Propiedades del Operador de Consecuencia Argumentativo

SEBASTIAN SARDIÑA* GUILLERMO R. SIMARI
Departamento de Ciencias de la Computación
Universidad Nacional del Sur
Av. Alem 1253 - (8000) Bahía Blanca
REPUBLICA ARGENTINA
e-mail: {ssardina,grs}@criba.edu.ar

PALABRAS CLAVES
Razonamiento No-monótono, Sistemas Argumentativos,
Operador de Consecuencia y Propiedades, Razonamiento Coherente

Resumen

Desde comienzos de la década de los '80 han aparecido un número significativo de formalismos que modelan, en ciertos aspéctos, el razonamiento de sentido común. El esfuerzo fué orientado a definir una clase de razonamiento que admitiese información incompleta y potencialmente inconsistente, fundamentalmente motivado por la intuición de que esta parece ser una característica distintiva de las decisiones tomadas por un agente inteligente. La plétora de sistemas ha generado un interés creciente por lograr criterios unificadores y hallar patrones comunes entre todos ellos que puedan ayudar a comprender mejor la situación. Existen dos niveles donde estos patrones pueden buscarse: a nivel de la máquina formal usada para razonar y a nivel de la inferencia en sí misma.

Gabbay [Gab85] propuso concentrarse en el estudio a las propiedades formales de la relación de inferencia del sistema. Posteriormente, Makinson [Mak90] en un trabajo de gran trascendencia realizó un estudio sobre varios sistemas, centrándose en dos propiedades importantes llamadas propiedades cumulativas.

En [Sar98] se extiende este estudio a los sistemas argumentativos, en particular al sistema MTDR desarrollado por Simari y Loui [Sim92]. En aquel trabajo se desarrollan consideraciones con respecto a las propiedades cumulativas. Vreeswijk en [Vre92] realiza un estudio similar.

En este artículo se prolonga el estudio realizado en [Sar98] explorando ahora algunas propiedades importantes que hablan de la "coherencia lógica" de un razonamiento. Condiciones tales como *supraclasicalidad, absorción, equivalencia lógica a izquierda* y *debilitamiento a derecha*, necesarias en un "buen razonamiento", evalúan la coherencia de un razonamiento. Terminaremos analizando tres propiedades no-Horn que no se verifican en casi ningún sistema no-monótono.

*Becario de la UNS: *Beca de Iniciación a la Investigación para Graduados de la UNS.*

Propiedades del Operador de Consecuencia Argumentativo

PALABRAS CLAVES
Razonamiento No-monótono, Sistemas Argumentativos
Operador de Consecuencia y Propiedades, Razonamiento Coherente

1 Introducción: Historia y Motivaciones

En el campo del razonamiento no-monótono han proliferado una multitud de acercamientos, la mayoría desde puntos intuitivos diferentes y muchas veces difíciles de reconciliar. Se ha enfocado el problema desde distintos ángulos, obteniendo una serie de sistemas que parecieran ser *muy diferentes* tanto estructuralmente como también intuitivamente. Por tal motivo, ha surgido un gran interés en buscar un conjunto de patrones comunes o mínimos que un "buen" sistema debiera satisfacer. Por patrones comunes entendemos la satisfacción o no de propiedades sobre el operador de consecuencia asociado a un sistema. El objetivo es poder entender que es lo que está sucediendo en tal caótico campo, como también poseer un método para determinar la "clase" de razonamiento que un sistema dado modela.

Gabbay en [Gab85] es el primero en incursionar en dicha búsqueda. Kraus, Magidor y Lehmman realizan en [Kra90] un interesante estudio de algunas propiedades, centradas en sus modelos preferenciales. Luego vendría un muy prolijo estudio de Makinson [Mak89] y [Mak90], estudiando una serie de formalismos a la luz de varias propiedades interesantes. En este último trabajo, se resaltaron las propiedades cumulativas como posibles condiciones necesarias. Finalmente, en [Sar98] y [Sar97a] se extiende dicho estudio a los sistemas argumentativos y se debate la necesidad de las propiedades cumulativas.

¿Quedan otras propiedades importantes, además de las cumulativas, para analizar?. Las propiedades cumulativas *cut* y *monotonía cauta* nos indican el comportamiento de un sistema, ante la acumulación de una conclusión al conjunto de premisas iniciales. Nada nos dicen sobre la "coherencia lógica" del razonamiento modelado. Por ejemplo, si un agente concluyera que *si Sofía viene a la fiesta, entonces la fiesta estará divertida*, y además concluyera efectivamente que *Sofía vendrá a la fiesta*, ¿no sería **coherente** que infiriera que *la fiesta estará divertida*?. El presente trabajo intenta centrar la atención en estas cuestiones: **analizar lo que se entiende por un razonamiento "lógicamente coherente"** y, evaluar la coherencia del acercamiento argumentativo.

2 Buscando un Razonamiento Coherente

Como hemos dicho, intentamos buscar qué se entiende por un razonamiento lógico y qué tan lógico se comporta el razonamiento argumentativo. Recordemos que, nuestra estrategia se basa en la aplicación de propiedades al operador de consecuencia de un sistema lógico, el cual expresa todo lo inferido a partir de un conjunto inicial de premisas. Esta estrategia, intenta categorizar un razonamiento a partir de la "salida" de su máquina de inferencia. En esta sección, citaremos un conjunto de condiciones posibles, divididas en dos grupos según la forma en que estas propiedades se expresan.

2.1 Propiedades Horn

Las propiedades Horn tienen la forma: "*de la presencia de ciertas inferencias se asegura la presencia de algunas otras inferencias*". Para muchas de estas propiedades se hace referencia al operador clásico de consecuencia Cn. La primera propiedad que citaremos es tan simple como obvia. Nuestro razonamiento siempre debe ser coherente con las premisas iniciales. Como sabemos, el razonamiento no-monótono intenta obtener conclusiones plausibles, tentativas. Lo que sí parece ser indiscutible, es que todas las consecuencias clásicas de nuestro conocimiento inicial, que es el conocimiento totalmente seguro, deben formar parte de nuestras conclusiones. Si un agente $\mathcal{A}$ conoce[1] inicialmente, que *todos los hombres son mortales* y que *José es un hombre*, debería indiscutiblemente concluir a través de su razonamiento, no sólo dicho conocimiento, sino además que *José es mortal*. Dicho comportamiento lo modela la siguiente condición:

$$Cn(A) \subseteq C(A) \quad \text{(Supraclasicalidad)} \tag{1}$$

Supraclasicalidad parece ser una propiedad universalmente verificada por cualquier tipo de razonamiento basado en alguna noción de consecuencia. Las próximas dos condiciones expresan la influencia de la lógica subyacente de un sistema sobre las conclusiones plausibles del mismo. Para ver la primera propiedad, refirámosnos a nuestro anterior ejemplo: para el agente $\mathcal{A}$, ¿debería ser lo mismo, para su razonamiento, conocer que *José es mortal porque es hombre, y todos los hombres son mortales*, que conocer explícitamente que *Juan es mortal*?. Pareciera que si.

$$Cn(A) = Cn(B) \Rightarrow C(A) = C(B) \quad \text{(Equivalencia Lógica a Izquierda)} \tag{2}$$

Equivalencia Lógica a Izquierda indica la necesidad de que conjuntos lógicamente equivalentes (dado por la igualdad de sus consecuencias clásicas), tengan exáctamente las mismas consecuencias. En otras palabras, las consecuencias de un conjunto de conocimiento inicial debe depender de su significado y no de su forma. La próxima propiedad, debilitamiento a derecha, expresa el hecho de que uno debe aceptar como consecuencia tentativa todo lo que es lógicamente implicado por algo que uno acepta como consecuencia tentativa. Si sabemos que *el hecho que un individuo sea mortal* es una consecuencia lógica de que *dicho individuo es un hombre* y llegamos a la conclusión plausible a través de nuestro razonamiento, que nuestro amigo *José es un hombre*, entonces deberíamos aceptar que efectivamente *José es mortal*. Notemos que, en este caso, **no conocemos inicialmente si José es hombre**, sino que llegamos a dicha creencia a través del razonamiento.

$$x \to y \in Cn(A),\, x \in C(A) \Rightarrow y \in C(A) \quad \text{(Debilitamiento a Derecha)} \tag{3}$$

Intuitivamente, pareciera trival coincidir que las conjunciones de nuestras conclusiones deben ser también conclusiones, como lo expresa la siguiente propiedad:

$$x, y \in C(A) \Rightarrow x \wedge y \in C(A) \quad \text{(Conjunción de Conclusiones)} \tag{4}$$

Una propiedad íntimamente relacionada a la propiedad supraclásica dada en (1), es la propiedad de cumulatividad subclásica, la cual nos indica que el razonamiento se mantiene estable al acumular una conclusión clásica. Nótese que la propiedad cumulativa es mucho más amplia, ya que la acumulación se refiere a cualquier conclusión del sistema, y no únicamente a las inferencias clásicas. Para un estudio de las propiedades cumulativas referirse a [Sar98].

$$A \subseteq B \subseteq Cn(A) \Rightarrow C(A) = C(B) \quad \text{(Cumulatividad Subclásica)} \tag{5}$$

Ahora bien, las propiedades que hemos descripto modelan distintos aspectos "lógicos" que parecen ser necesarios en un razonamiento coherente. Todas ellas pueden enmarcarse como partes de una propiedad mucho más general, ambiciosa y no tan trivial: **absorción total**.

$$C(A) = C(Cn(A)) \quad \text{(Absorción a Derecha)} \tag{6}$$

$$C(A) = Cn(C(A)) \quad \text{(Absorción a Izquierda)} \tag{7}$$

$$C(A) = C(Cn(A)) = Cn(C(A)) \quad \text{(Absorción Total)} \tag{8}$$

[1] Observese la diferencia entre "conocer" y "concluir". El pimero indica que el dato pertenece al conjunto de premisas. El segundo se refiere a que se llega al dato a través del razonamiento y no necesariamente debe pertenecer a las premisas.

Makinson afirma en [Mak89] que "un acercamiento le merece el nombre de *lógico* siempre que conlleve a una operación de inferencia C que verifique el principio de absorción total". Continua diciendo: "en otras palabras, sólo si las proposiciones que están permitidas inferir a partir de un conjunto A, forman una teoría clásica ($C = CnC$), la cual además depende solo del contenido lógico de A más que de la forma en que éste se presenta ($C = CCn$)". Coincidimos plenamente en dicha restricción y deseamos que nuestro sistema argumentativo la verifique.

El impacto de *absorción total* puede verse en la siguiente proposición.

Proposición 2.1: Sea C un operador cualquiera. Si C satisface la absorción a izquierda $C(A) = Cn(C(A))$, entonces:

(a) siempre que $x, y \in C(A)$ entonces $x \wedge y \in C(A)$ (conjunción de conclusiones)

(b) siempre que $x \in C(A)$ y $x \rightarrow y \in Cn(A)$ entonces $y \in C(A)$ (debilitamiento a derecha)

Opuestamente, si C satisface la absorción a derecha $C(A) = C(Cn(A))$ entonces:

(c) siempre que $Cn(A) = Cn(B)$ entonces $C(A) = C(B)$ (equivalencia lógica a izquierda)

(d) siempre que $A \subseteq B \subseteq Cn(A)$ entonces $C(A) = C(B)$ (cumulatividad subclásica)

Prueba.

(a) si $x, y \in C(A)$ entonces $x \wedge y \in Cn(C(A)) = C(A)$.

(b) si $x \in C(A)$ y $x \rightarrow y \in Cn(A)$, por propiedad de consecuencia clásica $y \in Cn(C(A)) = C(A)$, dado que $x \rightarrow y \in C(A)$ por supraclasicalidad.

(c) $C(A) = C(Cn(A)) = C(Cn(B)) = C(B)$.

(d) si $A \subseteq B \subseteq Cn(A)$ entonces $Cn(A) \subseteq Cn(B) \subseteq Cn(Cn(A)) = Cn(A)$ de manera que $Cn(A) = Cn(B)$ y se puede aplicar (3).

■

2.2 Tres Propiedades no-Horn

Existen condiciones diferentes en forma de las que ya hemos enunciado, interesantes tanto desde el punto de vista matemático como intuitivo.

Sintácticamente, estas condiciones pueden ser formuladas con hipótesis negativas o equivalentemente con conclusiones disyuntivas. Las nuevas condiciones serán de la forma: "**de la ausencia de ciertas inferencias en la relación, deducimos la ausencia de algunas otras inferencias**". A estas condiciones se las llaman *condiciones no-Horn*.

La primera condición se la llama **negación racional** (negation rationality):

$$A \not\mid\!\sim z \Rightarrow A \cup \{x\} \not\mid\!\sim z \text{ ó } A \cup \{\neg x\} \not\mid\!\sim z \tag{9}$$

Como afirma Kraus, Lehmann y Magidor en [Kra90] la negación racional dice que las inferencias no están hechas solo a base de la ignorancia. Si aceptamos z como consecuencia plausible a partir de A, debemos aceptar también z como consecuencia o bien de $A \cup \{x\}$ o bien de $A \cup \{\neg x\}$. De hecho, son las únicas dos alternativas que existen. Imaginemos que esperamos que *la fiesta estará divertida*, pero no sabemos si *Juan vendrá o no a ella*. Es obvio que, o bien *la fiesta estará divertida con la presencia de Juan*, o bien, *la fiesta estará divertida si Juan no viene a ella*. De otra manera, si la fiesta no estuviera divertida tanto si Juan viniera como si no lo hiciera, entonces *¿cómo es que deducimos que la fiesta estaría divertida?*.

La segunda condición se la llama **racionalidad disyuntiva** (disjunctive rationality):

$$A \cup \{x \vee y\} \not\mid\!\sim z \Rightarrow A \cup \{x\} \not\mid\!\sim z \text{ ó } A \cup \{y\} \not\mid\!\sim z \tag{10}$$

Volviendo a las ideas intuitivas, esta propiedad afirma que las inferencias hechas a partir de una disyunción de proposiciones deben ser soportadas por al menos una de ellas. Otra vez, pareciera ser un requerimiento razonable. Si no aseguramos que *si Pedro viene a la fiesta, la fiesta estará divertida*, ni aseguramos que *si María viene a la fiesta, la fiesta estará divertida*, entonces, ¿cómo pudimos sostener que *si al menos Pedro o María vienen, la fiesta estará divertida*?.

La tercera condición se la denomina **monotonía racional** (rational monotony):

$$A \mid\!\sim z \Rightarrow A \cup \{x\} \mid\!\sim z \text{ ó } A \mid\!\sim \{\neg x\} \quad (11)$$

La monotonía racional expresa el hecho que sólo información adicional, la negación de lo que se espera, debe forzarnos a eliminar conclusiones plausibles previamente inferidas. Esta es una importante herramienta para minimizar la actualización que un agente debe efectuar al aprender nueva información. Siguiendo con nuestro ejemplo, supongamos que concluímos que *la fiesta estará divertida*, pero no decimos que *aún si Pedro viene, la fiesta estará divertida*, ¿i.e. pensamos que la presencia de Pedro puede arruinar la fiesta. ¿No deberíamos, inicialmente, sostener tentativamente que *Pedro no vendrá a la fiesta*?.

En resumen, estas tres propiedades no-Horn modelan distintos aspectos racionales del razonamiento. Si bien casi ningún acercamiento al razonamiento no monótono las cumple, en [Kra90] se afirma que "cualquier razonador racional debería, en nuestra opinión, soportarlas". Por otra parte, Makinson opina que sería bueno que se cumpla hasta la segunda condición, ya que la última (monotonía racional) es "demasiado pedir". Veremos, en la sección próxima, el comportamiento del sistema argumentativo MTDR ante estas condiciones.

3 Comportamiento Argumentativo

En esta sección estudiaremos el comportamiento del sistema argumentativo MTDR, propuesto por Simari y Loui en [Sim92], sobre las propiedades estudiadas en la sección anterior. Esto nos permitirá ver qué tan "coherente" es el razonamiento argumentativo.

La idea básica de la argumentación es ver el razonamiento como un proceso de construír primeramente argumentos en favor de una conclusión, y luego seleccionar el mejor entre ellos. En otras palabras, una sentencia será inferida si los argumentos que la soportan pueden defenderse exitosamente contra los argumentos que soportan la sentencia opuesta. Daremos por familiares los conceptos de *estructura lógica rebatible, derivación rebatible, argumento, contra-argumento, derrota, especificidad* y *justificación*, los cuales pueden encontrarse definidos en [Sim92] y extendidos en [Gar93] y [Sim94].

3.1 Operador de Consecuencia Argumentativo

Daremos a continuación, la definición del operador de consecuencia asociado al sistema argumentativo MTDR. Esta definición fue dada en [Sar98] y [Sar97a]. Luego, en este mismo trabajo, reformularemos dicha definición en base a las conclusiones arribadas.

Definición 3.1: (**Operador de Consecuencia para el Sistema MTDR**) Sea Δ un conjunto de reglas rebatibles y sea $\mathcal{K}$ un conjunto de fórmulas de primer orden, tal que KB $= (\mathcal{K}, \Delta)$ es una base de conocimiento válida. Definimos el operador de consecuencia argumentativo C_Δ o simplemente C como:

$$C(\mathcal{K}) = Cn(\mathcal{K}) \oplus Justified(\mathcal{K}, \Delta)$$

Donde $Justified(\mathcal{K}, \Delta)$ es el conjunto de literales justificados en el sistema, a partir de la estructura lógica rebatible KB. $\square$

3.2 MTDR y Propiedades Horn

Inicialmente, debemos ver la coherencia lógica con respecto a las premisas iniciales, lo cual es lo mínimo que podemos requerir de un sistema.

Proposición 3.1: (*Satisfacción de Supraclasicalidad*) El operador C_Δ o simplemente C del sistema MTDR es supraclásico, esto es, $Cn(\mathcal{K}) \subseteq C(\mathcal{K})$.
PRUEBA.
La demostración es trivial a partir de la definición 3.1. $Cn(\mathcal{K}) \subseteq Cn(\mathcal{K}) \cup B$ para cualquier conjunto B. Luego, solo basta con tomar $B = Justified(\mathcal{K}, \Delta)$. ■

Pasemos ahora a estudiar tres propiedades íntimamente relacionadas, como se ha visto en la proposición 2.1: *absorción a derecha, equivalencia lógica a izquierda* y *cumulatividad subclásica.*

3.2.1 Absorción a Derecha

Dado que en todo el proceso de argumentación, se hace uso de la clausura deductiva de $\mathcal{K}$ y dada la idempotencia del operador clásico Cn se ve que tomar $\mathcal{K}$ ó $Cn(\mathcal{K})$ como premisas es equivalente. En el sistema argumentativo, no tiene relevancia como está expresado el conjunto inicial de conocimiento. Solo interesa lo que dicho conjunto $\mathcal{K}$ "expresa". Por ejemplo, para el sistema resulta irrelevante si $p \in \mathcal{K}$ ó si $q, q \rightarrow p \in \mathcal{K}$, en ambos casos $\mathcal{K}$ "está expresando" la proposición 'p'.

Proposición 3.2: (*Comportamiento ante la Absorción a Derecha*) El operador C_Δ o simplemente C del sistema MTDR satisface absorción a derecha¿ $C(\mathcal{K}) = C(Cn(\mathcal{K}))$.
PRUEBA.
Solo debemos probar que $Justified(\mathcal{K}, \Delta) = Justified(Cn(\mathcal{K}), \Delta)$ ya que es trivial, por idempotencia de Cn, que $Cn(\mathcal{K}) = Cn(Cn(\mathcal{K}))$.

Si probamos que los argumentos que se construyen para la base de conocimiento $\mathrm{KB}_1 = (\mathcal{K}, \Delta)$ son exactamente los mismos que se construyen para $\mathrm{KB}_2 = (Cn(\mathcal{K}), \Delta)$, entonces estaremos probando que los literales justificados serán los mismos para ambas bases de conocimiento.

Sea $\langle A, h\rangle$ un argumento válido para KB_1, entonces:

- $\mathcal{K} \cup A |\!\sim h$. Es fácil ver que $Cn(\mathcal{K}) \cup A |\!\sim h$ también, por la definición de $|\!\sim$ [2].
- $\mathcal{K} \cup A \not|\!\sim \bot$, y por lo tanto, $Cn(\mathcal{K}) \cup A \not|\!\sim \bot$ también.
- $\nexists A' \subset A, talque \mathcal{K} \cup A' |\!\sim h$. Supongamos que $\exists A' \subset A$ tal que $Cn(\mathcal{K}) \cup A' |\!\sim h$. Por definición de $|\!\sim$, obtenemos que $\mathcal{K} \cup A' |\!\sim h$, lo cual es un absurdo. Luego, $\nexists A' \subset A, talque Cn(\mathcal{K}) \cup A' |\!\sim h$.

Por los tres puntos anteriores se ve que $\langle A, h\rangle$ es también un argumento válido para la base de conocimiento KB_2.

Similarmente, sea $\langle A, h\rangle$ un argumento válido para KB_2, entonces:

- $Cn(\mathcal{K}) \cup A |\!\sim h$ y entonces $\mathcal{K} \cup A |\!\sim h$ también.
- $Cn(\mathcal{K}) \cup A \not|\!\sim \bot$, y por lo tanto, $\mathcal{K} \cup A \not|\!\sim \bot$ también..
- $\nexists A' \subset A, talque Cn(\mathcal{K}) \cup A' |\!\sim h$. Supongamos que $\exists A' \subset A$ tal que $\mathcal{K} \cup A' |\!\sim h$. Por definición de $|\!\sim$, obtenemos que $Cn(\mathcal{K}) \cup A' |\!\sim h$, lo cual es un absurdo. Luego, $\nexists A' \subset A, talque \mathcal{K} \cup A' |\!\sim h$.

Finalmente, $\langle A, h\rangle$ es también un argumento válido para KB_1, y el conjunto de argumentos válidos para $\mathcal{K}$ coincide con el conjunto de argumentos válidos para $Cn(\mathcal{K})$. Luego, $Justified(\mathcal{K}, \Delta) = Justified(Cn(\mathcal{K}), \Delta)$ y $C(\mathcal{K}) = C(Cn(\mathcal{K}))$. ■

Al cumplirse la absorción a derecha, se asegura la verificación de las propiedades de *equivalencia lógica a izquierda* y de *cumulatividad subclásica*. En el sistema argumentativo MTDR obtendremos las mismas justificaciones para dos conjuntos de premisas lógicamente equivalentes, ya que, los argumentos que se construyen serán los mismos en ambos casos. Gerard Vreeswijk analiza la equivalencia lógica a izquierda para otro marco argumentativo, y liga su satisfacción a un concepto de "localidad" en la determinación de las comparaciones entre argumentos. Si bien es difícil aplicar dicho concepto en nuestro caso, la idea de localidad expresa el hecho que no se tiene en cuenta, cuando se comparan argumentos, la forma en que se arriban a las conclusiones

[2] Recordemos que la derivación rebatible $|\!\sim$ incluye dentro de si la derivación clásica $\vdash$.

clásicas del conjunto de premisas, usadas en un argumento. De esta forma, *todas las conclusiones clásicas del conjunto inicial de premisas* se ubican en un único "nivel" dentro del sistema.

Pasemos ahora a la condición de *cumulatividad subclásica*. Como todas la proposiciones pertenecientes a la clausura clásica del conjunto $\mathcal{K}$ de premisas se encuentran al mismo nivel, acumular una de ellas al conjunto $\mathcal{K}$ no cambiará su status dentro del sistema. Como se remarca en [Sar98], acumular una conclusión cualquiera, determina casi siempre, un cambio de status de dicha conclusión, lo cual repercute directamente en las conclusiones próximas. La excepción a esto ocurren en el caso que sea una conclusión clásica del conjunto de premisas la que se acumule. Dicha conclusión, no cambiará su status, debido a que toda la clausura clásica está ubicada en un mismo nivel de categorización.

¿Dónde reside la coherencia lógica de *absorción a derecha*?. La coherencia está en el hecho que cuando la *información es totalmente segura*, como es el caso de las premisas, no debiera importar su forma, sino su contenido.

3.2.2 Absorción a Izquierda

La condición de absorción a izquierda nos indica que el conjunto de inferencias que se obtiene de un sistema, conforman, en si, una teoría clásica. ¿Sería coherente que un razonador infiera 'a' e infiera 'b', pero no infiera '$a \wedge b$'?. ¿Es lógico que un agente $\mathcal{A}$ **no "crea"** en conclusiones clásicas de sus creencias?. ¿Qué pensaríamos de alguien que nos dice que *desea ir al cine, si el día es lindo* y que efectivamente *el día esta lindo*, pero simultáneamente nos asegura que *no tiene ganas de ir al cine*?. Seguramente, no confiaríamos demasiado en sus futuras afirmaciónes. Alli, es donde las propiedades de absorción a izquierda, debilitamiento a derecha y conjunción de conclusiones se enmarcan.

En [Sar97a] se puede leer que *"con respecto a la absorción a izquierda, solo podemos decir, que pareciera no cumplirse debido a que, según su definición estricta, solo se pueden justificar literales. Así, puede que los literales 'h_1' e 'h_2' estén justificados ($h_1, h_2 \in Justified(\mathcal{K}, \Delta)$), pero no así la fórmula $h_1 \wedge h_2$. De todas maneras, podemos dislumbrar que esta falla responde a la estricta definición de lo que se puede justificar."*

Lo que intentaremos mostrar es que -en la "filosofía" del sistema- se respeta la absorción a izquierda y, por lo tanto, se logra un razonamiento coherente. Lo que sucede, es que las restricciones técnicas, nos impiden lograr aquellas inferencias que no son literales, y así, sentencias como '$a \wedge b$', '$a \vee b$', '$a \rightarrow b$' y otras más complejas no pueden justificarse. Luego, según la definición 3.1 del operador de consecuencia argumentativo, dichas sentencias no pertenecerían al conjunto de inferencias. Luego de analizar el comportamiento interno del sistema y su coherencia, re-definiremos el operador de consecuencia C_Δ.

De ahora en adelante, flexibilizaremos el concepto de argumento, permitiendo que un argumento pueda ser una explicación tentativa no solo para literales sino para cualquier sentencia válida. Por ejemplo, $\langle A$,a $\wedge$ b$\rangle$ y $\langle B, a \rightarrow b\rangle$ serán argumentos válidos.

Coherencia Interna del MTDR

Supongamos que tengamos un argumento justificado $\langle A, p\rangle$ y que $p \cup \mathcal{K} \vdash q$. Estamos ante la siguiente situación: $p \in C(\mathcal{K}), q \in Cn(\{p\} \cup \mathcal{K})$. ¿Se cumple que $q \in C(\mathcal{K})$?. Informalmente, queremos ver si las consecuencias clásicas de una sentencia justificada están también justificadas.

Proposición 3.3: Sea KB $= (\mathcal{K}, \Delta)$ una base de conocimiento y C el operador argumentativo. Si $p \in C(\mathcal{K})$ y $q \in Cn(\{p\} \cup \mathcal{K})$ entonces $q \in C(\mathcal{K})$.
Prueba.
Como $p \in C(\mathcal{K})$, existe un argumento $\langle P, p\rangle$ justificado. Es fácil ver que $\langle P, q\rangle$ es un argumento válido para 'q', el cual tendrá el mismo árbol dialéctico que el argumento

$\langle P, p\rangle$: $\langle P, q\rangle$ estará justificado tal cual como lo está $\langle P, p\rangle$ y, por lo tanto, $q \in C(\mathcal{K})$. ■

Veamos que con esta proposición podemos garantizar la verificación de la condición (3) de debilitamiento a derecha: si $x \rightarrow y \in Cn(A)$ entonces $y \in Cn(A \cup \{x\})$, y si, además, $x \in C(A)$ entonces se asegura por la proposición anterior que $y \in C(A)$, por lo que *se verifica el debilitamiento a derecha.*

El problema es más complejo de ver, cuando la sentencia que es consecuencia clásica de algo justificado, es en realidad, una consecuencia clásica de un **conjunto de sentencias justificadas**. Supongamos que $q \in Cn(\mathcal{K} \cup \{p_1, p_2, \ldots, p_n\})$ tal que existen justificaciones para $p_1, p_2, \ldots, p_n$. Queremos que exista una justificación para 'q', logrando que $q \in C(\mathcal{K})$ y verificando, así, la absorción a izquierda.

Proposición 3.4: Sea $\langle T, h\rangle$ un argumento y $\langle S, j\rangle$ un subargumento de $\langle T, h\rangle$. Luego $\langle T, h\rangle$ es más específico que $\langle S, j\rangle$, i.e. $\langle T, h\rangle \succeq_{\mathbf{spec}} \langle S, j\rangle$. (La demostración puede hallarse en [Sim92]). ■

Observación 3.1: Dado dos conjuntos de reglas rebatibles T_1 y T_2, tal que $\langle T_1 \cup T_2, h\rangle$, podemos asumir, sin perder generalidad, que $h \in Cn(\mathcal{K} \cup \{h_1, h_2\})$ donde $T_1 \cup \mathcal{K} |\!\sim h_1$ y $T_2 \cup \mathcal{K} |\!\sim h_2$. Intuitivamente, todo lo que aporta T_1 para la derivación de 'h' esta reflejado en 'h_1', mientras que lo que aporta T_2 para la derivación de 'h' está reflejado en 'h_2'. Para esto, solo basta con asumir que existe una "regla fuerte" $a_1 \wedge a_2 \ldots \wedge a_n \leftrightarrow h_1$ en $\mathcal{K}$, donde h_1 es un literal "nuevo"
footnotePor "nuevo" queremos decir que dicho literal no se repite en ninguna otra fórmula. - y otra análoga para h_2 - tal que $a_1, a_2, \ldots, a_n \in Lit(T_1)$ son todos los literales de T_1, necesarios para que 'h' pueda ser obtenido a partir de T_1, T_2 y $\mathcal{K}$. Resumiendo, lo que hacemos para simplificar es "colisionar" todo lo que es necesario de T_1 para derivar 'h' en un único literal 'h_1'. Para T_2 y 'h_2' es análogo. Por la forma en que se tratan los argumentos, asumir esas dos reglas en $\mathcal{K}$ no afectará al sistema, pero nos simplificará la demostración siguiente.

Proposición 3.5: Sea KB $= (\mathcal{K}, \Delta)$ una base de conocimiento. Sea $\langle A, p_1\rangle$ y $\langle B, p_2\rangle$ dos argumentos justificados. Si $q \in Cn(\mathcal{K} \cup \{p_1, p_2\})$ entonces $\langle A \cup B, q\rangle$ es un argumento justificado para 'q' en KB.

PRUEBA.

Daremos una demostración un tanto esquemática. Debemos probar dos puntos:

(1) $\langle A \cup B, q\rangle$ es un argumento para 'q' en KB:
Como $A \cup \mathcal{K} |\!\sim \{p_1\}$, $B \cup \mathcal{K} |\!\sim \{p_2\}$, y $\{p_1, p_2\} \cup \mathcal{K} |\!\sim q$ entonces $A \cup B \cup \mathcal{K} |\!\sim \{q\}$. Además, $A \cup B \cup \mathcal{K} \not|\!\sim \perp$ ya que, si sucediera, entonces existiría un sub-argumento $A' \subseteq A$ en desacuerdo con un sub-argumento $B' \subseteq B$ y como a lo sumo uno de ellos puede quedar justificado, entonces los argumentos A y B no podrían estar justificados a la vez (recordemos, que los sub-argumentos de un argumento justificado están siempre justificados). Informalmente, dos argumentos justificados no pueden estar "enfrentados" ya que si lo estuvieran, a lo sumo uno de ellos estaría justificado.

(2) $\langle A \cup B, q\rangle$ está justificado en KB:
Supongamos por el absurdo que existe un derrotador $\langle C, c\rangle$ para $\langle A \cup B, q\rangle$ que evita su justificación. Según el sub-argumento de desacuerdo $\langle D, d\rangle$ -$D \subseteq A \cup B$- las siguientes situaciones son posibles:

  (a) $D \subseteq A$ ó $D \subseteq B$. Como el argumento en desacuerdo es un sub-argumento de A ó de B, entonces $\langle C, c\rangle$ sería un derrotador para $\langle A, p_1\rangle$, ó para $\langle B, p_2\rangle$, que evitaría su justificación. Pero A y B están justificados y por lo tanto no podría existir tal derrotador $\langle C, c\rangle$.

  (b) $D = A \cup B$, esto es, el argumento de desacuerdo es *todo el argumento*. Luego, estamos ante la siguiente situación: $\mathcal{K} \cup \{p_1, p_2\} \vdash d$ y $\mathcal{K} \cup \{c, d\} \vdash \perp$ y por lo tanto $\mathcal{K} \cup \{c, p_1, p_2\} \vdash \perp$. Pero, como la inferencia clásica $\vdash$ es contrapositiva, entonces $\mathcal{K} \cup \{p_2, c\} \vdash \neg p_1$ y $\mathcal{K} \cup \{p_1, c\} \vdash \neg p_2$. **¿Qué nos dice esto?**. Ahora podemos construír los argumentos $\langle C \cup B, \neg p_1\rangle$ y $\langle C \cup A, \neg p_2\rangle$

que derrotarán a los argumentos $\langle A, p_1\rangle$ y $\langle B, p_2\rangle$ respectivamente. Esta derrota se debe nuevamente, a la transitividad de dicho criterio, ya que $C \cup A \succeq_{\mathsf{spec}} C \succeq_{\mathsf{spec}} A \cup B \succeq_{\mathsf{spec}} B$ y $C \cup B \succeq_{\mathsf{spec}} C \succeq_{\mathsf{spec}} A \cup B \succeq_{\mathsf{spec}} A$. Luego, los argumentos $\langle A, p_1\rangle$ y $\langle B, p_2\rangle$ no estarían justificados. Hemos asumido que la unión del argumento C con el argumento A resulta en un argumento no derrotado ó ganador dentro del árbol de dialéctica. De hecho, si no lo fuera, entonces por el mismo proceso anterior y sabiendo que el árbol dialéctico es finito, el argumento C tampoco sería ganador y no evitaría así la justificación inicial de $\langle A \cup B, \neg c\rangle$, lo cual es un absurdo. Luego, tal derrotador $\langle C, c\rangle$ no podría existir.

(c) $D = A' \cup B'$ donde $A' \subseteq A$ y $B' \subseteq B$. Este es el **caso más interesante**, ya que el desacuerdo nace a partir de la mezcla de las suposiciones tentativas que asumen el argumento A y el argumento B.
Primero, veamos que podemos asumir que $d = \neg c$. Como $\mathcal{K} \cup \{d, c\} \vdash \bot$, entonces, $\mathcal{K} \cup \{d\} \vdash \neg c$, i.e. $\neg c \in Cn(\mathcal{K} \cup \{d\})$ y, por lo tanto, el argumento de desacuerdo puede re-escribirse como $\langle D, \neg c\rangle$.
Como $A' \cup B' \cup \mathcal{K} |\!\sim \neg c$ entonces, por la definición de derivación rebatible $|\!\sim$, existirá un conjunto $CA \subseteq Lit(A')$[3] y un conjunto $CB \subseteq Lit(B')$, tal que, $CA \cup CB \cup \mathcal{K} \vdash \neg c$.
Simplificando, pero sin perder generalidad, supongamos que en realidad $CA \cup \mathcal{K} \vdash a'$ y $CB \cup \mathcal{K} \vdash b'$ donde $\{a', b'\} \cup \mathcal{K} \vdash \neg c$. Lo que decimos aquí, es que el conjunto A' "argumenta a favor" de a', mientras que el conjunto B' "argumenta a favor" de b'. Cuando los dos conjuntos se fusionan entonces se obtiene una nueva consecuencia clásica $\neg c$ que proviene de la argumentación a favor, tanto de a' (a partir de A'), como de b' (a partir de B'). Esta última inferencia $\neg c$ no aparece a partir de A' ni de B' únicamente, pero si en la "unión de ambos". Para ver esta simplificación, refirámosnos a la observación 3.1.
Luego, tenemos un argumento $\langle A', a'\rangle$ y otro $\langle B', b'\rangle$, ambos justificados por ser sub-argumentos de argumentos justificados, tal que $\langle A' \cup B', \neg c\rangle$ está en desacuerdo con $\langle C, c\rangle$. Por consiguiente, estamos en las condiciones del inciso (b) anterior y tal derrotador $\langle C, c\rangle$ no podría existir, ya que si existiera $\langle A', a'\rangle$ y $\langle B', b'\rangle$ no estarían justificados, lo cual es un absurdo.

■

Veamos esta proposición con un ejemplo sencillo:

Ejemplo 3.1: Sea KB $= (\mathcal{K}, \Delta)$ la siguiente base de conocimiento:

$$\begin{aligned} \mathcal{K} &= \{ \text{ s}, \ (a, b) \rightarrow h\} \\ \Delta &= \{ \ s \succ\!\!- d1, s \succ\!\!- d2, d1 \succ\!\!- a, d2 \succ\!\!- b, s \succ\!\!- \neg h\} \end{aligned}$$

Supongamos que $A = \langle\{s \succ\!\!- d1, d1 \succ\!\!- a\}, a\rangle$ y $B = \langle\{s \succ\!\!- d2, d2 \succ\!\!- b\}, a\rangle$ están justificados y, por ende, $d1, d2 \in C_\Delta(\mathcal{K})$. Ahora bien, como $h \in Cn(\{a, b\} \cup \mathcal{K})$ esperamos que $h \in C_\Delta(\mathcal{K})$ a través del argumento $AB = \langle A \cup B, h\rangle$, pero se puede ver que el argumento $C = \langle\{s \succ\!\!- \neg h\}, \neg h\rangle$ derrota propiamente a AB evitando su justificación. Luego no se cumpliría que $h \in C_\Delta(\mathcal{K})$.

Al ver la construcción de la demostración en la proposición 3.5 deberíamos poder encontrar un derrotador para los argumento A y B usando el argumento C y la contrapositiva de la derivación clásica. Efectivamente , podemos ver que $\langle C \cup A, \neg b\rangle$ evita la justificación de $\langle B, b\rangle$ y, simétricamente, $\langle C \cup B, \neg a\rangle$ evita la justificación de $\langle A, a\rangle$. Graficamente podemos ver esto en la figura 1. □

Observación 3.2: Se puede ver en todo el estudio de esta interesante propiedad, lo importante que es el rol de la **propiedad contrapositiva** de la derivación clásica. De hecho, en sistemas aproximados al MTDR, como la programación en lógica rebatible

[3] $Lit(A)$ es el conjunto de todos los literales usados en las reglas rebatibles de A.

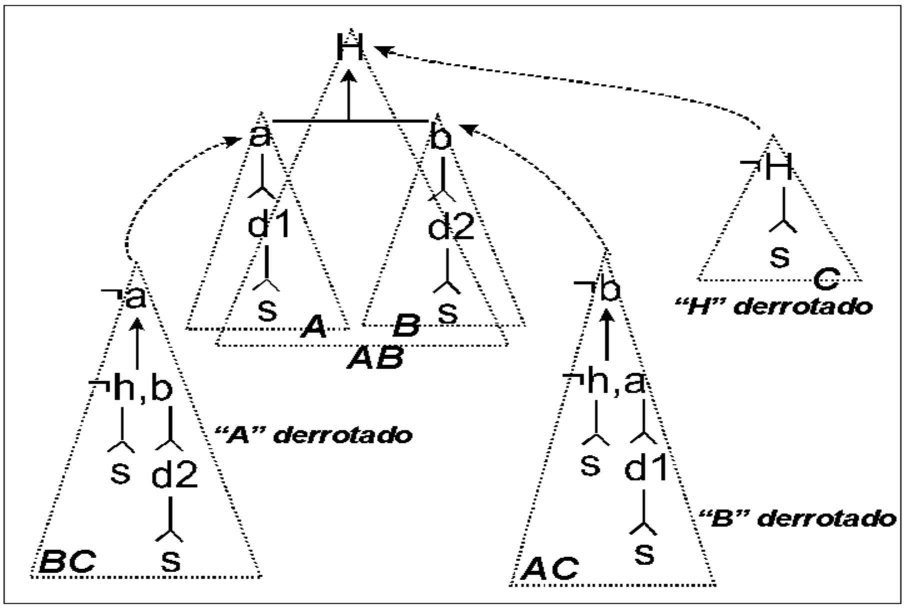

Figura 1: Unión de argumentos justificados.

de Alejandro García ([Gar97]), donde no se usa exactamente la derivación clásica para la construcción de argumentos, no se cumple la proposición 3.5, fallando, así, la absorción a derecha.

A través de la proposición anterior, podemos ver que el sistema argumentativo se comporta coherentemente a la hora dibujar inferencias. Estamos ante un sistema que respeta la absorción a izquierda, donde sus conclusiones inferidas son cerradas bajo la consecuencia clásica Cn: *el conjunto de inferencias conforma una teoría clásica.*

3.2.3 Resolviendo el inconveniente técnico

Como se pudo ver, la no justificación de sentencias que no sean literales evita obtener algunas consecuencias clásicas del conjunto de conclusiones. Además, hemos verificado que esto es solo un problema técnico, debido que, intrínsicamente, el sistema se comporta coherentemente. Ahora bien, ¿cómo serían los argumentos para sentencias que no sean solo literales?. Una primera idea para buscar dichos argumentos podría ser la siguiente:

- La unión de un argumento para 'a' con un argumento para 'b' podría ser un potencial argumento para $a \wedge b$.
- Un argumento para 'a' es también un argumento para '$b \rightarrow a$'.
- Un argumento para '$\neg a$' es también un argumento para '$a \rightarrow b$'.
- Un argumento para 'a' es también un argumento para '$a \vee b$'.

Así, si se desea encontrar un argumento para '$a \rightarrow (b \rightarrow c)$' entonces tanto un argumento a favor de '$\neg a$' u otro a favor de '$\neg b$' u otro que soporte 'c' serán suficientes. Finalmente, si se desea buscar un argumento que soporte la sentencia '$a \wedge (b \vee c)$' el sistema podría buscar un argumento para 'a' y otro para 'b', o bien, buscar uno para 'a' y otro para 'c'.

3.3 MTDR y Propiedades No-Horn

Luego de haber estudiado las propiedades Horn, pasaremos a ver cuál es el comportamiento argumentativo con respecto a las tres propiedades descriptas en la sección 2.2. Daremos ejemplos técnicos e intuitivos que muestran la no verificación de estas condiciones en el MTDR.

Ejemplo 3.2: (*Falla de Negación Racional y Monotonía Racional*) Sea la base de conocimiento $\mathrm{KB} = (\mathcal{K}, \Delta)$ tal que:

$$\begin{array}{rcl} \mathcal{K} & = & \{ \text{ d } \} \\ \Delta & = & \{ \ d \succ\!\!- i, i \succ\!\!- a, a \succ\!\!- c, q \succ\!\!- \neg c, \neg q \succ\!\!- \neg i \ \} \end{array}$$

Se puede ver que $c \in C_\Delta(\mathcal{K})$ ya que el argumento $A = \langle \{d \succ\!\!- i, i \succ\!\!- a, a \succ\!\!- c\}, c\rangle$ no posee ningún contra-argumento y por lo tanto está justificado.

Ahora bien, $c \notin C_\Delta(\mathcal{K} \cup \{q\})$ ya que el argumento $Q_1 = \langle \{q \succ\!\!- \neg c\}, \neg c\rangle$ logra derrotar por bloque al anterior argumento A. Además, $c \notin C_\Delta(\mathcal{K} \cup \{\neg q\})$ ya que el argumento $Q_2 = \langle \{\neg q \succ\!\!- \neg i\}, \neg i\rangle$ derrota al sub-argumento $\langle \{d \succ\!\!- i\}, i\rangle$ de A, evitando así que A esté justificado.

Con respecto a la monotonía racional, veamos que $\neg q \notin C_\Delta(\mathcal{K})$, por lo que, unido a lo anterior, no se satisface. $\square$

Intuitivamente, supongamos que un agente $\mathcal{A}$ razona sobre el voto de una persona en una elección para lo cual realizamos el siguiente renombre:

$$\begin{array}{rcl} d & = & desconoce\_plataformas \\ i & = & desinteresado\_votar \\ a & = & vota\_al\_azar \\ c & = & cualquier\_partido \\ q & = & afinidad\_partido\_A \end{array}$$

Si $\mathcal{K} = \{d\} = \{desconoce\_plataformas\}$, entonces el agente razonará que, al desconocer las plataformas de cada partido, no está interesado en votar, por lo que "votará al azar" concluyendo que cualquier partido puede ser el votado. Si luego se agrega el conocimiento que esta persona tiene afinidad con un cierto partido, entonces, hay un argumento para razonar que dicha persona, aún desconociendo las plataformas políticas, **no votará a cualquier** partido político:

$$Q_1 = \langle \{afinidad\_partido\_A \succ\!\!- \neg cualquier\_paritido\}, \neg cualquier\_partido\rangle$$

Por otra parte, si el conocimiento fuera que en realidad no hay afinidad con cierto partido, entonces, habrá un argumento válido para pensar que **no es cierto que esta persona no está interesada** en votar:

$$Q_2 = \langle \{\neg afinidad\_partido\_A \succ\!\!- \neg desintaresado\_votar\}, \neg desinteresado\_votar\rangle$$

Seguramente, no querrá votar a dicho partido, y por lo tanto no votará a cualquier partido. Así, vemos que tanto si tiene afinidad con el partido o como si efectivamente no la tiene, son - ambos - motivos suficientes para creer que no se votará a cualquier partido.

Sobre la monotonía racional, veamos que en primera instancia se infiere que se votará a cualquier partido, ya que se desconoce las plataformas políticas. De todas formas, no se puede inferir con el mismo conocimiento que no se posee afinidad con el partido A, aunque al agregar dicha afinidad al conocimiento se deseche la inferencia primera. En resumen, *no solo la negación de lo esperado* puede forzarnos a eliminar conlusiones plausibles previamente inferidas. Pensamos que, una restricción de este estilo resultaría en un sistema "menos rebatible" y por lo tanto coincidimos con Makinson en que monotonía racional es **demasiado pedir** para un sistema de razonamiento no-monótono.

Discutamos el comportamiento ante estas dos propiedades de manera abstracta: si vemos el ejemplo 3.2, podemos observar que no se cumple la condición de *negación racional* debido a que una sentencia y su negación pueden simultaneamente evitar la derivación de otra sentencia. Dado que en el sistema argumentativo, el concepto de derivación está flexibilizado por las reglas rebatibles, podríamos bien tener el caso que dos sentencias totalmente opuestas - $afinidad\_partido\_A$ y $\neg afinidad\_partido\_A$-

impidan, ambas, la derivación de otra proposición: *cualquier_partido*. En nuestro caso, el punto de derrota es distinto en cada caso. Parecería que estaríamos ante la situación que dos sentencias opuestas se comportan como si no lo fueran, ya que ambas "coinciden' en contradecir a la misma proposición, pero esto es entendible bajo el concepto de lo que una derivación rebatible y flexible significa.

La negación racional parece ser una propiedad demasiado restrictiva, ya que no creemos conveniente obligar a nuestro sistema a inferir la negación de todo lo que puede hacernos retractar de una inferencia plausible. Queremos un sistema que pueda desconocer tanto x como $\neg x$ y adaptarse correctamente al agregar a su conocimiento cualquiera de dichas proposiciones.

Estudiemos ahora, nuestra última propiedad no-Horn.

Ejemplo 3.3: (*Falla de Racionalidad Disjuntiva*) Sea la base de conocimiento KB $= (\mathcal{K}, \Delta)$ tal que:

$$\begin{aligned} \mathcal{K} &= \{\ \text{gallina} \ \vee \text{tweety}, gallina \vee tweety \rightarrow ave\} \\ \Delta &= \{\ ave \succ\!\!- alas, alas \succ\!\!- vuela, gallina \succ\!\!- \neg vuela, tweety \succ\!\!- \neg alas\ \} \end{aligned}$$

Este es el ya conocido ejemplo de las aves que vuelan citado ampliamente en la literatura del razonamiento no-monótono. Aquí, un agente $\mathcal{A}$ inferirá que el individuo sobre el cual está razonando, efectivamente vuela ya que se puede armar el siguiente argumento sin derrotadores:

$$A = \langle \{ave \succ\!\!- alas, alas \succ\!\!- vuela\}, vuela\rangle$$

Con esto, aseguramos que $vuela \in C_\Delta(\mathcal{K})$ pero sucede que $vuela \notin C_\Delta(\mathcal{K} \cup \{gallina\})$ y $vuela \notin C_\Delta(\mathcal{K} \cup \{tweety\})$. El primer caso se debe a que se puede armar un argumento $\langle \{gallina \succ\!\!- \neg vuela\}, \neg vuela\rangle$ que derrota a A y en el segundo caso se puede obtener el argumento $\langle \{tweety \succ\!\!- \neg alas\}, \neg alas\rangle$ que nos dice que efectivamente tweety no posee alas (por algún motivo) y derrota, así, a un sub-argumento de A y por ende a A también. No se cumple así, la propiedad de racionalidad disjuntiva. □

Podemos ver que, en el MTDR, la racionalidad disjuntiva no se soporta por una cuestión de "conocimiento más preciso". Si tenemos un conocimiento disjuntivo -como por ejemplo que un induviduo es una *gallina* ó que es *tweety*- tenemos una información menos precisa que si se conociera efectivamente alguna parte de la disjunción -que es una *gallina* ó, equivalentemente, que ese individuo es efectivamente *tweety*. Dicho conocimiento más preciso, podría llevarnos a cambiar drásticamente nuestras inferencias. En nuestro caso, saber que es una *gallina* o que es *tweety*, sólo nos asegura el hecho que estamos frente a un ave y por lo tanto inferimos tentativamente que vuela. Si refinamos nuestro conocimiento un poco más, entonces lograremos, en ambos casos y por distintas causas, razonar que dicho individuo no vuela. Lo que sucede es que, implicitamente, tanto si es una *gallina* como si es *tweety* no podrá volar pero no podemos hacer explicito en nuestro sistema dicho razonamiento en este punto. Si bien parecería, que esta propiedad es deseable, hace a nuestro sistema menos rebatible ante información menos precisa.

3.4 Redefiniendo el Operador de Consecuencia Argumentativo

Dado que en el sistema argumentativo MTDR se respeta internametne la coherencia lógica a la hora de construír las argumentaciones, redefiniremos nuestro operador de consecuencia argumentativo de manera que esta coherencia interna se vea explícitamente reflejada. ¿Qué necesitamos para lograr esto?. Podemos pensar en un *nuevo sistema argumentativo* que, una vez obtenidas todas las justificaciones posibles, obtenga luego todas las consecuencias clásicas a partir de ellas. Así, nuestro nuevo sistema argumentativo estaría formado por la máquina argumentativa original y, por "sobre ella", una máquina de inferencia clásica. Con esto, no necesitamos modificar el sistema original.

Es importante remarcar que, este agregado a nuestro sistema MTDR **no es arbitrario**, sino que se basa en el hecho que el mecanismo de razonamiento argumentativo dado por el MTDR respeta la coherencia lógica que estamos agregando. Si no lo hiciera, estaríamos **obligando** al sistema a respetar dicha coherencia cuando en realidad no lo hace, no siendo natural ni sano tal modificación.

Luego, nuestro nuevo sistema argumentativo (el MTDR y por sobre él una máquina de inferencia clásica) tendría el siguiente operador de consecuencia asociado para una base de conocimiento KB $= (\mathcal{K}, \Delta)$:

$$C_\Delta(\mathcal{K}) = Cn(Justified(\mathcal{K}, \Delta))$$

Observemos que hemos incluído dentro de los justificable, todo lo que es derivable clásicamente a partir de las premisas. De hecho, para cualquier litreal $a \in Cn(\mathcal{K})$ existe un argumento $\langle \phi, a \rangle$ que está justificado ya que no puede tener ningún derrotador. De hecho, cualquier argumento que lo contra-argumente sería inconsistente con el conjunto de premisas $\mathcal{K}$.

4 Conclusiones

En primer lugar, en este trabajo, hemos identificado algunas propiedades interesantes que juzgan el razonamiento de un sistema de inferencia desde el punto de vista de la *coherencia lógica*. Dividimos a dichas propiedades en dos grupos: Horn y no-Horn, según la sintaxis con que estas se expresan. Para cada condición se dio una explicación intuitiva de que es lo que la propiedad nos dice y porque deseamos que un "buen" sistema la respete.

Posteriormente, hemos estudiado la propiedad de absorción total, dividida en dos partes, las cuales con sintaxis similares, expresan muy distintas ideas. Otras propiedades, como equivalencia lógica a izquierda, conjunción de conclusiones, debilitiamiento a derecha y cumulatividad subclásica son consecuencias directas de estas dos propiedades.

Luego, analizamos el comportamiento argumentativo ante tres propiedades no-Horn y pudiendo observar la no satisfacción de ninguna de ellas. Hemos dado ejemplos y explicaciones técnicas e intuitivas del porqué de este comportamiento.

Finalmente, se redefinió el operador de consecuencia argumentativo en función de lo estudiado, de manera que se refleje explícitamente la coherencia lógica que el sistema internametne logra verificar.

Referencias

[Gab85] D.M.Gabbay: *Theoretical Foudations for Nonmonotonic Reasoning in Expert Systems.* K. Apt *Eds.*, Logics and Models of Concurrent Systems, 1985, Springer-Verlag, Berlin.

[Gar97] A.J.García. *La Programación en Lógica Rebatible: Su Definición Teórica y Computacional.* Tesis de Magister, Departamento de Ciencias de la Computación, Universidad Nacional del Sur, 1997.

[Gar93] A.J.García, C.I.Chesñevar,G.R.Simari, *Making Argumentative Systems Computationally Attractive*, *Proceedings of the XIII International Conference of the Chilenean Society for Computer Science*, Octubre 1993.

[Kra90] S.Kraus, D.Lehmann, M.Magidor, Nonmotonic Reasoning, Preferential Models and Cumulative Logics. *Artificial Intelligence*, **44**:167-207, 1990.

[Mak89] D.Makinson. *General Theory of Cumulative Inference*, In *Proceedings of the 2nd. International Workshop on Nomonotonic Reasoning (LNAI 346)* , 1989.

[Mak90] D.Makinson. *General Patterns of Nonmonotonic Reasoning*, In *Handbook of Logic in Artificial Intelligence and Logic Programming* D.Gabbay, *Ed.*, Vol. III, Oxford University Press,1990.

[Sar97a] S. Sardiña. *Patrones Generales en Sistemas de Razonamiento Nomonóntono.* Tesis de Licenciatura, Departamento de Ciencias de la Computación, Universidad Nacional del Sur, 1997.

[Sar98] S.Sardiña, G.R.Simari. *El Operador de Consecuencia en los Sistemas Argumentativos.* En *Anales del IV Congreso Internacional de ingeniería Informática, ICIE'98*, pp. 20-35, Buenos Aires, 1998.

[Sim94] G.R.Simari, C.I.Chesñevar, A.J.García, *The Role of Dialectics in Defeasible Argumentation.* En *XIV International Conference of the Chilean Computer Society*, Concepción, Chile, Octubre 1994.

[Sim92] G.R. Simari, R.P. Loui. *A Mathematical Treatment of Defeasible Reasoning and its Implementation.* Artificial Intelligence **53**:125-157, 1992.

[Vre92] G. Vreeswijk. *Nonmonotonicity and Partiality in Defeasible Argumentation.* En *Nonmonotonic Reasoning and Partial Sematics*, Ellis Horwood, *Ed.*, 1992.