boldsymbol{\mathcal{T}}\langle T, h\rangle$ entonces $\langle S, j\rangle$ será marcado como "U" en $\boldsymbol{\mathcal{T}}^*\langle T, h\rangle$ si y sólo si cualquier hijo de $\langle S, j\rangle$ es marcado como "D" en $\boldsymbol{\mathcal{T}}^*\langle T, h\rangle$. El nodo $\langle S, j\rangle$ será marcado como "D" en $\boldsymbol{\mathcal{T}}^*\langle T, h\rangle$ si y sólo si existe al menos un hijo marcado como "U" en $\boldsymbol{\mathcal{T}}^*\langle T, h\rangle$.

□

La idea intuitiva subyacente a la definición 2.14 consiste en identificar aquellos argumentos que no están derrotados con respecto a un conjunto vacío de argumentos y luego desde las hojas a la raíz, se va ascendiendo, identificando en el trayecto a aquellos argumentos que cuentan con defensas y a aquellos que no. Obviamente, si luego del marcado de un árbol, un argumento es etiquetado como "U", tal argumento es una extensión del sistema, o lo que es lo mismo, está justificado, o justifica la conclusión que sustenta. Esta última idea es la que se encuentra expresada en la definición 2.15.

Definición 2.15. Literal justificado

[García y Simari, 2004]

Sea $\langle T, h\rangle$ una estructura de argumento y $\boldsymbol{\mathcal{T}}^*\langle T, h\rangle$ su árbol dialéctico marcado asociado. El literal *h es justificado* si y sólo si la raíz de $\boldsymbol{\mathcal{T}}^*\langle T, h\rangle$ es marcada como "U". Cuando tal caso se verifica se dirá que *T* es una justificación para *h*.

□

El proceso de identificación de una justificación para un literal se hace por tanto mediante un procedimiento bottom-up de marcado. Un argumento será una justificación para la conclusión que sustenta si todos sus derrotadores son argumentos derrotados, en términos de un árbol dialéctico, un argumento estará justificado si el árbol dialéctico para un argumento determinado tiene marcado el nodo raíz como 'U'.

3. Casos problemáticos

DeLP es un formalismo capaz de modelar información tentativa y potencialmente contradictoria mediante la construcción, comparación y evaluación de argumentos a favor y en contra de ciertas afirmaciones o literales. A continuación, se presentarán una serie de ejemplos o situaciones en las que *DeLP* parece tener problemas. Los ejemplos han sido discutidos en diversos trabajos (Touretzky *et al* (1991); Loui y Stiefvater, 1992; Prakken y Vreeswijk, 2000; Horty, 2001; 2012; Prakken, 2002, Caminada, 2004; Antoniou, 2006; Dimopoulos, 2009; Alessio, 2012; 2015; Bodanza, 2011; Bodanza y Alessio, 2010; 2014; 2016).

Caso 1: Al, la gallina salvaje

A continuación se propondrá el ejemplo de *Al, la gallina salvaje*, originalmente propuesto por Touretzky et al (1991) y discutido en Loui y Stiefvater (1992) y Horty (2001). Un ejemplo equivalente puede encontrarse en Prakken y Vreeswijk (2000) y en Prakken (2000), el ejemplo del *pingüino genéticamente modificado*.

Supóngase que se dispone de la siguiente información. Existe una gallina salvaje muy peculiar llamada Al. Se sabe que en general las gallinas salvajes vuelan. Obviamente por una regularidad entre las gallinas es común que estas no vuelen. También se sabe que en general las aves son animales voladores. A su vez se sabe que todas las gallinas salvajes son gallinas y que las gallinas son aves. Supóngase que $WC(z)$, $C(x)$, $B(x)$, y $F(x)$ representan respectivamente a las proposiciones: *x es una gallina salvaje, x es una gallina, x es un ave, y x es un ave voladora.* Tal información puede expresarse en el siguiente programa:

$$\Delta = \{F(x) \prec B(x),\ \backsim F(x) \prec C(x),\ F(x) \prec WC(x)\}$$
$$\Pi_G = \{C(x) \longleftarrow WC(x),\ B(x) \longleftarrow C(x)\}$$
$$\Pi_F = \{\mathrm{WC}(a)\}$$

A partir de tal programa, los siguientes tres argumentos pueden construirse:

$\langle A_1, h_1 \rangle$	$\langle \{F(a) \prec B(a)\}, F(a) \rangle$
$\langle A_2, h_2 \rangle$	$\langle \{\backsim F(a) \prec C(a)\}, \backsim F(a) \rangle$
$\langle A_3, h_3 \rangle$	$\langle \{F(a) \prec WC(a)\}, F(a) \rangle$

Dadas las relaciones de especificidad, es claro que $\langle A_2, h_2 \rangle$ es estrictamente más específico que $\langle A_1, h_1 \rangle$ y $\langle A_3, h_3 \rangle$ lo es con respecto a $\langle A_2, h_2 \rangle$.

$\langle A_2, h_2 \rangle$ es estrictamente más específico que $\langle A_1, h_1 \rangle$ porque todo lo que activa a $\langle A_2, h_2 \rangle$ también activa a $\langle A_1, h_1 \rangle$, a saber, WC(a) y C(a) es todo lo que activa a $\langle A_2, h_2 \rangle$, que también activa a $\langle A_1, h_1 \rangle$, más en particular:

- WC(a) es el hecho disponible en Π_F que junto a Π_G permite que la regla $\neg F(a) \prec C(a)$ sea disparada ya que $C(x) \longleftarrow WC(x) \in \Pi_G$. A su vez, es WC($a$) activador de $\langle A_1, h_1 \rangle$ porque WC(a) dispara la regla $C(x) \longleftarrow WC(x)$ que a su vez activa la regla $B(x) \longleftarrow C(x)$ que permite obtener la consecuencia B(a) que dispara la regla $F(a) \prec B(a)$.
- C(a) es un hecho derivado de $C(x) \longleftarrow WC(x)$ y $WC(a)$. C(a) dispara la regla $\neg F(a) \prec C(a)$, lo que permite que el argumento $\langle A_2, h_2 \rangle$ concluya. A su vez, $C(a)$ junto a la regla $B(x) \longleftarrow C(x)$ activa al argumento $\langle A_1, h_1 \rangle$.

Nótese que no habría diferencia en este ejemplo si como hechos se supusiese que Al es un ave y es una gallina salvaje, o que es un ave y una gallina, o que es una gallina salvaje y una gallina. Pero si Π_G fuese vacío, el comportamiento frente al ejemplo, sería diferente porque en ese caso no se relacionarían por especificidad.

Además de que todo lo que activa a $\langle A_2, h_2 \rangle$ debe activar a $\langle A_1, h_1 \rangle$, para que $\langle A_2, h_2 \rangle$ sea considerado estrictamente más específico que $\langle A_1, h_1 \rangle$ se debe verificar que al menos, algo de lo que activa a $\langle A_1,h_1\rangle$ no activa a $\langle A_2, h_2\rangle$. En este caso $B(a)$ es un activador de $\langle A_1,h_1\rangle$ pero no lo es de $\langle A_2,h_2\rangle$.

Teniendo en cuenta que $\langle A_2, h_2\rangle$ es un contraargumento más específico que $\langle A_1, h_1\rangle$, $\langle A_2, h_2\rangle$ derrota a $\langle A_1,h_1\rangle$ según definición 2.6 y 2.8.

Por otro lado, $\langle A_3, h_3\rangle$ derrota a $\langle A_2, h_2\rangle$ puesto que $\langle A_3, h_3\rangle$ es un contraargumento más específico que $\langle A_2, h_2\rangle$ según definición 2.6 y 2.8.

$\langle A_3, h_3\rangle$ es estrictamente más específico que $\langle A_2, h_2\rangle$ porque todo lo que activa a $\langle A_3, h_3\rangle$ también activa a $\langle A_2, h_2\rangle$, a saber, WC(a). Por otro lado, algo de lo que activa a $\langle A_2, h_2\rangle$, C(a), no activa a $\langle A_3, h_3\rangle$.

Más en particular, WC(a) es un hecho disponible en Π_F que permite que la regla $F(a)\prec CW(a)$ sea disparada activando de ese modo al argumento $\langle A_3, h_3\rangle$. A su vez, es WC(a) un activador de $\langle A_2, h_2\rangle$ porque WC(a) dispara la regla $C(x)\longleftarrow WC(x)$ que a su vez activa al argumento $\langle A_2, h_2\rangle$. Por su parte, C(a) activa a $\langle A_2, h_2\rangle$ pero tal hecho es incapaz de activar al argumento $\langle A_3, h_3\rangle$.

Atendiendo a las relaciones de derrota, en el proceso de justificación, los argumentos $\langle A_1,h_1\rangle$ y $\langle A_3,h_3\rangle$ serán marcados como "U", es decir, son argumentos finalmente no derrotados y los literales h_1 *y* h_3 deben considerarse como justificados. Los árboles marcados asociados a cada argumento se representan en la figura 1.

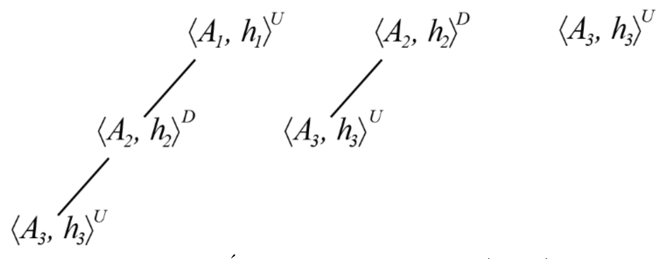

Figura 1: *Árbol marcado asociado a* $\boldsymbol{T}^*\langle T_1, h_1\rangle$, $\boldsymbol{T}^*\langle T_2, h_2\rangle$ *y* $\boldsymbol{T}^*\langle T_3, h_3\rangle$ *respectivamente*

Ahora bien, intuitivamente pareciera que las razones por las que alguien estaría dispuesto a aceptar que Al vuela no se deben al hecho de que Al sea un ave, sino al hecho de que Al es una gallina excepcional, una gallina salvaje. Esto lleva a afirmar que *DeLP* considera como justificado a un argumento que, aunque sustenta una conclusión correcta, se encuentra basado en razones incorrectas, a saber $\langle A_1,h_1\rangle$.

El caso 1 podría parecer no demasiado problemático, tal como lo señala Horty (2001), ya que en fin de cuentas el argumento sustenta una conclusión correcta. Sin embargo, el siguiente ejemplo permite señalar la posibilidad de considerar como justificados a argumentos que sustentan conclusiones que son dada la información disponible, simplemente, incorrectas.

Caso 2: Empleados millonarios

Suponga que posee la siguiente información. Usted sabe que *Beth es una nueva empleada de Microsoft en el departamento X. Sabe además que tales empleadas por lo general poseen al menos medio millón por una peculiar participación en las ganancias de la empresa. También sabe que hay un grupo de empleados de Microsoft que usualmente no son millonarios, los nuevos empleados. Pero los empleados estándar de la empresa tienden a ser millonarios.* La información al respecto puede ser modelada en *DeLP* de la siguiente manera. Si *NMEX(b), NME(b), ME(b), 1M(b), <1/2M(b), >1/2M(b)* representa respectivamente las siguientes proposiciones: *Beth es una nueva empleada de Microsoft en el departamento X, Beth es una nueva empleada de Microsoft, Beth es una empleada de Microsoft, Beth tiene un millón dólares, Beth tiene menos de medio millón de dólares y Beth tiene más de medio millón de dólares,* es posible construir en *DeLP* el siguiente programa:

$$\Delta = \{1M(x) \prec ME(x), <1/2\mathrm{M}(x) \prec NME(x), >1/2M(x) \prec NMEX(x)\}$$
$$\Pi_G = \{ME(x) \longleftarrow NME(x), NME(x) \longleftarrow NMEX(x), >1/2M(x) \longleftarrow 1M(x),$$
$$\sim <1/2M(x) \longleftarrow 1M(x), \sim 1M(x) \longleftarrow <1/2M(x), \sim >1/2M(x) \longleftarrow <1/2M(x),$$
$$\sim <1/2M(x) \longleftarrow >1/2M(x)\}$$
$$\Pi_F = \{\mathrm{NMEX}(a)\}$$

Debe notarse que en el programa se han introducido relaciones de conflicto entre ciertos predicados, las razones son claras, *poseer un millón de dólares* implica que *no es cierto que se posea menos que medio millón*, y viceversa. Por otro lado, *poseer al menos medio millón* implica que *no es cierto que se posea menos de medio millón* y viceversa. También se han explicitado las relaciones de implicación entre ciertos predicados, en particular en relación a lo siguiente:

poseer un millón implica *poseer al menos medio millón*, *ser empleado nuevo en el departamento X de Microsoft* implica *ser empleado nuevo en Microsoft*, y *ser empleado nuevo en Microsoft* implica *ser empleado de Microsoft*. Tres argumentos interesantes, construidos a partir de la información suministrada, deben considerarse:

$\langle A_1,h_1\rangle$	$\langle\{1M(b) \prec ME(b)\}, 1M(b)\rangle$
$\langle A_2,h_2\rangle$	$\langle\{<1/2M(b) \prec NME(b)\}, <1/2M(b)\rangle$
$\langle A_3,h_3\rangle$	$\langle\{>1/2M(b) \prec NMEX(b)\}, >1/2M(b)\rangle$

Es fácil comprobar que en *DeLP* $\langle A_3, h_3\rangle$ es un derrotador de $\langle A_2, h_2\rangle$ y $\langle A_2, h_2\rangle$ es un derrotador de $\langle A_1, h_1\rangle$.

$\langle A_2, h_2\rangle$ derrota a $\langle A_1, h_1\rangle$ porque en primer lugar, $\langle A_2, h_2\rangle$ contra-argumenta a $\langle A_1, h_1\rangle$ debido a que $\sim <1/2M(x) \longleftarrow 1M(x)$ y $\sim 1M(x) \longleftarrow <1/2M(x)$, y todo lo que activa a $\langle A_2, h_2\rangle$ también activa a $\langle A_1, h_1\rangle$ pero algo de lo que activa a $\langle A_1, h_1\rangle$ no activa a $\langle A_2, h_2\rangle$. Puntualmente, a partir del hecho de que Beth sea una empleada nueva de Microsoft en el departamento X ($NMEX(a)$) se sabe que Beth es una empleada de Microsoft ($ME(a)$) y que es una empleada Nueva de Microsoft ($NME(a)$). Los hechos que activan al argumento $\langle A_2, h_2\rangle$ son $NMEX(a)$ y $NME(a)$, tales hechos también activan a $\langle A_1, h_1\rangle$, ahora bien, $ME(a)$ es un hecho que permite activar a $\langle A_1, h_1\rangle$ pero no activa a $\langle A_2, h_2\rangle$

Por otro lado, $\langle A_3, h_3\rangle$ derrota a $\langle A_2, h_2\rangle$ porque en primer lugar, $\langle A_3, h_3\rangle$ contra-argumenta a $\langle A_2, h_2\rangle$ y $\langle A_3, h_3\rangle$ es estrictamente más específico que $\langle A_2, h_2\rangle$ porque $\{NMEX(a)\}$ es el conjunto de hechos que activa a $\langle A_3, h_3\rangle$ y que activa a $\langle A_2, h_2\rangle$. Pero, como ya se dijo, $NME(a)$ es un activador de $\langle A_2, h_2\rangle$ que no es activador de $\langle A_3, h_3\rangle$. Como puede evidenciarse en Π_G, $\langle A_3, h_3\rangle$ contra-argumenta a $\langle A_2, h_2\rangle$ porque $\sim >1/2M(x) \longleftarrow <1/2M(x)$, $\sim <1/2M(x) \longleftarrow >1/2M(x)$.

Tanto $\langle A_3, h_3\rangle$ como $\langle A_1, h_1\rangle$, calificarán luego del procedimiento de justificación, como podrá constatarse en la figura 2, como argumentos justificados alcanzando un resultado no intuitivo: *Beth es millonaria*. Ahora bien, atendiendo a la información total disponible, no hay razones para creer que lo sea. La única razón que podría llevar a la aceptación de que es millonaria se debería a que fuese una empleada de Microsoft estándar, pero no lo es.

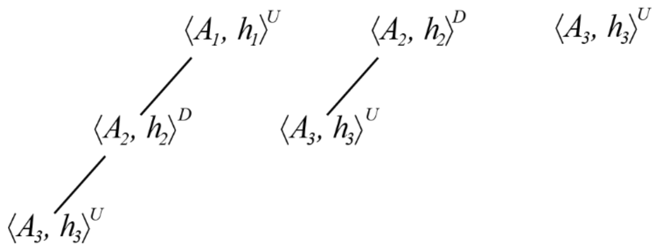

Figura 2: *Árbol marcado asociado a* $\mathcal{T}^*\langle T_1, h_1\rangle$, $\mathcal{T}^*\langle T_2, h_2\rangle$ *y* $\mathcal{T}^*\langle T_3, h_3\rangle$ *respectivamente del caso 2*

Caso 3: Ana y el escepticismo sobre si es o no adinerada

Considere la siguiente información sobre Ana*: Ana es defensora pública y renta un pequeño departamento en un lugar llamado Brentwood. En general las personas que allí viven son adineradas, aunque una pequeña cantidad de residentes de Brentwood, los que rentan, no lo son. Por otro lado, los abogados son en general adinerados con excepción de una subclase, los defensores públicos.* Esta información puede ser modelada en *DeLP* de la siguiente manera. *AD(a), AB(a), VB(a), RB(a), DP(a)* representan respectivamente las siguientes proposiciones: *Ana es adinerada, Ana es abogada, Ana vive en Brentwood, Ana renta en Brentwood y Ana es defensora pública.* Esta información puede ser expresada en el siguiente programa:

$\Delta = \{\ AD(x) \prec AB(x),\ \ {\sim}AD(x) \prec DP(x),\ \ {\sim}AD(x) \prec RB(x),\ AD(x) \prec VB(x)\}$

$\Pi_G = \{\ VB(x) \longleftarrow RB(x), AB(x) \longleftarrow DP(x)\}$

$\Pi_F = \{DP(a), RB(a)\}$

A partir del anterior programa se pueden construir los siguientes argumentos:

$\langle A_1,h_1\rangle$	$\langle\{\ AD(a) \prec AB(a)\}, AD(a)\rangle$
$\langle A_2,h_2\rangle$	$\langle\{\ AD(a) \prec VB(a)\}, AD(a)\rangle$
$\langle A_3,h_3\rangle$	$\langle\{{\sim}AD(a) \prec DP(a)\}, {\sim}\mathrm{AD}(a)\rangle$
$\langle A_4,h_4\rangle$	$\langle\{{\sim}AD(a) \prec RB(a)\}, {\sim}\mathrm{AD}(a)\rangle$

Es fácil comprobar que en *DeLP,* $\langle A_3, h_3\rangle$ es un derrotador propio de $\langle A_1, h_1\rangle$, y que $\langle A_3, h_3\rangle$ es un derrotador por bloqueo para $\langle A_2, h_2\rangle$ y viceversa. A su vez, $\langle A_4, h_4\rangle$ es un derrotador propio de $\langle A_2, h_2\rangle$, y $\langle A_4, h_4\rangle$ y $\langle A_1, h_1\rangle$ se derrotan recíprocamente.

$\langle A_3, h_3\rangle$ es un derrotador propio de $\langle A_1, h_1\rangle$ porque $\langle A_3, h_3\rangle$ contra-argumenta a $\langle A_1, h_1\rangle$ y todo lo que activa a $\langle A_3, h_3\rangle$, *DP(a),* activa también a $\langle A_1, h_1\rangle$, pero algo que activa a $\langle A_1, h_1\rangle$, *AB(a),* no activa a $\langle A_3, h_3\rangle$.

$\langle A_3, h_3\rangle$ es un derrotador por bloqueo de $\langle A_2, h_2\rangle$ ya que son incomparables por especificidad y las conclusiones de ambos argumentos son contradictorias.

$\langle A_4, h_4\rangle$ es un derrotador propio de $\langle A_2, h_2\rangle$ porque $\langle A_4, h_4\rangle$ contra-argumenta a $\langle A_2, h_2\rangle$ y todo lo que activa a $\langle A_4, h_4\rangle$, i.e. *RB(a),* activa también a $\langle A_2, h_2\rangle$, pero algo que activa a $\langle A_2, h_2\rangle$, a saber *VB(a),* no activa a $\langle A_4, h_4\rangle$.

$\langle A_4, h_4\rangle$ es un derrotador por bloqueo de $\langle A_1, h_1\rangle$ ya que son incomparables por especificidad y las conclusiones de ambos argumentos son contradictorias.

Atendiendo al procedimiento de justificación en *DeLP*, ninguno de los cuatro argumentos cuenta como justificados, y por tanto ningún literal, de los sustentados por los cuatro argumentos cuenta entre las creencias del sistema. Sin embargo, es claro que los argumentos $\langle A_3,h_3\rangle$ y $\langle A_4,h_4\rangle$ deberían considerarse como justificados, es decir, intuitivamente no habría dudas de que Ana no es adinerada, puesto que se encuentra sustentada tal conclusión en base a los argumentos de información más específica, Ana renta en Brentwood y es defensora pública. Sin embargo, para *DeLP,* esto no es tan obvio. En la Figura 3 se puede observar los diversos árboles marcados asociados a cada argumento considerado.

En el árbol correspondiente al argumento $\langle A_1, h_1\rangle$ puede verse que este cuenta con dos nodos hijos que lo derrotan, $\langle A_3, h_3\rangle$ y $\langle A_4, h_4\rangle$, $\langle A_3, h_3\rangle$ es un derrotador propio de $\langle A_1, h_1\rangle$ y $\langle A_4, h_4\rangle$ es un derrotador por bloqueo. La línea que sigue a $\langle A_3, h_3\rangle$ continúa con una derrota por bloqueo de $\langle A_2, h_2\rangle$. Nótese que $\langle A_3, h_3\rangle$ tiene un solo hijo porque según la definición de árbol aceptable, $\langle A_1, h_1\rangle$ no podría usarse para defenderse a sí mismo. $\langle A_2, h_2\rangle$ por su parte cuenta con un derrotador propio $\langle A_4, h_4\rangle$. En $\langle A_4, h_4\rangle$ finaliza esa línea porque el único argumento del ejemplo que derrota a $\langle A_4, h_4\rangle$ es $\langle A_1, h_1\rangle$, pero no puede usarse como defensa de sí mismo. Lo mismo sucede en la línea iniciada con $\langle A_4, h_4\rangle$. El etiquetamiento del árbol iniciando por las hojas, lleva a que $\langle A_1, h_1\rangle$ sea marcado como 'D' y el literal '*Ana es adinerada*' no esté justificado en $\langle A_1, h_1\rangle$.

Por su parte el árbol para $\langle A_2, h_2\rangle$, que sustenta el literal '*Ana es adinerada*' permite ilustrar el hecho de que cuenta con dos nodos hijos, sus derrotadores $\langle A_4, h_4\rangle$ y $\langle A_3, h_3\rangle$. $\langle A_4, h_4\rangle$ es un derrotador propio de $\langle A_2, h_2\rangle$ y $\langle A_3, h_3\rangle$ es un derrotador por bloqueo de $\langle A_2, h_2\rangle$. Por su parte, $\langle A_4, h_4\rangle$ tiene un nodo hijo que lo derrota, $\langle A_1, h_1\rangle$, a su vez derrotado por $\langle A_3, h_3\rangle$. Obviamente que no se puede derrotar a $\langle A_3, h_3\rangle$ en ninguna de las dos líneas debido a que solo sería posible mediante la derrota por bloqueo de $\langle A_2, h_2\rangle$ pero $\langle A_2, h_2\rangle$ se estaría defendiendo a sí mismo, y la noción de árbol aceptable restringe ese caso. De modo que el marcado del árbol, iniciando por las hojas, lleva a marcar a $\langle A_2, h_2\rangle$ como 'D' y el literal '*Ana es adinerada*' no se puede determinar justificado en A_2.

El árbol asociado a $\langle A_3, h_3\rangle$ cuenta de una única línea de argumentación porque sólo $\langle A_2, h_2\rangle$ derrota a $\langle A_3, h_3\rangle$. $\langle A_2, h_2\rangle$ es derrotado a su vez por $\langle A_4, h_4\rangle$ que es bloqueado por $\langle A_1, h_1\rangle$ de modo que el marcado de la raíz del árbol asigna a $\langle A_3, h_3\rangle$ la etiqueta 'D'. Por lo tanto, el literal '*Ana no es adinerada*' no se encuentra justificado en A_3.

El árbol asociado a $\langle A_4, h_4\rangle$ está conformado por una línea de argumentación. $\langle A_4, h_4\rangle$ es bloqueado por $\langle A_1, h_1\rangle$, $\langle A_1, h_1\rangle$ es derrotado propiamente por $\langle A_3, h_3\rangle$ quien es bloqueado por $\langle A_2, h_2\rangle$. Obviamente que $\langle A_2, h_2\rangle$ no puede ser derrotado debido a que de lo contrario no se trataría de una línea de argumentación aceptable. Por ello, el marcado de $\langle A_4, h_4\rangle$ es 'D' y la justificación de *"Ana no es adinerada"* no puede hacerse.

Como puede notarse en la figura 3, ningún argumento puede considerarse como una justificación para la conclusión que sustenta.

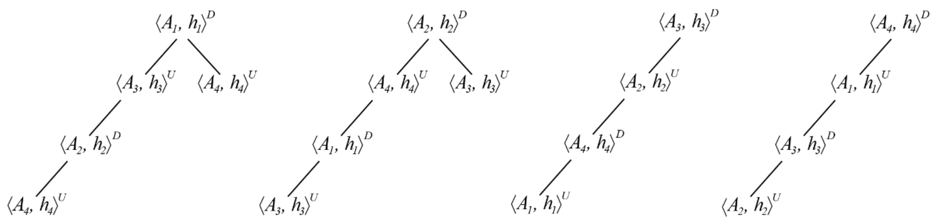

Figura 3: *Árbol marcado asociado a* $\mathcal{T}^*\langle T_1, h_1\rangle$, $\mathcal{T}^*\langle T_2, h_2\rangle$, $\mathcal{T}^*\langle T_3, h_3\rangle$ y $\mathcal{T}^*\langle T_4, h_4\rangle$ *respectivamente del caso 3*

Caso 4. La isla florida y desierta

Imagínese que un biólogo se encuentra estudiando la distribución de las aves en una remota zona de islas (*las islas remotas*) donde ha identificado una curiosa especie de canarios llamada *canarios del cabo*. Por un particular comportamiento de la especie, los nidos se encuentran distribuidos en su mayoría, aunque no únicamente en la isla Florida, un sector muy pequeño de la zona de islas remotas. Una subespecie de estos canarios, llamados *canarios del cabo rojo,* por su parte, tiene, por lo general, sus nidos distribuidos en toda la zona de islas. Ahora bien, considerando a un individuo particular de tal especie, llamado Frank, que se sabe, es un canario del cabo rojo ¿Qué podrá concluir el biólogo con respecto de la ubicación del nido de Frank? ¿Qué conclusiones podrían obtenerse en *DeLP*? Para responder los interrogantes, primeramente será preciso construir los argumentos. Atendiendo a la información considerada estos son los siguientes.

A. *El nido de Frank se encuentra en cualquier parte de la zona de las islas remotas porque Frank es un canario del cabo rojo*
B. *El nido de Frank se encuentra en la isla Florida porque Frank es un canario del cabo.*

Si *NIF(f), NCI (f), CC(f), CCR(f)* representa respectivamente las siguientes proposiciones: *Frank tiene su nido en la isla Florida, Frank tiene su nido en el complejo de islas, Frank es un canario del cabo, Frank es un canario del cabo rojo.* Tal información puede ser expresada en el siguiente programa

$\Delta = \{NIF(x) \prec CC(x), NCI(x) \prec CCR(x)\}$
$\Pi_G = \{NCI(x) \leftarrow NIF(x), CC(x) \leftarrow CCR(x)\}$
$\Pi_F = \{CCR(a)\}$

La anterior información permite construir los siguientes argumentos:

$\langle A_1,h_1 \rangle \qquad \langle \{NIF(f) \prec CC(f)\}, NIF(f) \rangle$
$\langle A_2,h_2 \rangle \qquad \langle \{NCI(f) \prec CCR(f)\}, NCI(f) \rangle$

Es fácil comprobar que en *DeLP* $\langle A_2, h_2 \rangle$ y $\langle A_1, h_1 \rangle$ son consistentes por lo que ambos argumentos están justificados. Sin embargo, no parece que el biólogo tenga equivalentes razones para buscar únicamente en la isla florida o en el complejo de islas. Intuitivamente, el argumento $\langle A_2, h_2 \rangle$ es más razonable que el argumento $\langle A_1,h_1 \rangle$. Además de que es más específico, la conclusión es menos *fuerte*. Sin embargo, para *DeLP*, ambos argumentos parecen buenas razones y por lo tanto, basado en ambos argumentos, para el biólogo será lo mismo buscar en la isla florida que en el complejo de islas.

Caso 5. La paradoja de la derrota bloqueante

Considere el siguiente ejemplo, aunque abstracto, y reflexione sobre el significado del mismo. Antes, se presentará la información en un programa, luego, los argumentos construibles a partir de ella.

$\Delta = \{G(x) \prec F(x), J(x) \prec F(x)\}$
$\Pi_G = \{\sim G(x) \leftarrow J(x)\}$
$\Pi_F = \{\mathrm{F}(j)\}$

A partir del anterior programa, los siguientes argumentos se pueden construir

$\langle A_1,h_1 \rangle \qquad \langle \{G(f) \prec F(f)\}, G(f) \rangle$
$\langle A_2,h_2 \rangle \qquad \langle \{\sim G(f) \leftarrow J(f), J(f) \prec F(f)\}, \sim G(f) \rangle$

Observación 1: Dada la definición de argumento, la regla $\backsim G(f) \longleftarrow J(f)$ no debe estar en el argumento $\langle A_2,h_2\rangle$ pero se la ha colocado allí con vistas a explicitar la relación de contra-argumentación entre ambos argumentos.

Si se atiende al concepto de derrota propuesto en *DeLP* ambos argumentos se bloquean y por lo tanto ninguno está justificado. No es el caso de que $\langle A_1,h_1\rangle$ sea más específico que $\langle A_2,h_2\rangle$ y viceversa. Podría pensarse a simple vista que $\langle A_1,h_1\rangle$ es más específico porque todo lo que activa a $\langle A_1,h_1\rangle$ activa a $\langle A_2,h_2\rangle$, pero lo que activa a $\langle A_2,h_2\rangle$ que no activa a $\langle A_1,h_1\rangle$, i.e. J(*f*), es un activador trivial de $\langle A_2,h_2\rangle$.

Retomando la idea del ejemplo, es razonable que ninguno de los argumentos sea aceptado, debido a que la aceptación de F(*f*), dadas las reglas en el sistema, lleva a aceptar una contradicción. Ahora bien, y a pesar de que la derrota por bloqueo en este ejemplo parece inofensiva, la perspectiva puede cambiar si se atiende a la siguiente situación.

Supóngase que luego de un aprendizaje, el agente modelado por *DeLP* es capaz de construir el siguiente argumento:

$\langle A_3,h_3\rangle$ $\quad\quad$ $\langle\{G(f) \prec H(f)\}, G(f)\rangle$

La base de conocimientos se actualiza de la siguiente manera:

$\Delta = \{G(x)\prec F(x),\ J(x)\prec F(x),\ G(x) \prec H(x)\}$

$\Pi_G = \{\backsim G(x)\longleftarrow J(x)\}$

$\Pi_F = \{F(j), H(j)\}$

Atendiendo a la información nueva, $\langle A_3,h_3\rangle$ es incomparable por especificidad con $\langle A_2,h_2$. $\langle A_3,h_3\rangle$ es un argumento que será considerado finalmente como derrotado puesto que hay una derrota por bloqueo con $\langle A_2,h_2\rangle$. El árbol dialéctico asociado a cada argumento (Figura 4) permitirá visualizar mejor lo dicho hasta aquí.

$$\begin{tikzcd}[arrows=-]
& \langle A_1, h_1\rangle^D \arrow[dl] & & \langle A_2, h_2\rangle^D \arrow[dl] \arrow[dr] & & \langle A_3, h_3\rangle^D \arrow[dr] & \\
\langle A_2, h_2\rangle^U & & \langle A_3, h_3\rangle^U & & \langle A_1, h_1\rangle^U & & \langle A_2, h_2\rangle^U
\end{tikzcd}$$

Figura 4: *Árbol marcado asociado a* $\boldsymbol{\mathcal{T}}^*\langle T_1, h_1\rangle$, $\boldsymbol{\mathcal{T}}^*\langle T_2, h_2\rangle$ *y* $\boldsymbol{\mathcal{T}}^*\langle T_3, h_3\rangle$ *respectivamente del caso 5.*

El árbol asociado a $\langle T_1, h_1\rangle$ cuenta con un único derrotador debido a que $\langle T_2, h_2\rangle$ es un bloqueador de $\langle T_1, h_1\rangle$, y la cuarta condición de la definición 2.12 establece la imposibilidad de bloquear a un bloqueador, por lo cual no es posible jugar como defensa de $\langle T_1, h_1\rangle$ a $\langle T_3, h_3\rangle$. Por tal motivo, $\langle T_1, h_1\rangle$ debe ser rotulado como 'D' y la conclusión sustentada por el argumento no puede considerarse justificada en tal argumento. Por su parte, el argumento $\langle T_2, h_2\rangle$ debe etiquetarse como 'D' porque este argumento cuenta con derrotadores etiquetados con 'U'. Vale recordar que la noción de línea de argumentación aceptable prohíbe la autodefensa, razón por la cual, $\langle T_2, h_2\rangle$ no puede jugarse luego de $\langle T_1, h_1\rangle$ y $\langle T_3, h_3\rangle$. El árbol asociado a $\langle T_3, h_3\rangle$ muestra que este argumento, el *nuevo,* debe considerarse como derrotado y a la conclusión sustentada por él como no justificada debido a que $\langle T_2, h_2\rangle$ es un bloqueador de $\langle T_3, h_3\rangle$ y por lo tanto no puede ser defendido mediante el uso de $\langle T_1, h_1\rangle$.

El presente ejemplo ilustra la siguiente situación: *una vez establecida una derrota bloqueante entre un par de argumentos, si luego se intentan construir argumentos que sustenten cualquiera de ambas conclusiones y tales argumentos son incomparables con aquellos, entonces, estos argumentos también contarán entre los no justificados*. Lo anterior es razonable pero sólo si el bloqueo es razonable. En este caso, la aceptación de reglas paradójicas '$J(f) \prec F(f)$' y '$G(f) \prec F(f)$' genera la imposibilidad de creer en $G(f)$ o en $\backsim G(f)$. De modo que cuando el sistema cuenta con información paradójica como la señalada, sancionará como derrotado a cualquier argumento que sustente $G(f)$ o $\backsim G(f)$ y lo hará, no porque hayan razones conflictivas irresolubles frente a las conclusiones en cuestión (lo que sería razonable), sino, porque el sistema no cuenta con una manera de manejar este tipo de situaciones, i.e. derrotas bloqueantes originadas por argumentos que sustentan conclusiones contradictorias pero que están basados en las mismas evidencias.

4. Argumentos filtrados en DeLP*

Luego de haber presentado los ejemplos y haber constatado que resultados no intuitivos son obtenidos en *DeLP*, cabe la posibilidad de pensar una estrategia que minimice los resultados contraintuitivos. Pero antes ¿Por qué aparecen los resultados precedentes? ¿Hay algo en común entre todos los ejemplos considerados? ¿De qué modo pueden ser neutralizados? Al respecto diversas ideas se han expuesto en la literatura, y en general han sido tratados como problemas diferentes. Por ejemplo, el caso 1 y 2 ha sido discutido como un problema del restablecimiento por Loui y Stiefvater (1992); Horty (2001); Prakken y Vreeswijk (2002); Prakken (2002); Bodanza y Alessio (2010; 2014); Alessio (2015). Las razones esgrimidas por algunos se basan a cuestiones de representación, otros en cuestiones adicionales. Por otro lado, el caso 3 ha sido discutido también como un problema de restablecimiento por Prakken (2002) y como un problema de separación de fases según Horty (2001). Aunque Dimopoulos et al (2009) y Alessio (2012), han tratado al caso 3 como un problema semántico (en el sentido de Dung, 2015). El caso 4, por su parte, ha sido considerado como un problema de representación por Horty (2012) y como un problema asociado al criterio de preferencia por especificidad en el trabajo propuesto por Bodanza y Alessio (2014). El caso 5 no ha sido tratado como un caso problemático, al menos que el autor tenga noticias al respecto. Las posiciones subyacentes a las propuestas y sugerencias pueden ser resumidas brevemente como sigue:

- Algunos de los principios de la argumentación rebatible no son válidos, como por ejemplo el restablecimiento; o la separación de las fases construcción, comparación y evaluación; o el proceso de justificación presenta algunos problemas (Chesñevar et al, 1994; Horty, 2001; Dimopoulos et al, 2009; Alessio, 2012, Prakken y Vreeswijk, 2002).
- La aplicación de ciertos principios de la argumentación rebatible debe ser limitado frente a ciertos casos (Bodanza y Alessio, 2010).
- El lenguaje formal debe ser más expresivo a fin de que este permita bloquear lo que deba bloquear (Loui y Stiefvater, 1992; Prakken, 2002; Horty, 2012).

- La derrota por especificidad tiene un efecto no neutralizable por el restablecimiento bajo ciertas condiciones (Bodanza y Alessio, 2014;; Alessio, 2015).

Los trabajos citados abordan la cuestión relacionada a los sistemas argumentativos en general y no tanto como problemas de sistemas específicos. Ahora bien, el presente trabajo dirije la cuestión específicamente en lo que respecta al sistema *DeLP* y se pretende proponer una solución ceñida a este formalismo, aunque claramente, la idea puede ser extendida a otros sistemas, pero esto último excede las pretensiones del presente trabajo.

La razón de tal objetivo se debe básicamente a que es llamativo que en *DeLP* aparezcan los problemas destacados ya que *DeLP* es un sistema diseñado para modelar, aunque claro no exclusivamente, razonamiento default, *y* se encuentra dotado de un criterio de resolución de conflictos basado en especificidad (al estilo de sus sucesores como lógica default o redes de herencia), y cuenta con un mecanismo refinado de justificación.

Por otro lado, en general, como se puede constatar en la literatura, no hay acuerdo sobre la causa del problema y se suponen que los ejemplos ilustran problemas diferentes. Sin embargo, parece que los ejemplos problemáticos comparten un rasgo común (RC):

RC: *son argumentos que sustentan conclusiones basadas en reglas que no deberían ser aplicables cuando se atiende a la evidencia total disponible.*

Esto hace que ciertos argumentos impidan la justificación de otros cuando no deberían hacerlo, como en el ejemplo de Ana, o se permite la justificación de argumentos que no deberían contar como tales, como en los ejemplos de *Al*, *Beth* y *Frank*, o la misma evidencia lleva a aceptar afirmaciones contradictorias (como en el caso 5).

Específicamente, en el ejemplo de Al, no es razonable que la regla rebatible: "*Por lo general las aves vuelan*" sea aplicable al caso de Al atendiendo a la información disponible, porque este no es un ave prototípica. Tampoco es razonable que la regla "*Por lo general las gallinas no vuelan*" sea aplicada a Al porque tampoco es una gallina prototípica, nuevamente y atendiendo a la información disponible. Obviamente, Al es una excepción de las excepciones de las aves normales (con respecto al volar) y por ello es razonable creer que vuela.

Por su parte, en el ejemplo de Beth, no es adecuado que las reglas *"Por lo general los empleados de Microsoft son millonarios"* y *"Por lo general los empleados nuevos poseen menos de medio millón"* sean aplicadas para el caso de Beth puesto que es una empleada excepcional en ambos casos.

En el caso de Frank, la regla más específica permite concluir información más general por lo que parece más razonable creer en ella y de alguna manera limita la validez de la regla más general, a pesar de que no hay conflicto entre ellas. En términos más práctico sería extraño que alguien le diga a otro *"busca en la Isla Florida"* cuando en realidad es más razonable decir, *"busca en cualquier parte de la zona de islas remotas",* porque Frank no es un canario del cabo estándar.

Cuando distintas piezas de evidencia llevan a sostener afirmaciones rivales, se dice que tales argumentos se bloquean, lo cual es razonable. En el caso 5, las reglas empleadas se encuentran basadas en la misma evidencia, y tal evidencia lleva a aceptar afirmaciones contradictorias, las reglas que autorizan a realizar tales inferencias deberían estar bloqueadas.

Ahora bien, es interesante, en este contexto, tener en cuenta el criterio de aceptabilidad o razonabilidad de un argumento default. En general se puede considerar a un argumento default como *bueno*, o *aceptable*, cuando no exista información disponible más específica que lo contradiga. Esto es la esencia también del criterio de especificidad como recurso para dirimir argumentos default en conflicto. De modo que, cuando hay un argumento default más específico que contradice a otro, este último ha quedado, a partir de ese momento como un argumento obsoleto o *inaceptable*. Sin embargo, como se observa en los ejemplos problemáticos, los argumentos para los que existen objeciones más específicas siguen desempeñando funciones en el sistema, de hecho son tales que impiden la justificación de argumentos razonables o permiten la justificación de argumentos irrazonables. Esto sugiere que la regla '*a menos que…'* (en ausencia de información contraria más específica) que es emblemática de los sistemas que modelan razonamiento default, no es tomada demasiado en serio en *DeLP*.

Parecería que la causa de esto se debe, tal como se ha discutido ampliamente en (Horty, 2001; Alessio, 2015; Bodanza y Alessio, 2014; 2016), al restablecimiento y a la separación de las fases de construcción y evaluación.

Pero todo ello no es más que la expresión o la forma de manifestarse del problema de fondo: *Argumentos que no deberían interferir en el proceso de justificación, interfieren. Tales argumentos son los argumentos para los que existe información más específica que los contradice* (o que reporta información más general como en el ejemplo de Frank). De modo que la tarea consiste en implementar algún recurso que permita evitar a los *argumentos entrometidos.*

Para lograrlo se empleará un criterio de selección de argumentos posterior a la fase de construcción pero anterior a la de evaluación. Esta idea constituye una extensión de la propuesta en (Alessio, 2015) que fue aplicada a los casos 1 a 4.

El criterio busca seleccionar solo aquellos argumentos no paradójicos para los que no exista evidencia más específica contraria, o para los que no exista evidencia más específica que lleva a aceptar una conclusión de mayor grado de generalidad. En términos más claros, serán seleccionados sólo aquellos argumentos para los que no exista evidencia más específica que lleve a cambiar la conclusión de un argumento. A tales argumentos se los denominará como argumentos máximamente específicos. La razón que motiva este criterio es simple: un razonamiento default está autorizado a inferir cuando las reglas default están activas y lo estarán cuando no exista información más específica contraria disponible. Por otro lado, también parece razonable establecer que una regla default debe estar autorizada a inferir cuando no se dé la situación de que *exista información más específica que lleve a aceptar algo más general*, tal como lo ha ilustrado Horty (2012) con la propuesta del ejemplo de Frank.

Con vistas a implementar la idea y poder seleccionar a tales argumentos, la especificidad de *DeLP* debe ser utilizada de una manera diferente. Ahora se la requerirá para *preseleccionar* argumentos y como tal es necesario que tenga en cuenta si están en conflicto o si se dan casos como en el ejemplo de Frank. Por ello, se le exigirá como condición adicional que se aplique entre argumentos en conflicto, al estilo de Dung y Son (2000) o cuyas conclusiones se impliquen de cierta manera. Esto es expresado en la definición 4.1 que reemplazará a la definición 2.3.

Definición 4.1 (definición 2.3 modificada) Especificidad Generalizada*

Sea $\mathcal{P} = (\Pi, \Delta)$ un *de.l.p.* Sea Π_G es el conjunto de todas las reglas estrictas (sin incluir hechos). Sea $\mathcal{F}$ el conjunto de todos los literales para los que existe una derivación rebatible a partir de $\mathcal{P}$ ($\mathcal{F}$ será considerado un conjunto de hechos). Sean $\langle T_1, h_1\rangle$ y $\langle T_2, h_2\rangle$ dos argumentos obtenidas a partir de $\mathcal{P}$. Se dirá que $\langle T_1, h_1\rangle$ es *estrictamente más específico** que $\langle T_2, h_2\rangle$, notado como $\langle T_1, h_1\rangle \succ_{esp} \langle T_2, h_2\rangle$, si y sólo si:

i. Para todo $H \subseteq \mathcal{F}$: Si $\Pi_G \cup H \cup T_1 \mid\!\sim h_1$ y $\Pi_G \cup H \nvdash h_1$ entonces $\Pi_G \cup H \cup T_2 \mid\!\sim h_2$, yexiste al menos un $H' \subseteq \mathcal{F}$ tal que $\Pi_G \cup H' \cup T_2 \mid\!\sim h_2$ y $\Pi_G \cup$ H' $\nvdash h_2$ y $\Pi_G \cup H' \cup T_1 \not\mid\!\sim h1$, y

ii. o bien $\Pi_G \cup \{h_1, h_2\} \vdash \perp$ o bien $\Pi_G \cup \{h_2\} \vdash h_1$ y $\Pi_G \cup \{h_1\} \nvdash h_2$.

□

i. dice que todo lo que activa a $\langle T_1, h_1\rangle$ activa a $\langle T_2, h_2\rangle$ y que existe algo que activa a $\langle T_2, h_2\rangle$ que no activa a $\langle T_1, h_1\rangle$. *ii.* exige que la comparación se haga entre argumentos en conflicto o que el argumento más general implique la conclusión del más específico.

El término *activar*, o *conjunto de activadores* se usa para referirse a un conjunto H de hechos, cuando se lo emplea junto a una estructura de argumento $\langle T, h\rangle$ y el conjunto de reglas estrictas Π_G para derivar el literal h de $\langle T, h\rangle$. Debe tenerse en cuenta que el conjunto de activadores no es trivial y esto quiere decir que no puede derivarse el literal del argumento sólo con la presencia de H y Π_G. Es mandatorio que sea empleado al menos una regla rebatible perteneciente al conjunto T. Una definición precisa de *activador* será útil para más adelante.

Definición 4.2 Activador

Sea $\mathcal{P} = (\Pi, \Delta)$ un *de.l.p.* Sea $\mathcal{F}$ el conjunto de todos los literales para los que existe una derivación rebatible a partir de $\mathcal{P}$ y sea $\langle T, h\rangle$ una estructura de argumento obtenidas a partir de $\mathcal{P}$. Se dirá que un conjunto $H \subseteq \mathcal{F}$ *activa* o es un *activador* de $\langle T, h\rangle$ si y sólo si $\Pi_G \cup H \cup T \mid\!\sim h$ y $\Pi_G \cup H \nvdash h$. La clase de todos los subconjuntos H que activan a $\langle T, h\rangle$ será denotada como $Act(\langle T, h\rangle)$

□

La definición 4.1 permitirá inhibir el comportamiento de los casos 1 a 4, sin embargo, el caso 5 no cae dentro de la caracterización. Para hacerlo se propondrá una definición que pretende capturar la idea general que el caso 5 ilustra. La situación ejemplificada por tal caso consiste en que la misma información, o *el mismo conjunto de activadores*, lleva a aceptar afirmaciones contradictorias. Lo cual es extraño al menos desde un punto de vista intuitivo.

Definición 4.3 Argumento paradójico

Sea $\boldsymbol{\mathcal{P}} = (\Pi, \Delta)$ un *de.l.p.* Sea $\boldsymbol{\mathcal{F}}$ el conjunto de todos los literales para los que existe una derivación rebatible a partir de $\boldsymbol{\mathcal{P}}$. Sean $\langle T_1, h_1\rangle$ y $\langle T_2, h_2\rangle$ dos argumentos obtenidos a partir de $\boldsymbol{\mathcal{P}}$. Se dirá que $\langle T_1, h_1\rangle$ es un *argumento paradójico* si y sólo si existe un argumento $\langle T_2, h_2\rangle$ tal que $Act(\langle T_1, h_1\rangle) = Act(\langle T_2, h_2\rangle)$ y $\Pi_G \cup \{h_1, h_2\} \vdash \perp$.

□

Ahora bien, una vez que se sabe cuáles son los argumentos más específicos y cuáles los paradójicos, se procederá a seleccionar aquellos para los que no existen excepciones o información paradójica. Previamente será importante identificarlos. Para ello se propone la definición 4.4.

Definición 4.4 Máxima Especificidad

Una argumento $\langle T_1, h_1\rangle$ se dice *máximamente específico* con respecto a h_1 si y sólo si $\langle T_1, h_1\rangle$ no es paradójico y no existe un argumento $\langle T_2, h_2\rangle$ tal que $\langle T_2, h_2\rangle$ sea estrictamente más específico* que $\langle T, h\rangle$ para cualquier $\langle T, h\rangle$ tal que $\langle T, h\rangle$ es un sub-argumento de $\langle T_1, h_1\rangle$.

□

Ya se sabe cuáles son los máximamente específicos, estos argumentos, serán coleccionados en un conjunto, el conjunto de los máximamente específicos (Definición 4.5).

Definición 4.5. Conjunto de Argumentos Máximamente Específicos

Dado el conjunto *ARG* de todas las estructuras argumentativas que pueden construirse en base a un programa $\boldsymbol{\mathcal{P}}$, se define: $ARG^* =_{\text{df}} \{\langle T, h\rangle \in ARG$: $\langle T, h\rangle$ es máximamente específico con respecto a $h\}$.

□

El conjunto *ARG*, es el conjunto de todos los argumentos que se construyen en función de la definición 2.1. Un subconjunto de tal conjunto, el conjunto *ARG** está constituido por todos aquellos argumentos que son máximamente específicos, según la definición 4.4. La modificación propuesta para *DeLP*, *DeLP*,* comenzará el proceso de comparación y continuará con la evaluación exclusivamente con los miembros de *ARG**.

*DeLP** sólo determinará la justificación de los argumentos entre los miembros de *ARG**. Todos los argumentos que no verifican la condición de *máxima especificidad* son simplemente dejados fuera del proceso restante. Solo los miembros de *ARG** se contra-argumentan, derrotan, prefieren o se encuentran involucrados en un proceso de justificación. Esto llevara a que la estructura de *DeLP* sea ahora, en *DeLP*,* la siguiente:

Fase 1: Construcción de argumentos
Fase 2: Selección de los máximamente específicos
Fase 3: Contra-argumentación y derrota.
Fase 4: Justificación.

Como puede fácilmente evidenciarse en el esquema presentado, las relaciones ya sean de contra-argumentación o derrota se darán exclusivamente entre argumentos máximamente específicos.

Definición 4.6 Ámbito de la Contra-argumentación* y Derrota*

Si *R* es una relación de contra-argumentación o derrota, $R \subseteq ARG^* \times ARG^*$

□

La noción de línea de argumentación también requerirá una modificación.

Definición 4.7 (definición 2.9 modificada). Línea de argumentación*

Sea $\mathcal{P} = ((\Pi, \Delta)$ un *delp* y $\langle T_0, h_0\rangle$ una estructura de argumento máximamente específica obtenida a partir de $\mathcal{P}$. Una *línea de argumentación* para $\langle T_0, h_0\rangle$ es una secuencia de estructuras de argumento a partir de $\mathcal{P}$, notado como $\Lambda=[\langle T_0, h_0\rangle, \langle T_1, h_1\rangle, \langle T_2, h_2\rangle, \langle T_3, h_3\rangle,\ldots]$,donde cada elemento de la secuencia $\langle T_i, h_i\rangle$, i > 0, es un derrotador de su predecesor $\langle T_{i-1}, h_{i-1}\rangle$

□

La modificación de la definición 2.9, en la definición 4.7, es sólo en relación al punto de inicio de la línea de argumentación "$\langle T_0, h_0\rangle$ es *una estructura de argumento máximamente específica*".

Nótese que no fue necesario explicitar la condición de máxima especificidad para $\langle T_i, h_i\rangle$ porque la definición 4.6 exige, para que $\langle T_i, h_i\rangle$ sea un derrotador, que $\langle T_i, h_i\rangle \in ARG^*$.

La definición 2.10 no requiere que sea ajustada debido a que presupone a la definición 4.7. Ahora bien, la definición 2.11, de argumentos concordantes, relevante para el proceso de justificación, exige la condición de consistencia entre los argumentos, dos argumentos pueden ser consistentes pero uno de ellos puede no ser máximamente específico, de modo que esta definición requiere una actualización.

Definición 4.8 (definición 2.11 modificada). Argumentos concordantes*

Sea $\mathcal{P} = ((\Pi, \Delta)$ un *delp* y $\langle T_1, h_1\rangle$ y $\langle T_2, h_2\rangle$ estructuras de argumento máximamente específicos obtenidas a partir de $\mathcal{P}$. Se dirá que $\langle T_1, h_1\rangle$ y $\langle T_2, h_2\rangle$ son *argumentos concordantes* si y sólo si el conjunto $\Pi \cup T_1 \cup T_2$ no es contradictorio. Más generalmente un conjunto de estructuras de argumentos máximamente específicos $\{\langle T_i,h_i\rangle\}\}_{i=1}^{n}$ es concordante si y sólo si $\Pi\cup\bigcup_{i=1}^{n} T_i$ no

es contradictorio.

□

Las otras nociones, subsidiarias de estas no requieren una atención especial. Ahora bien, resta preguntarse si una estructura de argumento puede ser justificada si no es máximamente especifica. Como es de esperar, la respuesta a tal situación debe ser negativa, ningún argumento que no sea máximamente específico debe contar como una justificación para la afirmación que sustenta. Al mismo tiempo, todo argumento justificado será un argumento máximamente específico. Las siguientes definiciones pretenden capturar esta intención.

Definición 4.9 (definición 2.13 modificada). Árbol Dialéctico*

Sea $\mathcal{P}$ =(Π, Δ) un *delp* y $\langle T_0, h_0\rangle$ una estructura de argumento máximamente específica obtenida a partir de $\boldsymbol{\mathcal{P}}$. Un *árbol dialéctico* para $\langle T_0, h_0\rangle$, notado $\boldsymbol{\mathcal{T}}\langle T_0, h_0\rangle$, es definido como sigue:

- La raíz del árbol es etiquetado con $\langle T_0, h_0\rangle$.
- Sea N un nodo no raíz del árbol etiquetado $\langle T_n, h_n\rangle$, y Λ= $[\langle T_0, h_0\rangle,\langle T_1, h_1\rangle,\langle T_2, h_2\rangle,\ldots, \langle T_n, h_n\rangle]$ la secuencia de etiqueta de los caminos de la raíz a N.
  Sean $\langle S_1, j_1\rangle,\langle S_2, j_2\rangle,\ldots, \langle S_k, j_k\rangle$ todos los derrotadores para $\langle T_n, h_n\rangle$.
  Para cada derrotador $\langle S_i, j_i\rangle$ ($1 \leq i \leq$ k), tal que, la línea de argumentación Λ'=$[\langle T_0, h_0\rangle,\langle T_1, h_1\rangle,\langle T_2, h_2\rangle,\ldots,\langle T_n, h_n\rangle,\langle S_i, j_i\rangle]$sea aceptable, existe un nodo hijo N_i de N etiquetado con $\langle S_i, j_i\rangle$.
  Si no existe ningún derrotador para $\langle T_n, h_n\rangle$ o no existe un $\langle S_i, j_i\rangle$ tal que Λ' es aceptable entonces N es una hoja.

La modificación de la definición 2.13 está dada en las condiciones iniciales "$\langle T_0, h_0\rangle$ *una estructura de argumento máximamente específica*", de la misma manera que en la definición 4.7. La razón de ello es sencilla. Las derrotas sólo se dan entre los argumentos máximamente específicos, entonces sería posible construir un árbol dialéctico para un argumento que no fuera máximamente específico, argumento que no tendría derrotadores porque como ya se dijo, la derrota es sólo una relación entre los miembros de *ARG**. De modo que en fin de cuentas, un argumento no máximamente especifico estaría también justificado. Exigir entonces la construcción del árbol solo para los máximamente específicos parece bastante razonable. Como podrá advertirse, a los nodos no raíces no debe exigirse, en la definición 4.9, la máxima especificidad, porque para ser un nodo no raíz, este debe ser un derrotador de su padre, y si es un derrotador, según definición 4.6, debe ser un argumento máximamente específico.

El marcado del árbol dialéctico será de la misma manera que en *DeLP* estándar porque en la definición 2.14 se dan las instrucciones para marcar un árbol dialéctico que desde ahora es un árbol dialéctico para argumentos máximamente específicos.

De lo dicho hasta aquí también se sigue el siguiente teorema, que también justifica la no necesidad de la modificación de la definición 2.15.

Teorema 1

Sea $\langle T, h\rangle$ una estructura de argumento. Si T es una justificación para h, $\langle T, h\rangle$ es máximamente específico.

Prueba. Si T es una justificación para h, por definición 2.15, $\langle T, h\rangle$ ha sido marcado como 'U' y para que $\langle T, h\rangle$ esté marcado como 'U' en $\boldsymbol{\mathcal{T}}^*\langle T, h\rangle$ según la definición 2.14 cualquier hijo de $\langle S, j\rangle$ es marcado como 'D' en $\boldsymbol{\mathcal{T}}^*\langle T, h\rangle$, pero además el árbol dialéctico $\boldsymbol{\mathcal{T}}\langle T, h\rangle$ de $\langle T, h\rangle$ que ha sido marcado luego como $\boldsymbol{\mathcal{T}}^*\langle T, h\rangle$, se construye a partir de $\langle T, h\rangle$ si $\langle T, h\rangle$ es máximamente específico, de modo que no podría ser que T sea una justificación para *h y* $\langle T, h\rangle$ no sea máximamente específico.

■

Teorema 2

Sea $\langle T, h\rangle$ una estructura de argumento. Si $\langle T, h\rangle$ no es máximamente específico, T no es una justificación para h.

Prueba. Si $\langle T, h\rangle$ no es máximamente específico, entonces no puede construirse un árbol dialéctico en torno a $\langle T, h\rangle$, y si no puede construirse un árbol dialéctico en torno a $\langle T, h\rangle$ no puede etiquetarse con 'U' por lo que tampoco podrá considerarse a T como una justificación para h.

■

En la sección siguiente se ilustrará el comportamiento de *DeLP** frente a los casos problemáticos y a una serie de ejemplos interesantes. También se intentará caracterizar las variaciones entre los argumentos finalmente derrotados en *DeLP* y *DeLP** como aquellos argumentos finalmente no derrotados en ambas versiones. Esto permitirá tener una adecuada percepción de las modificaciones generadas a partir de la definición 4.1.

Antes de pasar a la sección siguiente, será importante recordar que en *DeLP** pueden existen otros tipos de derrotas además de la basada en especificidad. Por ello, podrán existir derrotas asimétricas independientemente de que las derrotas estrictas basadas en especificidad sean inexistentes, puesto que el proceso de preselección elimina tales relaciones entre los argumentos. En base a esto puede demostrarse el siguiente teorema.

Teorema 3

Sean $\langle T, h\rangle$ y $\langle S, j\rangle$ estructuras de argumentos que pertenezcan a ARG^*, si $\langle T, h\rangle$ derrota a $\langle S, j\rangle$, $\langle T, h\rangle$ no es más específico que $\langle S, j\rangle$.

Prueba. Si $\langle T, h\rangle$ derrota a $\langle S, j\rangle$, $\langle T, h\rangle$ no es más específico que $\langle S, j\rangle$ porque si $\langle T, h\rangle$ fuese más específico que $\langle S, j\rangle$, $\langle S, j\rangle$ no sería máximamente específico, es decir, $\langle S, j\rangle$ no pertenece a ARG^*, lo cual es contradice el supuesto.

■

5. Representación del conocimiento en *DeLP**

Tanto en el caso 1, como en el caso 2 se señaló que el argumento $\langle A_1, h_1\rangle$ debe considerarse como derrotado porque existe evidencia más específica que permite inferir una conclusión contradictoria a la sustentada por el más general, a saber, $\langle A_2, h_2\rangle$.

En el caso 1, *DeLP** sancionará como único argumento máximamente específico a $\langle A_3, h_3\rangle$, i.e. $\langle\{F(a) \prec WC(a)\}, F(a)\rangle$ debido a que $\langle A_2, h_2\rangle$ cuenta con un derrotador más específico y $\langle A_1, h_1\rangle$ también. Esta situación los deja fuera del proceso de comparación y evaluación. El único argumento para el que se puede construir un árbol dialéctico es $\langle A_3, h_3 \rangle$. Como tal, el árbol tendrá un solo nodo, y estará etiquetado como 'U'.

En el caso 2, $\langle \{<1/2M(b) \prec NMEX(b)\}, <1/2M(b) \rangle$ es el único argumento justificado debido a que tanto $\langle A_2, h_2 \rangle$ como $\langle A_1, h_1 \rangle$ no verifican la condición de ser máximamente específico.

El caso 3, el ejemplo de Ana, es un ejemplo interesante porque permite ilustrar un cambio importante: *DeLP** permitirá la justificación de argumentos que en *DeLP* contaban como derrotados. En *DeLP** tanto $\langle A_3,h_3 \rangle$ y $\langle A_4,h_4 \rangle$ son ambos argumentos máximamente específicos i.e. los argumentos que sostienen que Ana no es adinerada. Dado que para $\langle \{\backsim AD(a) \prec DP(a)\}, \backsim \mathrm{AD}(a) \rangle$ y $\langle \{\backsim AD(a) \prec RB(a)\}, \backsim \mathrm{AD}(a) \rangle$ no hay argumentos que los derroten en el proceso de justificación, podrán contar como justificados. $\langle A_1, h_1 \rangle$ y $\langle A_2,h_2 \rangle$ no pueden interferir en la justificación de $\langle A_3,h_3 \rangle$ y $\langle A_4,h_4 \rangle$ simplemente porque no califican como máximamente específicos y por lo tanto, no pueden participar en el proceso de comparación y evaluación de argumentos.

En el caso 4 se puede apreciar un cambio radical en el comportamiento de *DeLP** con respecto a *DeLP*. Ahora es posible que argumentos consistentes no puedan contar como justificados. En el ejemplo de Frank, sólo es máximamente específico $\langle \{NCI(f) \prec CCR(f)\}, NCI(f) \rangle$. Al no tener derrotadores, $NCI(f)$ está justificado en $NCI(f) \prec CCR(f)\}$. Por su parte, el argumento $\langle \{NIF(f) \prec CC(f)\}, NIF(f) \rangle$ no califica para el proceso de comparación y evaluación de argumentos por no ser máximamente específico, debido a que existe un argumento más específico que sanciona una conclusión más general. El comportamiento de *DeLP** frente a ejemplos de este tipo es un comportamiento más cauto que *DeLP,* puesto que prefiere creer en algo más general pero mejor sustentado (más específico).

El caso 5 ilustra una situación interesante. Tal ejemplo muestra que la misma información, *el mismo conjunto de activadores*, lleva a aceptar afirmaciones contradictorias. Esto, en sí mismo no es problemático porque en el fondo, el proceso de justificación de *DeLP* impide que las conclusiones de ambos argumentos, frente a un programa como el definido, se sancionen como justificadas. El inconveniente aparece cuando hay al menos un argumento incomparable con aquellos. Cuando tal situación se da, *DeLP* impide también la aceptación de tal argumento. Pero bajo estas condiciones esto no es razonable, porque los argumentos que impiden la aceptación del nuevo argumento, son argumentos paradójicos. Con vistas a impedir que tales argumentos se entrometan en el estado de otros argumentos se estableció la condición de que un argumento paradójico no puede calificar para el proceso de comparación y evaluación y tal noción fue incorporada a la de *argumento máximamente específico*. Por ello, si el programa es:

$$\Delta = \{G(x) \mathrel{-\!\!\prec} F(x),\ J(x) \mathrel{-\!\!\prec} F(x),\ G(x) \mathrel{-\!\!\prec} H(x)\}$$
$$\Pi_{\mathrm{G}} = \{\sim G(x) \longleftarrow J(x)\}$$
$$\Pi_{\mathrm{F}} = \{\mathrm{F}(j), \mathrm{H}(j)\}$$

y los argumentos construidos son:

$\langle A_1,h_1\rangle$	$\langle\{G(f) \mathrel{-\!\!\prec} F(f)\}, G(f)\rangle$
$\langle A_2,h_2\rangle$	$\langle\{\sim G(f) \longleftarrow J(f), J(f) \mathrel{-\!\!\prec} F(f)\}, \sim G(f)\rangle$
$\langle A_3,h_3\rangle$	$\langle\{G(f) \mathrel{-\!\!\prec} H(f)\}, G(f)\rangle$

sólo es máximamente específico $\langle A_3,h_3\rangle$ y dado que no contará con derrotadores, $\langle A_3,h_3\rangle$ será rotulado como 'U' y por lo tanto, la conclusión h_3 podrá considerarse como una conclusión justificada. Por su parte, $\langle A_1,h_1\rangle$ y $\langle A_2,h_2\rangle$ simplemente no califican para el proceso de comparación y evaluación por no ser máximamente específicos. Lo importante de esto, es que tales argumentos no podrán interferir en el estado de otros argumentos.

A continuación se presentan una serie de ejemplos que ilustran el comportamiento de *DeLP** en contraste con *DeLP*.

Ejemplo 5.1

$$\Delta = \{P(x) \mathrel{-\!\!\prec} A(x),\ Q(x) \mathrel{-\!\!\prec} B(x),\ R(x) \mathrel{-\!\!\prec} C(x),\ S(x) \mathrel{-\!\!\prec} E(x),\ \sim R(x) \mathrel{-\!\!\prec} C(x) \wedge P(x),\ \sim Q(x) \mathrel{-\!\!\prec} B(x) \wedge R(x),\ \sim S(x) \mathrel{-\!\!\prec} D(x) \wedge Q(x)\}$$
$$\Pi = \{A(a), B(a), C(a), D(a), E(a)\}$$

$\langle A_1,h_1\rangle$	$\langle\{P(a) \mathrel{-\!\!\prec} A(a); \sim R(a) \mathrel{-\!\!\prec} C(a) \wedge P(a)\}, \sim R(a)\rangle\}$
$\langle A_2,h_2\rangle$	$\langle\{R(a) \mathrel{-\!\!\prec} C(a); \sim Q(a) \mathrel{-\!\!\prec} B(a) \wedge R(a)\}, \sim Q(a)\rangle$
$\langle A_3,h_3\rangle$	$\langle\{Q(a) \mathrel{-\!\!\prec} B(a); \sim S(a) \mathrel{-\!\!\prec} D(a) \wedge Q(a)\}, \sim S(a)\rangle$
$\langle A_4,h_4\rangle$	$\langle\{S(a) \mathrel{-\!\!\prec} E(a)\}, S(a)\rangle$

En *DeLP,* el argumento $\langle A_4,h_4\rangle$ no contaría como justificado, sin embargo, en *DeLP** sí debido a que $\langle A_3,h_3\rangle$ no podría derrotarlo ya que no califica como máximamente específico. Este ejemplo ilustra el hecho de que un argumento finalmente derrotado para *DeLP* está finalmente no derrotado en *DeLP**

Ejemplo 5.2

$\Delta = \{P(x) \prec A(x),\ Q(x) \prec B(x),\ R(x) \prec C(x),\ S(x) \prec D(x),\ \backsim R(x) \prec C(x) \wedge P(x),\ \backsim Q(x) \prec B(x) \wedge R(x),\ \backsim S(x) \prec D(x) \wedge Q(x),\ \backsim T(x) \prec E(x) \wedge S(x)\}$
$\Pi = \{A(a), B(a), C(a), D(a), E(a)\}$

$\langle A_1,h_1\rangle \qquad \langle\{P(a) \prec A(a); \backsim R(a) \prec C(a) \wedge P(a)\}, \backsim R(a)\rangle\}$
$\langle A_2,h_2\rangle \qquad \langle\{R(a) \prec C(a); \backsim Q(a) \prec B(a) \wedge R(a)\}, \backsim Q(a)\rangle$
$\langle A_3,h_3\rangle \qquad \langle\{Q(a) \prec B(a); \backsim S(a) \prec D(a) \wedge Q(a)\}, \backsim S(a)\rangle$
$\langle A_4,h_4\rangle \qquad \langle\{S(a) \prec D(a); \backsim T(a) \prec E(a) \wedge S(a)\}, \backsim T(a)\rangle$

En *DeLP*,* el único argumento que contaría como justificado, en el ejemplo 5.2 es $\langle A_1,h_1\rangle$. Sin embargo, en *DeLP,* $\langle A_1,h_1\rangle$ y $\langle A_3,h_3\rangle$ deberían contar entre los no derrotados. Este ejemplo ilustra el hecho de que un argumento finalmente no derrotado para *DeLP* está finalmente derrotado en *DeLP**.

Ejemplo 5.3

$\Delta = \{V(x) \prec A(x),\ \backsim V(x) \prec P(x)\}$
$\Pi = \{P(t), A(x) \longleftarrow P(x)\}$

$\langle A_1,h_1\rangle \qquad \langle\{V(t) \prec A(t)\}, V(t)\rangle\}$
$\langle A_2,h_2\rangle \qquad \langle\{\backsim V(t) \prec P(t)\}, \backsim V(t)\rangle$

El ejemplo 5.3 ilustra el ejemplo canónico de Tweety, como podrá notarse, en *DeLP*,* el único argumento justificado, al igual que en *DeLP,* es $\langle A_2,h_2\rangle$ como es esperable.

Ejemplo 5.4
$\Delta = \{P(x) \prec C(x),\ \backsim P(x) \prec R(x)\}$
$\Pi = \{C(n), R(n)\}$

$\langle A_1, h_1\rangle$ $\langle\{P(n) \prec C(n)\}, P(n)\rangle\}$

$\langle A_2, h_2\rangle$ $\langle\{\sim P(n) \prec R(n)\}, \sim P(n)\rangle$

El ejemplo 5.4 ilustra el ejemplo canónico de Nixon, como podrá notarse, *DeLP* y DeLP* obtiene el mismo resultado como es esperable, ninguno de los dos argumentos puede, finalmente, considerarse como no derrotado.

El análisis de los ejemplos y la consideración del comportamiento de DeLP* frente a los casos problemáticos permite visualizar que las modificaciones propuestas llevan a aceptar argumentos en *DeLP** que en *DeLP* eran rechazados y se rechazan argumentos en *DeLP** que en *DeLP* son aceptados. Esto hace que el conjunto de argumentos que constituyen una justificación para la conclusión que sustentan de *DeLP* y *DeLP** sean incomparables. Sin embargo es posible caracterizar las variaciones. A continuación se propondrá una definición que permitirá facilitar la caracterización de los cambios.

Definición 4.10. Conjunto de Argumentos derrotados y no derrotados

Sea *ARG* el conjunto de todas las estructuras de argumentos construidas a partir de un programa $\mathcal{P}$, sea *ARG** el conjunto de argumentos que son máximamente específicos. A continuación se definirán el conjunto de argumentos no derrotados y derrotados para *ARG* y *ARG**

- $U\text{-}ARG =_{df} \{\langle T, h\rangle \in ARG$: La raíz de $\mathcal{T}^*\langle T, h\rangle$ es marcada como 'U'$\}$
- $D\text{-}ARG =_{df} \{\langle T, h\rangle \in ARG$: La raíz de $\mathcal{T}^*\langle T, h\rangle$ es marcada como 'D'$\}$
- $U\text{-}ARG^* =_{df} \{\langle T, h\rangle \in ARG^*$: La raíz de $\mathcal{T}^*\langle T, h\rangle$ es marcada como 'U'$\}$
- $D\text{-}ARG^* =_{df} \{\langle T, h\rangle \in ARG^*$: La raíz de $\mathcal{T}^*\langle T, h\rangle$ es marcada como 'D'$\}$

□

Teorema 4 ¿Qué argumentos aceptados en *DeLP*, son rechazados en *DeLP**?

Sea $\langle T, h\rangle$ una estructura de argumento. Si $\langle T, h\rangle \in$ *U-ARG* en DeLP y $\langle T, h\rangle \in$ *D-ARG** en *DeLP** entonces, existe al menos un argumento $\langle S, j\rangle \in$ *ARG* tal que, o

i. $\langle S, j\rangle$ contra-argumenta a $\langle T, h\rangle$ y $\langle S, j\rangle$ es estrictamente más específico que $\langle T, h\rangle$, o

ii. $\Pi_G \cup \{h\} \vdash j$ y $\Pi_G \cup \{j\} \nvdash h$ y $\langle S, j\rangle$ es estrictamente más específico que $\langle T, h\rangle$, o

iii. $Act(\langle S, j\rangle) = Act(\langle T, h\rangle)$ y $\Pi_G \cup \{j, h\} \vdash \bot$.

Prueba. $\langle T, h\rangle$ podría verificar una de las tres condiciones apuntadas y en DeLP calificar como un argumento aceptado. Más en particular:

- Si existiera un contraargumento más específico que $\langle T, h\rangle$, en *DeLP*, $\langle T, h\rangle$ podría calificar como miembro de *U-ARG* siempre y cuando, $\langle S, j\rangle$ fuese a su vez derrotado. Sin embargo, el sólo hecho de que $\langle S, j\rangle$ derrote por especificidad a $\langle T, h\rangle$ hace que $\langle T, h\rangle \in$ *D-ARG**.
- Si $Act(\langle S, j\rangle) = Act(\langle T, h\rangle)$ y $\Pi_G \cup \{j, h\} \vdash \bot$ $\langle T, h\rangle$ podría calificar como miembro de *U-ARG* si contase con una defensa de $\langle S, j\rangle$, sin embargo, *y a pesar de ello*, $\langle T, h\rangle$ contará como miembro del conjunto *D-ARG** por tratarse de un argumento paradójico.
- Si existiera un argumento $\langle S, j\rangle$ más específico que $\langle T, h\rangle$ y a su vez $\Pi_G \cup \{h\} \vdash j$ y $\Pi_G \cup \{j\} \nvdash h$, y $\langle T, h\rangle$ no tuviese ningún derrotador, $\langle T, h\rangle$ contaría como miembro de *U-ARG*, sin embargo, bajo la misma condición, $\langle T, h\rangle$ contaría entre los miembros de *D-ARG**.

■

Teorema 5 ¿Qué argumentos rechazados en *DeLP*, son aceptados en *DeLP**?

Sea $\langle T, h\rangle$ una estructura de argumento. Si $\langle T, h\rangle \in$ *D-ARG* en *DeLP* y $\langle T, h\rangle \in$ *U-ARG** en *DeLP** entonces, existe al menos un argumento $\langle S, j\rangle \in$ *ARG* tal que *i.* $\langle S, j\rangle$ es un bloqueador de $\langle T, h\rangle$ en *DelP*, pero $\langle S, j\rangle$ es paradójico y $\langle T, h\rangle$ no lo es, o

ii. $\langle S, j\rangle$ es un bloqueador de $\langle T, h\rangle$ en *DelP*, pero $\langle S, j\rangle$ es contra-argumentado por $\langle T', h'\rangle$ y $\langle T', h'\rangle$ es estrictamente más específico que $\langle S, j\rangle$.

Prueba. Si $\langle T, h\rangle$ es bloqueado en *DeLP* por $\langle S, j\rangle$, y $\langle S, j\rangle$ no es derrotado por un argumento más específico, pero sí está bloqueado por otro argumento $\langle T', h'\rangle$ tal que $Act(\langle S, j\rangle) = Act(\langle T', h'\rangle)$ y $\Pi_G \cup \{j, h'\} \vdash \perp$, $\langle T, h\rangle$ estará derrotado en *DeLP*, pero podrá contar como miembro del conjunto *U-ARG** porque su bloqueador no interferirá en el proceso de justificación de *DeLP** debido a que $\langle S, j\rangle$ no es máximamente específico.

Si $\langle T, h\rangle$ es bloqueado en *DeLP* por $\langle S, j\rangle$, y $\langle S, j\rangle$ es derrotado por un argumento más específico $\langle T', h'\rangle$, entonces debe existir un argumento $\langle T'', h''\rangle$ que sea un contraargumento más específico de $\langle T', h'\rangle$, de modo tal que $\langle T, h\rangle$ bajo tal circunstancia contará como miembro del conjunto *D-ARG*. Sin embargo, y bajo la misma circunstancia, $\langle T, h\rangle$ contará como miembro de *U-ARG** debido a que $\langle S, j\rangle$ *y* $\langle T', h'\rangle$ no calificarán para el proceso de comparación y evaluación en DeLP* y por lo tanto no interferirán en el proceso de justificación de $\langle T, h\rangle$

■

Restará, para un trabajo futuro responder a los siguientes interrogantes, ¿Qué argumentos aceptados en *DeLP*, son aceptados en *DeLP** y qué argumentos rechazados en *DeLP*, son rechazados en *DeLP**?

Conclusión

En el presente trabajo se han propuesto una serie de casos en los que *DeLP* obtiene resultados no intuitivos. Luego, se presentó brevemente la discusión que frente a tales ejemplos se ha dado en la literatura. Se detectó que en general no hay un tratamiento unitario de los mismos, sino que son estudiados de manera aislada porque se supone que las causas de cada caso son diferentes. Ahora bien, en el presente artículo se sostuvo que el problema se debe a que en *DeLP* no se consideraba seriamente la regla '*a menos que información contraria más específica esté disponible….*' que regula la aceptabilidad de razonamientos default. Al no ser considerada seriamente, se permite que ciertos argumentos interfieran en el estado de otros argumentos cuando no deberían hacerlo. Para evitarlo se propuso un criterio de selección de los argumentos que verifican la condición de ser *máximamente específicos*, o dicho de otra manera, son argumentos cuyas reglas defaults están activas porque no hay información más específica que indique lo contrario, o que lleve a concluir algo de mayor generalidad, o que la misma información lleve a aceptar afirmaciones contradictorias. Una vez que los argumentos son seleccionados se procede a compararlos y a evaluarlos como es usual en *DeLP*. Las conclusiones de los argumentos máximamente específicos que sean considerados como los argumentos que prevalecen frente a sus rivales, serán los literales justificados en *DeLP,* ahora denominado *DeLP**. De modo que en *DeLP**, todos los argumentos justificados son máximamente específicos, puesto que claramente todos aquellos que no verifican la propiedad han sido descartados previamente.

Referencias

ALESSIO, C.A. *Restablecimiento y especificidad en sistemas argumentativos.* Disertación doctoral, Universidad Nacional del Sur, Bahía Blanca: Argentina, 2015.

——— Un sistema argumentativo para razonamiento default: discusión y propuesta. En *II Encuentro de Jóvenes Investigadores: consolidando espacios del quehacer científico en San Juan*, 2013.

——— Una propuesta para modelar razonamiento default en sistemas argumentativos. En Milone, R. A. y Torres, J.M. (Eds.) (2012), pp. 19-26.

——— Aceptabilidad extendida para marcos argumentativos abstractos. *Epistemología e Historia de la Ciencia,* vol. 17, 2012.

——— Extensiones no justificadas en Lógica Default. En *Epistemología e historia de la ciencia*, vol. 15, pp. 13-19, 2010.

——— Razonamiento rebatible: origen, historia, estrategias. *Revista Cuadernos de la Universidad*, *42*, pp. 15-35, 2009.

ANTONIOU, G. Defeasible reasoning: A discussion of some intuitions. *International journal of intelligent systems* 21, 6, pp. 545-558, 2006

BARONI, P., & GIACOMIN, M. Semantics of abstract argument systems. In Rahwan, I. & Simari, G. R. (Eds.) (2009), pp. 25-44.

BENFERHAT, S. Computing Specificity in Default Reasoning. En Gabbay, D.; P. Smets (Eds) (2000), pp. 147-177.

BODANZA, G. A. Dudas razonables sobre restablecimiento. *Epistemología e historia de la ciencia,* vol., 17, pp. 94-101, 2011.

——— & ALESSIO, C. A. Sobre la aceptabilidad de argumentos en un marco argumentativo con especificidad. *Actas de la II Conferencia Internacional de Argumentación y Pensamiento Crítico*, pp. 74-81, 2010.

——— & ——— Reinstatement and the Requirement of Maximal Specificity in Argument Systems. *Logic, Language, Information, and Computation*, vol. 8652, pp. 81-93, 2014.

——— &——— Rethinking Specificity in Defeasible Reasoning and Its Effect in Argument Reinstamentent, manuscript, 2016.

BONDARENKO, A., DUNG, P. M., KOWALSKI, R. A., & TONI, F. An abstract, argumentation-theoretic approach to default reasoning. *Artificial intelligence*, 93(1), pp. 63-101, 1997.

CAMINADA, M. For the Sake of the Argument; explorations into argument-based reasoning. *PhD thesis Vrije Universiteit Amsterdam*, 2004.

CHESÑEVAR, C. I, SIMARI, G.R., GARCÍA, A.J. The Role of Dialectics in Defeasible Reasoning". In *Proc. of the XIV Conferencia Internacional de la Sociedad Chilena de Ciencia de la Computación*, pp. 335-344, 1994.

———, MAGUITMAN, A. G., & LOUI, R. P. Logical models of argument. *ACM Computing Surveys*, 32(4), pp. 337-383, 2000.

DIMOPOULOS, Y.; MORAITIS, P.; AMGOUD L. Extending Argumentation to Make Good Decisions. In *Proc. of ADT' 2009*, pp. 225-236, 2009.

DUNG, P.M. On the acceptability of arguments and its fundamental role in nonmonotonic reasoning, logic programming and n person games. *Artificial intelligence*, 77(2), pp. 321-357, 1995.

——— & SON, T. C. Default reasoning with specificity. *Lecture Notes in Artificial Intelligence*, Vol. 1861, pp. 792-806, 2000.

GABBAY, D. & GUENTHNER, F. (eds.) *Handbook of Philosophical Logic*, vol. 4, Netherlands: Springer Netherlands, 2002.

———, HOBBER, C., ROBINSON, J. (eds.) *Handbook of Logic in Artificial Intelligence and Logic Programming. Nonmonotonic Reasoning and Uncertain Reasoning*, vol. 3. Oxford: Oxford University Press, 1994.

———; P. Smets (Eds), *Handbook of Defeasible Reasoning and Uncertainty Management Systems. Algorithms for Uncertainty and Defeasible Reasoning*, vol. 5, Netherlands: Kluwer Academic Publisher, 2000.

GARCÍA, A.J. *La Programación en Lógica Rebatible: su definición teórica y computacional*. Tesis de Maestría en Ciencias de la Computación: Universidad Nacional del Sur, Bahía Blanca, Argentina, 1997.

——— *Programación en Lógica Rebatible: lenguaje, semántica operacional y paralelismo.* Tesis de Doctor en Ciencias de la Computación: Universidad Nacional del Sur, Bahía Blanca, Argentina, 2000.

——— & Simari, G. R. Defeasible logic programming: An argumentative approach. *Journal of Theory and Practice of Logic Programming*, 4 (1-2), pp. 95-138, 2004.

HEMPEL, C.G. *La explicación científica. Estudios sobre la filosofía de la ciencia*. Barcelona: Paidos, 1995.

HORTY, J. F. Some Direct Theories of Nonmonotonic Inheritance. In: Gabbay, D., Hobber, C., Robinson, J. (eds.) (1994), pp. 111–187.

——— Argument construction and reinstatement in logics for defeasible reasoning. *Artificial Intelligence and Law*, 9(1), pp. 1-28, 2001.

——— *Reasons as Defaults*. Oxford University Press, 2012.

———; Thomason, R.H., & Touretzky, D.S. A skeptical theory of inheritance in nonmonotonic semantics networks. *Artificial intelligence*, *42*(2), pp. 311-348, 1990.

KOWALSKI, R. & TONI, F. Abstract argumentation. *Artificial intelligence and law*, 4(3-4), pp. 275-296, 1996.

LOUI R. & STIEFVATER K, *Corrigenda to Poole's rules and a lemma of Simari-Loui*, Department of Computer Science, Washington University, 1992.

——— Defeat among arguments: a system of defeasible inference. *Computational intelligence*, *3*(1), pp. 100-106, 1987.

MILONE, R. A. y TORRES, J.M. *Ciencia y Metafísica: Selección de Trabajos*, Mendoza: FFL-UNCuyo. 2012.

POOLE, D. On the Comparison of Theories: Preferring the Most Specific Explanation. En *Proceedings of the 9th International Joint Conference on Artificial Intelligence,* pp. 144-147, 1985.

POLLOCK, J. L. Defeasible Reasoning. *Cognitive Science,* 11*,* pp. 481-418, 1987.

——— *Cognitive carpentry: A blueprint for how to build a person.* Massachusetts: MIT Press, 1995.

PRAKKEN, H. Intuitions and the Modelling of Defeasible Reasoning: Some Case Studies. In *Proc.* of *9th International Workshop on Non-Monotonic Reasoning,* pp. 91-102, 2002.

——— & SARTOR, G. A system for defeasible argumentation, with defeasible priorities. In *Proc.* of *Int. Conf. on Formal and Applied Practical Reasoning*, pp. 510-524, 1996.

——— & VREESWIJK, G.A.W. Logics for defeasible argumentation. En Gabbay, D. & Guenthner, F. (eds.) (2002), pp. 219-318.

RAHWAN, I. & SIMARI, G. R. (Eds.) *Argumentation in artificial intelligence,* Dordrecht-New York: Springer, 2009.

REITER, R. A Logic for Default Reasoning. *Artificial Intelligence, 13,* pp. 81-132, 1980.

———; Criscuolo, G. On Interacting of Default. *Proc. Int. Joint Conference on Artificial Intelligence,* pp. 270-276, 1981.

SIMARI, G.R. & LOUI, R.P. A Mathematical Treatment of Defeasible Reasoning and Its Implementation. *Artificial intelligence*, 53(2), pp. 125-157, 1992.

TOURETZKY, D.S., THOMASON, R.H., & HORTY, J.F. A skeptic's menagerie: Conflictors, Preemptors, Reinstaters, and Zombies in Nonmonotonic Inheritance. In *Proc. of the 12th Int. Joint Conference on Artificial Intelligence,* pp. 478-485, 1991.