II ENCUENTRO DEJÓVENES INVESTIGADORES

"Consolidando espacios del quehacer científico en San Juan"

2, 3 y 4 de Octubre 2013

UN SISTEMA ARGUMENTATIVO PARA RAZONAMIENTO DEFAULT: DISCUSIÓN Y PROPUESTA

Filosofía y discurso

Claudio A. Alessio

CONICET – UCCuyo

1 INTRODUCCIÓN

Horty en (2) ha propuesto una serie de ejemplos que parecen poner en duda la capacidad de los sistemas argumentativos para modelar adecuadamente razonamiento default. Aunque la crítica realizada por Horty es dirigida al sistema propuesto en (7), también puede aplicarse a otros sistemas similares como por ejemplo el definido en (9). En diversos trabajos (6; 8) se ha discutido y evaluado las razones que pueden llevar a tales sistemas a obtener los resultados intuitivamente inadecuados. Al parecer, la razón de ello se debe al empleo de un lenguaje formal demasiado simple y de un mecanismo de selección de argumentos basado en especificidad que hacen que el restablecimiento y el cálculo de aceptabilidad obtenga resultados problemáticos o incorrectos. El objetivo de este trabajo consiste en presentar la idea de que si la selección de los argumentos es definida de una manera alternativa, los resultados adecuados pueden ser obtenidos. En particular tal resultado es aplicado al sistema propuesto en (9).

2 SISTEMAS ARGUMENTATIVOS

Los sistemas argumentativos son una forma de modelar razonamiento revisable mediante la construcción y comparación de argumentos por y en contra de ciertas conclusiones. Los mismos pueden ser caracterizados en el siguiente esquema de tres pasos.

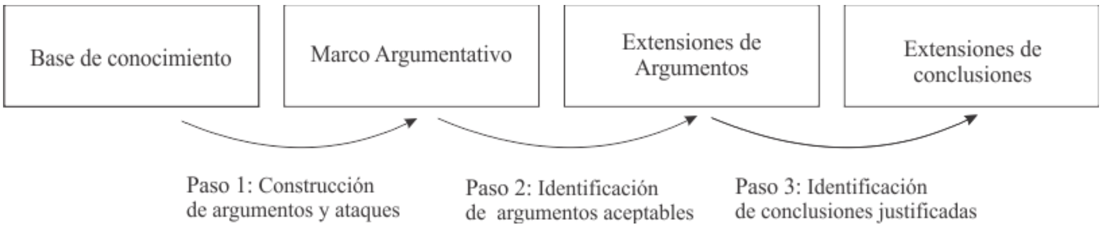

Fig. 1: Sistemas Argumentativos

A partir de una base de conocimientos dada, se genera un conjunto de argumentos y se determina en que forma estos argumentos se atacan (paso 1). El resultado es un marco argumentativo. Basado en este marco, el siguiente paso es determinar el conjunto de los argumentos que pueden ser aceptados, usando un criterio pre definido denominado semántica (paso 2). Luego de que el conjunto de argumentos aceptados es seleccionado, el conjunto de conclusiones justificadas es identificada (paso 3), las conclusiones de los argumentos aceptables.

Aunque los diferentes sistemas argumentativos pueden variar en el lenguaje formal y en consecuencia en la manera de modelar argumentos, como también en la forma en que se define las relaciones de ataque entre argumentos y en el tipo de semántica (crédula o escéptica); hay una noción fundamental que subyace a todos ellos, la noción de restablecimiento y la manera de construir y evaluar los argumentos.

El restablecimiento permite establecer que un argumento derrotado puede considerarse aceptable, i.e. justifica su conclusión, si todos sus derrotadores son a su vez derrotados por argumentos aceptables. El siguiente ejemplo ilustra la idea:

Ejemplo 1

A. Tweety vuela porque es un ave.
B. Tweety es un pingüino dado que esto ha sido observado, luego Tweety no vuela.
C. La observación de que Tweety es un pingüino no es confiable dado que fue hecha durante una tormenta de nieve.

En el marco definido por estos argumentos, parece tener sentido el hecho de que Tweety vuele, y en consecuencia, considerar al argumento A como aceptable dado que el argumento que podría llevar a su rechazo (B) ha sido derrotado por C.

Por otro lado, el restablecimiento permite que se separen las fases de construcción de argumentos y la evaluación final de los mismos como se observa en el mismo ejemplo. En (2) se ha presentado una serie de ejemplos que parecen poner en duda al restablecimiento y la consecuente separación de las fases de construcción. A continuación se presenta tal planteo.

3 RESTABLECIMIENTO Y CONSTRUCCIÓN DE ARGUMENTOS

En (2) se presentan una serie de ejemplos que ponen en duda dos mecanismos empleados en sistemas argumentativos: restablecimiento y la separación de las fases de construcción y evaluación de argumentos. Las cítricas de Horty son dirigidas a dos sistemas concretos: el sistema PS (7) y el sistema KT (3). Sin embargo, es posible extender las mismas críticas al sistema propuesto en (9), en adelante SL.

A continuación se presentan unos ejemplos similares a los propuestos por Horty. Luego de cada ejemplo se presenta una breve discusión sobre la aceptabilidad de ciertos argumentos desde un punto de vista intuitivo y se señala las respuestas que darían los PS, KT y SL.

Ejemplo 2

A. Tweety vuela porque es un ave, y las aves por lo general vuelan.
B. Tweety no vuela porque es un pingüino y los pingüinos no vuelan.
C. Tweety vuela porque es un pingüino mágico y los pingüinos mágicos por lo general vuelan.

Si A y C son conjuntamente admitidos, entonces una conclusión correcta (Tweety vuela) puede estar justificada en la base de razones incorrectas (vuela porque es ave). Claramente, la razón por la que Tweety vuela no es debida a la habilidad que usualmente las aves poseen sino a una habilidad específica que los pingüinos mágicos tienen. La aceptación de A puede ser incorrecta si el modelo es entendido en tanto que los argumentos ofrecen buenas explicaciones de lo que sostiene. Esto podría ser aceptable sólo si la subclase de aves (incluyendo a los pingüinos y los pingüinos mágicos) tuviera la habilidad de volar, pero aquí no es el caso. En los sistemas SL y PS es posible obtener los resultados anteriores.

Una situación similar, aunque más grave, aparece si la conclusión de A es más fuerte que la de C[1]:

Ejemplo 3

A. John es millonario porque es empleado de Microsoft y ellos tienden a ser millonarios.
B. John tiene menos de medio millón porque él es un nuevo empleado de Microsoft y ellos tienden a tener menos de medio millón.
C. John tiene al menos medio millón porque es un nuevo empleado de Microsoft en el departamento X y ellos tienden a tener al menos medio millón.

Aquí, la conclusión de A es más fuerte que la conclusión de C. El argumento más general A sustenta una conclusión más específica que la del argumento más

[1] Es más fuerte en el sentido que la conclusión de A implica la de C.

específico C que lo restablece, lo cual no es intuitivo. Desde un punto de vista práctico, no parece razonable defender el argumento A con C. En los sistemas KT, SL y PS es posible obtener los resultados anteriores.

Horty señala que el restablecimiento hace que los sistemas obtengan los resultados no intuitivos. La solución, no emplearlo. Sin embargo, Prakken, en (6) sostiene que los resultados extraños que aparecen en el sistema PS ocurren debido a que el lenguaje no es lo suficientemente expresivo como para bloquear las reglas rebatibles que no deben ser aplicadas.

Además de los ejemplos anteriores, Horty propone uno similar al siguiente para señalar que el proceso de evaluación de argumentos separado del de construcción, permite obtener un resultado inadecuado.

Ejemplo 4

A. Ana es adinerada porque es una abogada.
B. Ana no es adinerada porque es una defensora pública.
C. Ana es adinerada porque vive en Brentwood.
D. Ana no es adinerada porque renta en Brentwood.

Los argumentos cobran sentido si se tiene en cuenta la siguiente información rebatible: los abogados por lo general no son adinerados, los defensores públicos por lo general no son adinerados; por lo general los que viven en Brentwood son adinerados; y quienes rentan en Brentwood usualmente no son adinerados. Además se debe tener en cuenta de la siguiente información estricta obvia: Todos los defensores públicos son abogados. Todos los que rentan en Brentwood viven en Brentwood.

Frente a este ejemplo, el sistema PS no obtiene ninguna conclusión al respecto. Lo mismo en SL. Mientras que la semántica crédula de KT obtiene ambas y la versión escéptica ninguna.

Ahora bien, tal como se dijo más arriba, parecería que el responsable del problema se debe a la poca expresividad del lenguaje, sin embargo, es posible hacer que, por ejemplo, el sistema SL funcione adecuadamente conservando el mismo lenguaje de representación. Al parecer la aparición de los resultados no intuitivos se debe más bien a la manera en cómo se definen las relaciones entre argumentos que el efecto de la expresividad del lenguaje. La razón que motiva tal planteo es la siguiente.

En la famosa guía de falacias de Stephen Downes, se señala una que merece la atención en lo que respecta a la temática del presente artículo. La falacia es denominada falacia de la exclusión y consiste en la presentación de un argumento no deductivo que omite evidencia que puede socavarlo. El requerimiento de que toda la información relevante sea incluida es denominado principio de evidencia total. De modo que esta falacia consiste en un atentado contra tal principio. Dowens propone los siguientes ejemplos que ilustran la falacia:

Ejemplo 5: Jones es de Alberta, y la mayoría de los habitantes de Alberta votarán al Tory, de modo que Jones probablemente vote al Tory. La información excluida en el argumento es que Jones vive en Edmonton, de modo que es verdad que Jones es de Alberta, pero tal información llevaría a aceptar una conclusión distinta, dado que la mayoría de las personas de Edmonton votarán al partido liberal o N.D.P.

Ejemplo 6: Los Leafs probablemente ganarán este juego porque ellos han ganado nueve de los diez que han jugado últimamente. (Sin embargo, ocho de los juegos ganados han sido con equipos de los últimos lugares, ahora jugará con el que va en primer lugar.)

La prueba de la falacia consiste en señalar la evidencia excluida y mostrar que esta cambia la conclusión del argumento inductivo. Nótese el parecido con el clásico ejemplo de Tweety:

Ejemplo 7:Tweety es un ave y dado que por lo general las aves vuelan, es razonable concluir que Tweety vuela. Sin embargo se excluye la evidencia de que Tweety es un pingüino. Dado que al ser pingüino y que estas aves no poseen la habilidad de volar, el argumento para “Tweety vuela” debe ser rechazado.

Pero como se señaló más arriba, en (9) por ejemplo, el mecanismo de selección de argumentos permite el restablecimiento de un argumento falaz cuando se emplea el recurso del restablecimiento (Ver ejemplo 2-4). Sin embargo, la razón del comportamiento anómalo parecería no deberse al restablecimiento en sí, sino a la manera en que se define la relación de derrota en (9).

Pollock siempre ha señalado que la derrota por especificidad no es un principio general del razonamiento de sentido común. De hecho, en (5) responde explícitamente a la pregunta de si existe otro tipo de derrota a las de rebatimiento y socavamiento por él propuesta. La respuesta es sencillamente negativa. Pollock señala que la derrota por especificidad es un caso de lo que denomina *subpropertydefeaters* del silogismo estadístico. Esto es así porque, tanto el silogismo estadístico como el razonamiento default, son razonamientos que se deben ajustar al requerimiento de evidencia total. Según él, no existe ningún ejemplo de derrota por especificidad que no sea un caso de requerimiento de evidencia total en alguna de las variedades especiales de razonamiento rebatible y que en última instancia, son todos casos de derrota por socavamiento.

Ahora bien, qué es una derrota por socavamiento: Si las premisas de un argumento R son las razones para creer en una conclusión *h*, puede entenderse el socavamiento para R como la posesión de información que muestra que una o más razones para creer *h* se tornan irrelevantes, i.e. estas en realidad no constituyen una razón para la conclusión. Un ejemplo clásico es el siguiente: Supóngase el default, si algo se ve rojo en general es rojo. Existen varias razones que podrían cancelar este default. Por ejemplo, si se descubre que el observador del objeto es daltónico, o se sabe que el objeto se encuentra iluminado por una luz roja. En ambos casos, es claro que la conclusión, el objeto es rojo, no se encuentra garantizada; la "razón" que motivaba la conclusión –el objeto se ve rojo – resultó no ser una razón a la luz de los "nuevos" datos. Claramente hay una alta relación entre la falacia de la exclusión y la derrota por socavamiento.

Atendiendo a la falacia de la exclusión y a lo señalado por Pollock, el ejemplo de Tweety puede entenderse como socavamiento de la siguiente manera. El default "por lo general las aves vuelan", y el dato "Tweety es un ave" es una razón para creer que Tweety vuela, sin embargo, la información de que Tweety es un pingüino cancela el argumento, dado que la razón (las aves vuelan yTweety es un ave) que motivaba la conclusión -"Tweety vuela"- resultó no ser una razón a la luz de "Tweety es un pingüino". De modo que el argumento "Tweety vuela por ser un ave" es cancelado por el dato "Tweety es un pingüino". El argumento original parece ser un caso de falacia de exclusión dado que omite evidencia que puede socavarlo. Si se sabe que Tweety es pingüino no hay razón para creer que vuela por ser un ave. A su vez, es conveniente destacar el hecho de que si es pingüino ello no quiere decir que sea falso que Tweety sea ave o que sea falso de que vuele. Sólo señala que la razón no es adecuada para la conclusión si se tiene en cuenta tal información.

Continuando con el ejemplo 2, el argumento “Tweety no vuela porque es un pingüino”, parecería justificado. Sin embargo, a la luz de la información, “Tweety es un pingüino mágico”, tal argumento se encuentra cancelado, sin embargo, ese dato no invalida la cancelación del argumento “Tweety vuela por ser ave”, puesto que tal información socava el argumento que emplea la misma información pero no socava al socavador. Pensar el ejemplo 1 en estos términos permite señalar que no hay problema con el restablecimiento, sino que el origen del problema se debe más bien a como se definen las relaciones de derrota.

A continuación se propone un sistema argumentativo basado en SL donde el mecanismo de derrota ha sido modificado según lo anteriormente discutido.

4 PRELIMINARES

La propuesta consiste en construir un sistema formal capaz de modelar razonamiento default. El lenguaje está compuesto por un lenguaje de primer orden *L* al que se le agrega una relación binaria metalingüística “$\Rightarrow$”. Las reglas de inferencia asociadas a *L* son modus ponens y generalización.

Los miembros de la relación “$\Rightarrow$” se denominarán reglas rebatibles y tendrán la forma: $A_1,\dots, A_n \Rightarrow B \quad n\geq 1$.Donde $A_1,\dots, A_n$son literales de *L*. “$A_1,\dots, A_n \Rightarrow B$” expresa la siguiente idea: “las razones para creer en $A_1,\dots, A_n$proveen razones para creer en B” o simplemente: “$A_1,\dots, A_n$son razones para creer B”. Los literales del antecedente de la regla se encuentran en conjunción. Todo nombre de variable que aparece en ambos lados de la regla es la misma variable. Una instancia de una regla rebatible se obtiene por sustitución uniforme de cada variable por una constante de *L*.

El conjunto Sent(L) de sentencias de *L* es el conjunto de fórmulas bien formadas cerrado en *L*. Este conjunto puede ser particionado en dos subconjuntos SentN(L) y SentC(L), correspondiendo a la información *necesaria* y *contingente*. La información necesaria simplemente está constituida por sentencias de variables libres o implicaciones en *L*, mientras que la información contingente está constituida por literales instanciados:$Sent(L) = Sent_C(L) \cup Sent_N(L)$. Obviamente:$Sent_C(L) \cap Sent_N(L) = \varnothing$

El conocimiento de un agente *a* es representado por un par (K, Δ), donde es K un subconjunto de *Sent*(*L*) y Δ es un conjunto finito de reglas rebatibles. El par (K, Δ) será llamado una estructura lógica rebatible. Krepresenta la parte no rebatible del conocimiento de *a* mientras que Δ representa la información tentativa (K será denominado el contexto). Kpuede ser particionado en dos subconjuntos:

$$K_N = Sent_N(L) \cap K, \qquad K_C = Sent_C(L) \cap K$$

Claramente $K = K_N \cup K_C$. La única condición en K es la consistencia, i.e., $K \nvdash \perp$.

Dado un miembro *A* de *Sent*(*L*) y un conjunto $\Gamma=\{A_1, A_2, \ldots, A_n\}$, donde cada $A_i$es miembro de K o una instancia de Δ, se establecerá una relación meta-metalingüística $\mid\!\sim$, llamada consecuencia rebatible, entre Γ y *A* de la siguiente forma: Una fbf*A* será llamada una *consecuencia rebatible* de *Γ*sii existe una secuencia $B_1, \ldots, B_m$ tal que $A = B_m$ y para cada *i*, o bien B_i es un axioma de *L*, o B_i es en Γ, o B_i es una consecuencia directa de los miembros precedentes de la secuencia usando modus ponens o instanciación de una sentencia cuantificada universalmente. Se usará $\Gamma \mid\!\sim A$ como una abreviación de "*A* es una consecuencia rebatible de Γ".

5 ARGUMENTOS

Con vistas a facilitar las siguientes definiciones se introducirá el conjunto $\Delta^{\downarrow}$ consistente de todas las instancias de los miembros de Δ producidos por el uso de constantes individuales en *L*.

Definición 1. Dado una teoría rebatible (K, Δ), un subconjunto T de $\Delta^{\downarrow}$será denominado un *argumento* para $h \in Sent_C(L)$ en el contexto K, denotado por $\langle T, h \rangle$ si y sólo si:

i. $K \cup T \mid\!\sim h$,

ii. $K \cup T \not\mid\!\sim \perp$,

iii. $\nexists T' \subset T, K \cup T' \mid\!\sim h$.

Ejemplo 8.

Sea $K = \{P(a), Q(a)\}$ y $\Delta=\{P(x)\Rightarrow R(x), Q(x)\wedge R(x)\Rightarrow H(x), M(x)\Rightarrow N(x)\}$ el contexto y el conjunto de reglas rebatibles respectivamente. Por lo tanto, el subconjunto T de instancias de Δ

$$T=\{P(a)\Rightarrow R(a), Q(a)\wedge R(a)\Rightarrow H(a)\}$$

es una estructura de argumento para H(a), i.e. $\langle T, H(a)\rangle$.

Definición 2. Se dice que $\langle S, j\rangle$ es un *subargumento*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 $\langle S, j\rangle$ una estructura de argumento para j tal que $S \subseteq T$. Si $S \subset T$ se dirá que S es un *subargumento propio* de T.

Ejemplo 9. Dado que $T = \langle\{P(a)\Rightarrow R(a), Q(a)\wedge R(a)\Rightarrow H(a)\}, H(a)\rangle$ es un argumento para H(a) y $S = \langle\{P(a)\Rightarrow R(a)\}, R(a)\rangle$ es un argumento para R(a) y se verifica la propiedad de que $\{P(a)\Rightarrow R(a)\}\subseteq\{P(a)\Rightarrow R(a), Q(a)\wedge R(a)\Rightarrow H(a)\}$ entonces S es un subargumento de T.

6 INTERACCIONES ENTRE ARGUMENTOS

Diversos tipos de interacciones entre argumentos pueden ser definidos. Una de ellas es la de *subargumentación*. Sin embargo es posible definir relaciones de *concordancia*, i.e. cuando dos argumentos sustentan la misma conclusión; *conflicto* o *desacuerdo*, cuando ambos argumentos no pueden ser simultáneamente aceptados por el agente; *consistencia*, i.e. cuando los argumentos no son mutuamente excluyentes y *contraargumentación* que también supone la de desacuerdo. Algunas de estas serán definidas a continuación.

Definición 4. Dos estructuras de argumento, $\langle T_1, h_1\rangle$ y $\langle T_2, h_2\rangle$, están en *desacuerdo*si y sólo si: $K\cup\{h_1, h_2\}\vdash\perp$.

Ejemplo 10. Dados los siguientes argumentos $T = \langle\{A\Rightarrow B\}, B\rangle$ y $S = \langle\{C\Rightarrow\neg B\}, \neg B\rangle$ y $K = \{A, C\}$, se dice que T y S están en desacuerdo dado que $K\cup\{Con(T), Con(S)\}\vdash\perp$

En el siguiente ejemplo puede observarse un caso en que no se da la relación de desacuerdo aunque no pueden aceptarse conjuntamente:

Ejemplo 11. T = $\langle\{A\Rightarrow\neg B\}, \neg B\rangle$ y S = $\langle\{C\Rightarrow B, B\Rightarrow D\}, D\rangle$ y K = $\{A, C\}$

Tal ejemplo motiva la siguiente definición.

Definición 5. Se dice que un argumento $\langle T_1, h_1\rangle$*contraargumenta* a $\langle T_2, h_2\rangle$ si y sólo si existe un subargumento$\langle T, h\rangle$ de $\langle T_2, h_2\rangle$ tal que $\langle T_1, h_1\rangle$ y $\langle T, h\rangle$ están en desacuerdo, i.e. $K\cup\{h_1, h\}\vdash\perp$.

En el ejemplo11, T contraargumenta a S.

Las siguientes definiciones pretende capturar la idea de socavamiento anteriormente discutida.

Definición 6: Sea D = $\{a\in L$: es un literal básico y $K \cup\Delta^{\downarrow} \mid\!\sim a\}$. Sea $\langle T_1,h_1\rangle$ un argumento. Sea $S_1\subseteq$ D. Se dice que S_1 es un *activador no trivial* de $\langle T_1,h_1\rangle$sii

1. $K_g\cup S_1\cup T_1 \mid\!\sim h_1$; y
2. $K_g\cup S_1 \not\mid\!\sim h_1$

Definición 7: Sea D = $\{a\in L$: es un literal básico y $K \cup\Delta^{\downarrow} \mid\!\sim a\}$. Sea $\langle T_1,h_1\rangle$ un argumento. Sea $S \subseteq$ D. Se dice que Ses el conjunto de *todos los activadores no triviales* de $\langle T_1,h_1\rangle$sii

1. $\forall S_i\in S, K_g\cup S_i\cup T_1 \mid\!\sim h_1$; y
2. $K_g\cup S_i \not\mid\!\sim h_1$

Definición 8: Sean $\langle T_1,h_1\rangle$ y $\langle T_2,h_2\rangle$ argumentos para h_1 y h_2 respectivamente. Sea S el conjunto de todos los activadores para $\langle T_2,h_2\rangle$. Se dice que S es un socavador para $\langle T_1,h_1\rangle$, notado como $\langle S, \neg \langle T_1,h_1\rangle\rangle$sii

i. $\langle T_2,h_2\rangle$contraargumenta a $\langle T_1,h_1\rangle$ en h ($\langle T,h\rangle\sqsubseteq \langle T_1,h_1\rangle$); y
ii. $\forall S_i\in S$, $K_g\cup S_i\cup T_1 \mid\!\sim h_1$

Con esta definición en mente, es posible tratar los ejemplos problemáticos adecuadamente. Retomando el ejemplo 2:

$\langle T_1,h_1\rangle$ es el argumento “Tweety vuela por ser ave”. $\langle T_2,h_2\rangle$ el argumento para “Tweety no vuela por ser pingüino”. Claramente $\langle T_2,h_2\rangle$ es un argumento que contraargumenta a $\langle T_1,h_1\rangle$ y el dato Tweety es un pingüino es un activador del argumento $\langle T_2,h_2\rangle$ que también activa a $\langle T_1,h_1\rangle$ por el hecho de ser más específico. De modo que Tweety es pingüino es un socavador para $\langle T_1,h_1\rangle$. La situación es análoga cuando se considera a $\langle T_3,h_3\rangle$ -el argumento "Tweety vuela por ser un pingüino mágico"- en relación a $\langle T_2,h_2\rangle$. El conjunto de activadores de $\langle T_3,h_3\rangle$ está constituido por el dato: Tweety es un pingüino mágico. De modo que el dato: "Tweety es un pingüino mágico" es un activador de un argumento estrictamente más específico que contraargumenta a $\langle T_2,h_2\rangle$ entonces, "Tweety es un pingüino mágico" socava a $\langle T_2,h_2\rangle$.

Con este ejemplo, se puede visualizar que la inclusión de ciertos datos invalida la aplicación de ciertas inferencias, al estilo de la prueba contra la falacia de la exclusión. A su vez, también es posible demostrar que la solución es lo suficientemente general como para capturar la noción de derrota por especificidad dada en Simari y Loui, 1992 mediante el siguiente teorema

Teorema 1: Siempre que $\langle T_1,h_1\rangle$ sea un derrotador propio de $\langle T_2,h_2\rangle$, i.e. $\langle T_1,h_1\rangle >_{esp} \langle T_2,h_2\rangle$ y $\langle T_1,h_1\rangle \bowtie \langle T, h\rangle$, donde $\langle T, h\rangle \sqsubseteq \langle T_2,h_2\rangle$, en SL, entonces existe un socavador$\langle S, \neg\langle T_2,h_2\rangle\rangle$ tal que $S \in F$ donde F es el conjunto de activadores no triviales de $\langle T_1,h_1\rangle$.

Prueba: Dado que $\langle T_1,h_1\rangle$ es un derrotador propio de $\langle T_2,h_2\rangle$ es claro que $\langle T_1,h_1\rangle$contraargumenta$\langle T_2,h_2\rangle$ en h. Dado que $\langle T_1,h_1\rangle$ es estrictamente más específico que $\langle T,h\rangle$ es claro que todo lo que activa a $\langle T_1,h_1\rangle$ activa a $\langle T,h\rangle$. De modo que si para cualquier activador no trivial de $\langle T_1,h_1\rangle$ se verifica que tal activador es un activador de $\langle T,h\rangle$.

7 CONCLUSIÓN

En el presente trabajo se ha definido un mecanismo formal para sistemas argumentativosque modelan razonamiento default capaz de evitar el comportamiento anómalo que sistemas similares padecen, en particular en SL. El resultado se obtiene mediante la construcción de argumentos que han sido denominados socavadores. Estos cumplen la función de cancelar la inferencia de ciertos argumentos si es incluida información que puede activar un argumento estrictamente más específico que a la vez es un contraargumento. Se ha presentado un teorema que permite relacionar el mecanismo formal con el sistema SL. Resta poder definir las relaciones con el sistema PS y KT.

8 BIBLIOGRAFÍA

1) Dowens. S. 1996: *Guide to the Logical Fallacies, sd.*
2) Horty, J.: 2001, 'Argument Construction and Reinstatement in Logics For Defeasible Reasoning', *Artificial Intelligence and Law,* vol. 9, 1-28.
3) Kowalski, R& Toni, F. 'Abstract argumentation'. *Artificial Intelligence andLaw* 4:275–296, 1996.
4) Pollock, J.L.: Defeasible Reasoning. *Cognitive Science* (COGSCI) 11(4):481-518 (1987).
5) Pollock, J.L.: Defeasible reasoning with variable degrees of justification. *Artificial Intelligence* 133(1-2):233-282 (2001).
6) Prakken, H.: 2002, 'Intuitions and the Modelling of Defeasible Reasoning: Some Case Studies'. In Benferhat, S. & E. Giunchiglia (Eds.) *Proc. of the 9th International Workshop on Non-Monotonic Reasoning*, NMR 2002, Tolouse, France.
7) Prakken, H. and Sartor, G.: 1997, 'Argument-based Extended Logic Programming with Defeasible Priorities', *Journal of Applied Non-clasical Logics*, vol. 7, 25-75.
8) Prakken, H. and Vreeswikj, G.: 2002, 'Logics for Defeasible Argumentation'. In D. Gabbay and F. Guenthner (eds.), *Handbook of Philosophical Logic*, second edition, Vol 4, pp. 219-318. Kluwer Academic Publishers, Dordrecht etc., 2002.
9) Simari, G. and Loui, R.: 1992, 'A Mathematical Treatment of Defeasible Reasoning', *Artificial Intelligence*, vol. 53, pp. 125–157.