Universidad Nacional del Sur

Tesis de Doctorado en Ciencias de la Computación

Formalización de Sistemas Argumentativos
con Ataque y Soporte a Inferencias

Andrea Cohen

Bahía Blanca Argentina

2014

Prefacio

Esta Tesis se presenta como parte de los requisitos para optar al grado Académico de Doctorado en Ciencias de la Computación, de la Universidad Nacional del Sur y no ha sido presentada previamente para la obtención de otro título en esta Universidad u otra. La misma contiene los resultados obtenidos en investigaciones llevadas a cabo en el ámbito del Departamento de Ciencias e Ingeniería de la Computación durante el período comprendido entre el 1 de Abril de 2010 y el 4 de Diciembre de 2014, bajo la dirección del Dr. Alejandro J. García, Profesor Asociado del Departamento de Ciencias e Ingeniería de la Computación y del Dr. Guillermo R. Simari, Profesor Titular del Departamento de Ciencias e Ingeniería de la Computación.

....

Andrea Cohen
ac@cs.uns.edu.ar
Departamento de Ciencias e Ingeniería de la Computación
Universidad Nacional del Sur
Bahía Blanca, 4 de Diciembre de 2014

UNIVERSIDAD NACIONAL DEL SUR
Secretaría General de Posgrado y Educación Continua

La presente tesis ha sido aprobada el .../.../..., mereciendo la calificación de ......(.........)

Agradecimientos

Quiero agradecer a todas las personas e instituciones que hicieron posible la culminación de este doctorado. En primer lugar, a mis directores Alejandro García y Guillermo Simari: gracias por haberme enseñado y acompañado durante estos años, formándome como profesional y brindándome las herramientas necesarias para desempeñarme con éxito en el mundo de la investigación. Quiero agradecer también a la Universidad Nacional del Sur, por haberme brindado la oportunidad de desarrollar aquí mis actividades de investigación. Asimismo quiero agradecer al Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET) por haberme brindado el sustento necesario para el desarrollo y finalización de mi doctorado.

Gracias a todos los miembros del Departamento de Ciencias e Ingeniería de la Computación (DCIC) por haberme brindado un lugar de trabajo y por hacer que el venir a trabajar todos los días no sea una carga sino un disfrute. Por los mates compartidos, las charlas, las risas y llantos: gracias! En especial, gracias a todos mis compañeros de la sala de becarios, muchos de los cuales hoy son mis amigos. A los que están y los que estuvieron: gracias por todo! Esto no termina acá, me van a tener que seguir aguantando :)

A mi familia: gracias infinitas! En especial a mis papás, por haberme inculcado que todo esfuerzo conlleva su recompensa. Que lo que uno logra con dedicación queda para siempre. Gracias por ser como son y por haberme formado de esta manera. Los quiero hasta el infinito y más allá! A Gotti... Qué decir? Gracias es poco. Gracias por haberme acompañado durante esta etapa de mi vida, tanto en lo personal como en lo profesional. Disfruté mucho de trabajar junto a vos, a pesar de que algunas veces me enojé más de lo debido ;) Espero que esto haya sido solo el comienzo. Gracias por estar siempre, por darme ánimos cuando estuve bajoneada, por festejar nuesros logros. Gracias por hacerme feliz, y espero seguir siendo feliz a tu lado por el resto de mi vida :)

A todos mis amigos (no quiero olvidarme de nadie): gracias!! Son una parte indispensable de mi vida, con lo cual les agradezco el haber estado ahí en las alegrías y en las tristezas. Gracias por todos los momentos compartidos y por todos los que vendrán. En especial, quiero agradecer Gime y Lau: las adoro! Gracias eternamente por todo :)

Cada uno de ustedes sabe que aportó, de una manera u otra, su granito de arena en esta etapa de mi vida. Gracias, gracias y más gracias!

Noni

Resumen

Esta tesis comprende el estudio y formalización de herramientas para la representación de conocimiento y razonamiento en sistemas argumentativos. En particular, se proponen dos formalismos argumentativos que permiten la utilización y combinación de soporte y ataque a reglas de inferencia. Los formalismos aquí desarrollados abordan esta temática desde dos enfoques complementarios. El primero de ellos provee un marco unificado para modelar estas nociones en un contexto de argumentación abstracta, mientras que el segundo corresponde a un sistema argumentativo basado en reglas que permite expresar soporte y ataque a reglas rebatibles de un lenguaje de programación lógica. Para cada uno de los formalismos propuestos se abordará la definición del sistema, así como también el análisis de propiedades satisfechas por el mismo.

Los sistemas argumentativos definidos en esta tesis extienden a otras aproximaciones existentes en la literatura, incorporando características y elementos que aún no habían sido considerados conjuntamente por los formalismos desarrollados hasta el momento. Concretamente, los formalismos aquí propuestos incorporan en forma conjunta las nociones de backing y undercutting, de reconocida importancia en el área de argumentación. Asimismo, estos formalismos fueron concebidos de manera tal que es posible instanciar el marco argumentativo abstracto con el sistema argumentativo concreto, favoreciendo así su implementación computacional.

Abstract

This thesis concerns the study and formalization of tools for knowledge representation and reasoning in argumentation systems. We propose two argumentation formalisms that allow for the consideration and combination of attack and support for inference rules. The former provides a unified framework for modeling these notions within the context of abstract argumentation, whereas the latter corresponds to a rule-based argumentation system that allows for attack and support for defeasible rules in a logic programming setting. For each of the proposed systems we provide its formal definiton, and we analyze several properties that are satisfied.

The formalisms developed in this thesis extend other existing approaches in the literature by incorporating some elements and features that were not yet considered together by the formalisms developed so far. Specifically, the argumentation systems proposed here incorporate the notions of backing and undercutting in a joint manner; two notions of great importance within the community of argumentation. Also, these formalisms were conceived in such a way that it is possible to instantiate the abstract argumentation framework with the rule-based argumentation system, thus benefiting its computational implementation.

Índice general

Capítulo 1

Introducción

Esta tesis aborda el estudio y formalización de herramientas para la representación de conocimiento y razonamiento en sistemas argumentativos, las cuales incorporan la posibilidad de utilizar y combinar soporte y ataque a reglas de inferencia. Como parte de la contribución, se propone un formalismo que provee un marco unificado para modelar estas nociones en un contexto de argumentación abstracta. Por otra parte, se propone un sistema argumentativo basado en reglas que permite representar soporte y ataque a reglas rebatibles de un lenguaje de programación lógica. Este sistema permite la instanciación del marco argumentativo abstracto propuesto, poseyendo además características que favorecen su implementación computacional.

Como se verá a continuación, los sistemas argumentativos han demostrado ser de gran utilidad en diferentes dominios de aplicación. Los formalismos propuestos en esta tesis incorporan características y elementos que aún no habían sido considerados en forma conjunta por otras aproximaciones desarrolladas hasta el momento. En consecuencia, los resultados obtenidos en esta tesis brindan una contribución importante a los desarrollos en la comunidad de argumentación, significando esto un aporte dentro del área de Inteligencia Artificial en las Ciencias de la Computación.

En lo que resta de este capítulo se presentará brevemente el contexto en el cual se desarrolla esta tesis, y se describirán las principales contribuciones y resultados obtenidos. Finalmente, se indicará la organización del resto de la tesis, detallando brevemente el contenido de los capítulos siguientes.

Argumentación Rebatible

Argumentación es una forma de razonamiento en la cual, para una determinada afirmación (conclusión), se presta atención explícita a las justificaciones presentadas y a la resolución de los posibles conflictos entre ellas. En este tipo de razonamiento, una afirmación es aceptada o rechazada en función del análisis de los argumentos a su favor y en su contra. La forma en que los argumentos y las justificaciones para una afirmación son considerados permite definir un mecanismo de razonamiento automático, en el cual puede haber información contradictoria, incompleta e incierta [PV02, RS09].

En las últimas décadas, argumentación ha evolucionado como un atractivo paradigma para conceptualizar el razonamiento de sentido común [CML00, BCD07, BH08, AyC14]. Existen diferentes aproximaciones para modelar argumentación, tales como sistemas argumentativos abstractos [Dun95, Pra09] o sistemas argumentativos basados en reglas [PS97, GS04]. Por otra parte, este tipo de razonamiento es particularmente atractivo para la toma de decisiones, y dentro del área de Inteligencia Artificial existe especial interés en abordar este tipo de problemas. Concretamente, los sistemas argumentativos han sido utilizados en diversos dominios y aplicaciones como el razonamiento legal [PS02, Ver03a], sistemas para la toma de decisiones y negociación [APM00, RRJ+03, BH09b], y sistemas multi-agente [PSJ98, AMP02].

El uso de Soporte en Argumentación

Como se mencionó anteriormente, los sistemas argumentativos han demostrado ser de gran utilidad en diferentes dominios de aplicación. Una de las razones por las cuales argumentación resulta útil es que permite manejar los conflictos surgidos de la consideración de información inconsistente. Tales conflictos suelen capturarse mediante la noción de *derrota* entre argumentos [Dun95, GS04]. Sin embargo, en trabajos como [WGM+97, CGFE01], en los que se utiliza argumentación como herramienta de razonamiento en aplicaciones concretas, se observa que al evaluar la evidencia disponible no sólo existen derrotas entre argumentos, sino que también se identifican situaciones en las que los argumentos se *soportan* entre sí.

La noción de soporte ha estado presente en la literatura de argumentación desde sus trabajos fundacionales en la filosofía [Tou58]. Estudios subsiguientes en el área de argumentación computacional dejaron de lado la noción de soporte para concentrarse en la noción de derrota. En particular, uno de los trabajos de este área [Dun95] (muy influyente en los desarrollos de argumentación abstracta) considera un conjunto de argumentos y una relación de derrota entre ellos, la cual es utilizada para identificar los argumentos aceptados del sistema, sin contemplar la existencia de soporte. Siguiendo la aproximación de [Dun95], gran parte de los trabajos en la literatura se ha dedicado al desarrollo de formalismos argumentativos cuyo foco se encuentra en la relación de derrota. Sin embargo, en la última década, el estudio de la noción de soporte recobró atención entre los investigadores del área. Como resultado de estas investigaciones, se han desarrollado diversas propuestas que adoptan diferentes interpretaciones para la noción de soporte [CGGS13a, CLS13].

En particular, como se verá a continuación, las contribuciones de esta tesis conciernen el estudio de diferentes interpretaciones de la noción de soporte para la definición de formalismos argumentativos que contemplan el uso de soporte en diferentes entornos.

1.1. Contribuciones

Las principales contribuciones de esta tesis pueden sintetizarse de la siguiente manera:

Marco Argumentativo Abstracto con Soporte

Se propone un sistema de argumentación abstracta llamado *Marco Argumentativo con Backing y Undercutting (BUAF)* (por su nombre en inglés, *Backing-Undercutting Argumentation Framework*). El BUAF extiende al marco argumentativo abstracto introducido por Dung en [Dun95] mediante la incorporación de una relación especializada de soporte entre argumentos, una distinción entre diferentes tipos de ataque, y una relación de preferencia entre argumentos. De esta manera, el BUAF permite representar y utilizar en forma combinada dentro de un mismo formalismo dos nociones de reconocida importancia en el área de argumentación: el soporte existente entre las nociones de backing y warrant propuestas por Toulmin [Tou58] y la noción de undercutting defeater introducida por Pollock [Pol87]. En consecuencia, el BUAF provee un marco unificado que permite modelar ataque y soporte a inferencias en un contexto de argumentación abstracta, lo cual no había sido considerado en forma conjunta por los formalismos existentes en la literatura. Los conflictos entre los argumentos de un BUAF son resueltos mediante el uso de preferencias, dando lugar a un conjunto de derrotas que son utilizadas para determinar los argumentos aceptados del sistema. Finalmente, como parte de esta contribución, se incluye la formalización y demostración de diversas propiedades satisfechas por el BUAF.

Programación en Lógica Rebatible con Soporte y Ataque a Reglas Rebatibles

Se propone un sistema argumentativo basado en reglas llamado *Programación en Lógica Rebatible Extendida (E-DeLP)* (por su nombre en inglés, *Extended Defeasible Logic Programming*). E-DeLP extiende a la Programación en Lógica Rebatible (DeLP) [GS04] mediante la incorporación de dos nuevos tipos de reglas que permiten expresar razones a favor y en contra de reglas rebatibles, capturando así las nociones de undercutting defeater [Pol87] y backing [Tou58]. De esta manera, E-DeLP ofrece mejores capacidades de representación y razonamiento, proveyendo además los medios para modelar conjuntamente las ideas de Pollock y Toulmin en un lenguaje de programación lógica. En particular, esto no había sido considerado por ninguno de los sistemas argumentativos basados en reglas propuestos con anterioridad en la literatura. Cabe destacar que E-DeLP permite la instanciación del marco argumentativo abstracto (BUAF) propuesto en esta tesis. Asimismo, E-DeLP posee características que favorecen su implementación computacional ya que, siguiendo el enfoque orientado a consultas de la programación lógica, se propone una aproximación procedural que define un modelo de prueba dialéctico para E-DeLP. Por lo tanto, esto brinda la posibilidad de efectuar el cálculo de aceptabilidad para los argumentos de E-DeLP por demanda, en función de las consultas realizadas, sin tener que realizar el cálculo de aceptabilidad para todos los argumentos del sistema. Por último, como parte de esta contribución, se identificó y demostró una serie de propiedades satisfechas por E-DeLP, las cuales conciernen la aceptabilidad de los argumentos y la sensatez de las inferencias del sistema.

1.2. Publicaciones

A continuación se indican los artículos publicados como resultado de los trabajos llevados a cabo durante la realización de esta tesis. En particular, para cada uno de estos trabajos, se detalla brevemente su contribución y se indican los capítulos de esta tesis con los que se vinculan sus resultados.

- En el trabajo *"A survey of different approaches to support in argumentation systems"* [CGGS13b], publicado en *The Knowledge Engineering Review*, se estudiaron diferentes interpretaciones para la noción de soporte propuestas en la literatura, realizando un análisis comparativo destacando similitudes y diferencias entre ellas. Los resultados de este artículo fueron utilizados para los desarrollos de los Capítulos 3 y 4.

- En los trabajos:

  - *"An Argumentation Framework with Backing and Undercutting"* [CGS11a], [CGS12a], publicado en *XVII Congreso Argentino de Ciencias de la Computación (CACIC 2011)* y seleccionado para su publicación en *Computer Science & Technology Series: XVII Argentine Congress of Computer Science Selected Papers*; y
  - *"Backing and Undercutting in Abstract Argumentation Frameworks"* [CGS12b], publicado en *7th Int. Symposium on Foundations of Information and Knowledge Systems (FoIKS 2012), tomo 7153 de Lecture Notes in Computer Science*

  se definió un marco argumentativo abstracto que incorpora una relación de soporte. En particular, el formalismo desarrollado en estos trabajos adopta la interpretación de *backing* para la relación de soporte, caracterizando un *Marco Argumentativo con Backing y Undercutting (BUAF)*. Los resultados de estos trabajos se hallan estrechamente vinculados con el formalismo propuesto en el Capítulo 4.

- En los trabajos:

  - *"Extending DeLP with Attack and Support for Defeasible Rules"* [CGS10], publicado en *12th Ibero-American Conference on Artificial Intelligence (IBERAMIA 2010), tomo 6433 de Lecture Notes in Computer Science*; y
  - *"Backing and Undercutting in Defeasible Logic Programming"* [CGS11b], publicado en *11th European Conference on Symbolic and Quantitative Approaches to Reasoning with Uncertainty (ECSQARU 2011), tomo 6717 de Lecture Notes in Computer Science*

  se propuso un sistema argumentativo basado en reglas que extiende a la programación en lógica rebatible [GS04]. De esta manera, se definió el formalismo de *Pro-*

gramación en Lógica Rebatible Extendida (E-DeLP) que permite expresar razones a favor y en contra de reglas de inferencia, capturando así las nociones de *backing* y *undercutting defeater*. Los resultados de estos trabajos se hallan reflejados en los desarrollos del Capítulo 5.

1.3. Organización de la Tesis

El resto de esta tesis se encuentra organizado de la siguiente manera:

- En el Capítulo 2 se introducirá una serie de nociones básicas de argumentación, las cuales proveen las bases para los desarrollos de esta tesis. En primer lugar, se presentará una estructura conceptual que caracteriza los diferentes elementos presentes en los sistemas argumentativos. Seguidamente, se presentarán dos aproximaciones para la formalización de sistemas argumentativos: los marcos argumentativos abstractos y la programación en lógica rebatible. Finalmente, se motivará la existencia de una relación de soporte entre argumentos como mecanismo para ampliar las capacidades de representación y razonamiento de los sistemas argumentativos.

- En el Capítulo 3 se efectuará una revisión de diferentes interpretaciones para la noción de soporte propuestas en la literatura. Se realizará un estudio comparativo, destacando similitudes y diferencias entre estas interpretaciones, y se discutirá cómo son abordadas por diferentes formalismos argumentativos. En particular, el análisis se enfocará en el significado asociado a las interpretaciones de soporte, teniendo en cuenta las restricciones de aceptabilidad impuestas por cada una de ellas.

- En el Capítulo 4 se desarrollará una propuesta que contempla una forma alternativa de soporte entre argumentos, a diferencia de aquellas analizadas en el Capítulo 3. Se desarrollará un formalismo de argumentación abstracta, llamado *Marco Argumentativo con Backing y Undercutting (BUAF)*, que permite combinar dos nociones importantes del área de argumentación presentadas en los capítulos 2 y 3: el soporte existente entre las nociones de backing y warrant propuestas por Toulmin [Tou58]; y la noción de undercutting defeater introducida por Pollock [Pol87].

- En el Capítulo 5 se abordarán las ideas de Toulmin y Pollock en el contexto de un sistema argumentativo basado en reglas. Se desarrollará un formalismo de *Programación en Lógica Rebatible Extendida (E-DeLP)* que captura las nociones de undercutting defeater y backing, permitiendo expresar argumentos a favor y en contra de reglas rebatibles. Se abordarán diferentes alternativas para el cálculo de aceptabilidad de argumentos en E-DeLP, una de las cuales emplea los desarrollos del Capítulo 4.

- En el Capítulo 6 se presentarán conclusiones y trabajo a futuro.
- En el Apéndice A se incluyen todas las propiedades introducidas a lo largo de esta tesis, junto a sus correspondientes demostraciones.
- En el Apéndice B se listan todos los acrónimos utilizados en esta tesis.

Capítulo 2

Argumentación como Mecanismo de Razonamiento

Argumentación es una forma de razonamiento en la cual, para una determinada afirmación, se presta atención explícita a las justificaciones presentadas y a la resolución de los posibles conflictos entre ellas. En este tipo de razonamiento, una afirmación es aceptada o rechazada en función del análisis de los argumentos a su favor y en su contra. La forma en que los argumentos y las justificaciones para una afirmación son considerados permite definir un mecanismo de razonamiento automático que contempla la posibilidad de contar con información contradictoria, incompleta e incierta [PV02, RS09].

El estudio de argumentación ha sido abordado desde diferentes enfoques: a nivel lógico puede considerarse como una forma de modelar inferencia rebatible, y a nivel dialógico puede verse como una forma de interacción entre agentes inteligentes. Esta forma de razonamiento ha sido ampliamente estudiada en disciplinas como la filosofía [Tou58, Pol87], y desde los años '70 es posible encontrar estudios en las Ciencias de la Computación que han contribuido notablemente a la noción de argumento. A mediados de los años '80 comenzaron los desarrollos del área de argumentación desde un punto de vista computacional, donde los argumentos son explícitamente construidos y comparados como medios para resolver problemas en una computadora.

En las últimas décadas, argumentación ha evolucionado como un atractivo paradigma para conceptualizar el razonamiento de sentido común [CML00, BCD07, BH08, AyC14]. Existen diferentes aproximaciones para modelar argumentación, tales como sistemas de argumentación abstracta [Dun95, Pra09], sistemas que utilizan lógica clásica [BH01, BH09a], o sistemas argumentativos basados en reglas [PS97, GS04, DKT06, AK07]. Por otra parte, este tipo de razonamiento es particularmente atractivo para la toma de decisiones, y dentro del área de Inteligencia Artificial existe especial interés en abordar este tipo de problemas. Concretamente, el proceso argumentativo ha sido empleado en diversos dominios y aplicaciones como el razonamiento legal [PS02, Ver03b], sistemas para la toma de decisiones y negociación [APM00, RRJ$^+$03, BH09b, AV12], y sistemas multi-agente [PSJ98, AMP02], entre muchos otros.

En lo que resta de este capítulo se introducirá una serie de nociones básicas de argumentación, las cuales proveen las bases para los desarrollos de esta tesis. En primer lugar se presentará la estructura conceptual propuesta en [PV02], la cual provee una caracterización de los diferentes elementos presentes en los sistemas argumentativos. Luego, se mostrará cómo estos elementos se ven reflejados en los diferentes pasos involucrados en el proceso de razonamiento argumentativo. Seguidamente, se presentarán dos aproximaciones para la formalización de sistemas argumentativos: los marcos argumentativos abstractos [Dun95] y la programación en lógica rebatible [GS04]. Se introducirán las principales características de estos sistemas, haciendo referencia a cómo instancian los diferentes elementos de la estructura conceptual antes mencionada. Finalmente, partiendo de la propuesta de Toulmin [Tou58], se motivará la existencia de una relación de soporte entre argumentos como mecanismo para ampliar las capacidades de representación y razonamiento de los sistemas argumentativos.

2.1. Una Caracterización de los Sistemas Argumentativos

En [PV02] los autores proponen una estructura conceptual que describe los principales elementos con los que se puede caracterizar a un sistema argumentativo. Se identifican cinco elementos, algunos de los cuales pueden hallarse presentes de manera implícita: un *lenguaje de representación*; una definición de la noción de *argumento*; la identificación de *conflictos* entre argumentos; la obtención de las *derrotas* entre argumentos; y, finalmente, la determinación del *estado de aceptabilidad* de los argumentos, el cual puede ser utilizado para definir una noción de consecuencia lógica del sistema.

- **Lenguaje de Representación:** Los sistemas argumentativos se construyen a partir un lenguaje de representación que es utilizado para modelar la información acerca del dominio sobre el cual se efectuará el razonamiento. Asociada a este lenguaje de representación se define una noción de consecuencia lógica, la cual es utilizada para definir la noción de argumento (segundo elemento de la estructura conceptual). Algunos sistemas utilizan un lenguaje de representación particular (*e. g.*, [GS04]), mientras que otros sistemas dejan la lógica subyacente parcialmente especificada o completamente sin especificar (*e. g.*, [Dun95]). En particular, esta última clase de sistemas puede ser instanciada utilizando diferentes lógicas alternativas, caracterizando así la noción de *marco argumentativo*.

- **Noción de Argumento:** La noción de argumento corresponde a una prueba (o la existencia de una prueba) en la lógica subyacente. En cuanto al formato para la representación de los argumentos, en la literatura se destacan principalmente tres alternativas. La primera de ellas define a los argumentos como árboles de inferencias basados en premisas. Otra alternativa consiste en la construcción de argumentos a partir de una secuencia de prueba o derivación. Finalmente, algunos sistemas simplemente definen un argumento como un par *premisas-conclusión*, dejando implícito el hecho de que el operador de consecuencia asociado a la lógica subyacente permite obtener una prueba de la conclusión a partir de las premisas. Cabe destacar que algunos formalismos argumentativos se abstraen de la forma en que los argumentos son construidos, dejando su estructura interna completamente sin especificar. Tal es el caso del marco argumentativo propuesto por Dung en [Dun95], el cual trata la noción de argumento como una primitiva, centrándose exclusivamente en las interacciones entre los argumentos.

- **Conflictos entre Argumentos:** El uso de argumentación como mecanismo de razonamiento presupone, de alguna manera, la existencia de conflictos entre argumentos. Esta noción de conflicto, también llamada ataque o contra-argumentación, puede tener origen en diferentes situaciones. La literatura de argumentación usualmente distingue tres tipos de conflictos. El primero de ellos, conocido como *ataque por refutación* (en inglés, *rebutting attack*), ocurre cuando dos argumentos poseen conclusiones contradictorias. Otro tipo de conflicto, llamado *ataque a una suposición* (en inglés, *assumption attack*) surge cuando un argumento realiza una suposición de no existencia de prueba (*i. e.*, se asume que no es posible obtener una prueba para una determinada fórmula del lenguaje lógico subyacente) y otro argumento prueba la conclusión que el primero asumió como no demostrable. Existe en la literatura otra forma de conflicto que se encuentra relacionada con el ataque a una suposición: el *ataque a una premisa* (en inglés, *undermining attack*). Este tipo de ataque está dirigido sobre una premisa de un argumento para la cual no existe una prueba fehaciente (*i. e.*, una premisa que no forma parte de los hechos de la base de conocimiento). Por último, Pollock [Pol87] identificó otra situación de conflicto, actualmente conocido como *ataque por socavamiento* (en inglés, *undercutting attack*), en la que un argumento ataca una regla de inferencia de otro argumento. Dados los tipos de conflicto identificados en la literatura, podemos observar que el ataque por refutación es simétrico, mientras que los otros tipos de ataque son asimétricos. Generalmente, los tipos de conflicto arriba mencionados son inferidos a partir de las características del lenguaje de representación. Es por esto que en sistemas abstractos como el propuesto en [Dun95], en los que hay abstracción del lenguaje de representación y de la forma en que los argumentos son construidos, se asume que los conflictos entre argumentos se encuentran establecidos explícitamente de antemano.

Observación 2.1 *Ante la falta de una traducción al idioma español globalmente aceptada por la comunidad de argumentación, en esta tesis se empleará la terminología en inglés correspondiente a los siguientes tipos de conflicto:* ataque por refutación*,* ataque a una suposición*,* ataque a una premisa *y* ataque por socavamiento*. Concretamente, nos referiremos a estos tipos de ataque como* rebutting attack*,* assumption attack*,* undermining attack *y* undercutting attack *respectivamente.*

- **Derrotas entre Argumentos:** La noción de conflicto o ataque entre argumentos no implica evaluación alguna. Luego, es necesaria la definición de una noción de derrota, la cual conlleva la evaluación de pares de argumentos conflictivos para determinar qué ataques son exitosos. De esta manera, se define una relación binaria de derrota que captura la intuición de "ataca y no es más débil" (en su versión débil) o "ataca y es más fuerte" (en su versión fuerte). En particular, cada uno de los tipos de ataque identificados en el tercer elemento de esta estructura conceptual posee su contrapartida como derrota. En la literatura de argumentación existen diferentes alternativas para la resolución de conflictos y la obtención de derrotas. Una de las más populares en el área de Inteligencia Artificial es la utilización de un criterio de comparación entre argumentos llamado *especificidad*, el cual prefiere argumentos basados en información más específica, valiéndose únicamente de la estructura lógica de los argumentos y abstrayéndose de la información del dominio de aplicación. Sin embargo, en trabajos como [Vre91, Pol95, PS96] los autores establecen que el criterio de especificidad no corresponde a un principio general del razonamiento de sentido común, sino que simplemente es uno de muchos estándares que pueden o no ser utilizados. Más aún, propuestas como [Kon88, Vre91] sugieren que la información acerca del dominio suele ser la herramienta principal para evaluar los argumentos del sistema. Es por esto que sistemas argumentativos como [GS04] se encuentran parametrizados con respecto al criterio de comparación entre argumentos, siendo este provisto por el usuario y posiblemente hallándose relacionado con el dominio de aplicación.

- **Estado de Aceptabilidad de los Argumentos:** El objetivo de un sistema argumentativo consiste en determinar qué argumentos son finalmente aceptados. Para esto es necesario considerar las derrotas entre los argumentos del sistema. Sin embargo, no resulta suficiente considerar aisladamente la relación de derrota para cada par de argumentos conflictivos, ya que esto permite identificar únicamente la fuerza relativa de dos argumentos individuales. En contraste, para determinar el estado de aceptación de un argumento, es necesario considerar las interacciones de todos los argumentos disponibles en el sistema. Es decir, dado un argumento, deben considerarse sus derrotadores, los derrotadores de sus derrotadores, y así sucesivamente. Por ejemplo, consideremos un sistema argumentativo con los argumentos $\mathcal{A}, \mathcal{B}$ y $\mathcal{C}$, donde el argumento $\mathcal{A}$ derrota al argumento $\mathcal{B}$ y, por otra parte, el argumento $\mathcal{C}$ derrota al argumento $\mathcal{A}$. En este caso, el argumento $\mathcal{C}$ provee una defensa para $\mathcal{B}$ ante la derrota de $\mathcal{A}$. Esta situación es capturada en la literatura bajo la noción de *reinstalación* (en inglés, *reinstatement*). Luego, los argumentos aceptados del sistema serían $\mathcal{C}$ y $\mathcal{A}$. La definición de un mecanismo para el cómputo del estado de aceptabilidad de los argumentos permite obtener las inferencias de un sistema argumentativo. Típicamente, este mecanismo divide a los argumentos en al menos dos categorías: argumentos *aceptados* y argumentos *rechazados*. Algunas veces también se considera una tercer categoría, correspondiente a argumentos *en discusión*; es decir, argumentos cuyo estado de aceptación es *indeciso* dado que no se encuentran aceptados ni rechazados. En la literatura existen diversas propuestas para el cálculo de aceptabilidad de los argumentos de un sistema argumentativo. En particular, estas alternativas pueden definirse siguiendo una aproximación declarativa o una aproximación procedural. La aproximación declarativa simplemente establece ciertas condiciones que un conjunto de argumentos aceptables debe cumplir, sin especificar cómo se computarán tales conjuntos. Como se verá en la siguiente sección, un ejemplo de esta aproximación está dado por la definición de *semánticas de aceptabilidad* para los marcos argumentativos propuestos en [Dun95]. Por otra parte, las aproximaciones procedurales brindan un procedimiento para determinar si un argumento en particular se encuentra aceptado o no. En general, este tipo de procedimientos simula el juego argumentativo que ocurre en una discusión, y es modelado a través de árboles de argumentos conocidos como *árboles de dialéctica*. De esta manera, podemos considerar que la aproximación declarativa define la semántica de un sistema argumentativo, mientras que la aproximación procedural constituye su teoría de prueba.

A partir de la caracterización propuesta en [PV02] podemos observar cómo los elementos de un sistema argumentativo intervienen en las diferentes etapas del proceso de razonamiento argumentativo. La Figura 2.1 provee una esquematización de los cinco pasos del proceso argumentativo: *i*) construcción de argumentos, *ii*) construcción de interacciones entre argumentos, *iii*) comparación de argumentos, *iv*) cómputo de aceptabilidad de argumentos, y *v*) obtención de conclusiones del sistema. A continuación se provee una descripción de cada uno de estos pasos.

(*i*) Los argumentos son construidos a partir de una base de conocimiento expresada en el lenguaje de representación subyacente. Dependiendo del lenguaje y de las reglas para la construcción de argumentos, es posible obtener diferentes tipos de argumentos.

(*ii*) Los argumentos obtenidos en el paso anterior no pueden ser considerados aisladamente. En efecto, la mayoría de los argumentos de un sistema argumentativo interactúan de alguna manera. Por ejemplo, podemos considerar una relación de conflicto o ataque entre argumento.

(*iii*) La idea detrás de este paso consiste en asignar un peso a cada argumento. Luego, mediante la consideración de estos pesos es posible comparar los argumentos. Existen en la literatura diferentes criterios para asignar pesos a los argumentos.

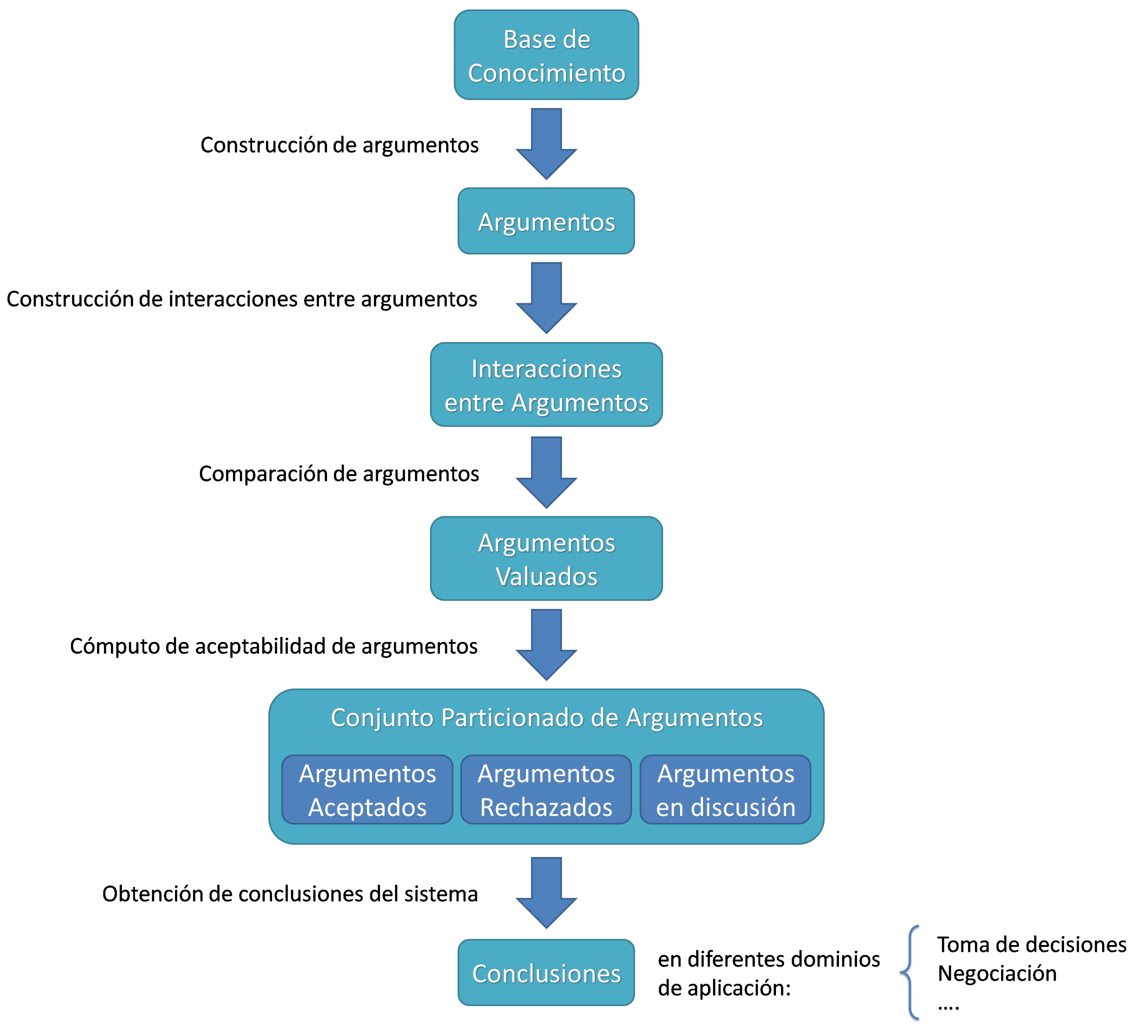


Figura 2.1: Los cinco pasos del proceso de razonamiento argumentativo.

Por ejemplo, puede utilizarse el criterio de especificidad mencionado anteriormente. Otra alternativa, utilizada en [AC02a], consiste en considerar prioridades explícitas e implícitas.

(*iv*) El objetivo de un sistema argumentativo es identificar el conjunto de argumentos que deben ser aceptados, los cuales constituyen los "mejores" argumentos del sistema. El cálculo del estado de aceptabilidad de los argumentos contempla las interacciones entre estos, así como también los valores asignados por el criterio de comparación.

(*v*) El estado de aceptabilidad de los argumentos determina las conclusiones del sistema. De esta manera, es posible identificar un conjunto de inferencias a partir de la selección de los argumentos aceptados. Luego, es posible utilizar estas conclusiones como herramienta de apoyo en diferentes dominios de aplicación, tales como sistemas para la toma de decisiones o negociación.

2.2. Formalización de Sistemas Argumentativos

En esta sección se presentarán dos variantes que constituyen diferentes aproximaciones para la formalización de sistemas argumentativos: un sistema argumentativo abstracto y un Sistema Argumentativo Basado en Reglas (SABR). Los sistemas argumentativos del primer tipo se abstraen de la mayoría de los elementos de la estructura conceptual introducida en la Sección 2.1, para concentrarse en el análisis del estado de aceptación de los argumentos del sistema. En contraste, los SABR definen todos los elementos identificados en la estructura conceptual. Es por esto que, usualmente, la literatura clasifica a los SABR como sistemas argumentativos *concretos*, a diferencia de los sistemas argumentativos *abstractos*, también conocidos como *marcos argumentativos*.

Los SABR cuentan con una base de conocimiento que permite almacenar información expresada en un lenguaje lógico de representación. En particular, existe un conjunto de reglas de inferencia con las cuales, a partir de cierta información (antecedente) es posible inferir nueva información (consecuente). Estas reglas son almacenadas en la base de conocimiento en conjunto con otra información en forma de hechos o presuposiciones, representando evidencia que se obtiene del entorno. A partir de esta evidencia es posible utilizar las reglas de inferencia para construir argumentos a favor o en contra de una afirmación. Luego, mediante un análisis exhaustivo, se evalúan los argumentos construidos para decidir cuáles de ellos son aceptados y, de esta manera, concluir si a partir de la información existente puede garantizarse o no dicha afirmación.

Cabe destacar que los SABR son no monótonos, dado que la incorporación de nueva información al sistema puede generar nuevos argumentos que contradigan a otros ya existentes y, por lo tanto, invalidar afirmaciones que antes se hallaban garantizadas. Es por esto que los SABR permiten representar conocimiento de sentido común y proveen una forma de razonamiento atractiva para aplicaciones de Inteligencia Artificial. En particular, este tipo de sistemas ha resultado de especial interés en sistemas multi-agente, dado que proveen un mecanismo de razonamiento sofisticado para los agentes involucrados en este dominio [RA06, GRTS07].

A continuación se presentarán dos formalismos asociados a diferentes alternativas para la especificación de sistemas argumentativos. Por una parte, se presentará el marco argumentativo abstracto propuesto por Dung en [Dun95], correspondiente a la categoría de sistemas argumentativos abstractos. Por otra parte, se introducirá el formalismo de programación en lógica rebatible propuesto en [GS04], el cual constituye un SABR. A partir de estos formalismos se ilustrarán, respectivamente, las diferentes características de los sistemas argumentativos abstractos y concretos. Asimismo, estos formalismos resultan de especial interés, ya que serán tomados como base para los desarrollos de esta tesis.

2.2.1. Marcos Argumentativos Abstractos (AF)

En esta sección se presentarán las nociones básicas correspondientes a los marcos argumentativos abstractos introducidos por Phan M. Dung en [Dun95]. Como se mencionó anteriormente, el formalismo propuesto por Dung se abstrae de los primeros elementos de la estructura conceptual presentada en la Sección 2.1. Concretamente, los marcos argumentativos de Dung se abstraen del lenguaje lógico de representación, del proceso de construcción de argumentos y de la identificación de los conflictos entre ellos, para concentrarse en el cálculo de aceptabilidad de los argumentos.

Dado el gran nivel de abstracción y simplicidad de estos marcos, en la actualidad son utilizados como base teórica para la gran mayoría de los avances en el área de argumentación. En la literatura existe una gran cantidad de aproximaciones que extienden la teoría de Dung, definiendo nuevas semánticas de aceptabilidad [BCG11], proponiendo marcos argumentativos con preferencias [AC02b], marcos argumentativos bipolares [CLS05], o marcos argumentativos con derrotas recursivas [BCGG09], entre otros. En particular, los *Marcos Argumentativos con Backing y Undercutting* [CGS11a, CGS12b] propuestos en el Capítulo 4 constituyen otra extensión al formalismo propuesto por Dung.

2.2.1.1. Especificación de los Marcos Argumentativos Abstractos

La propuesta de Dung comienza con la definición de un *Marco Argumentativo Abstracto (AF)* (por su nombre en inglés, *Abstract Argumentation Framework*). Básicamente, un AF es un par ordenado donde el primer elemento es un conjunto de argumentos y el segundo elemento es una relación binaria de derrota[1] entre ellos. De esta manera, en [Dun95] se considera que los argumentos son entidades abstractas indivisibles, cuyo único rol está determinado por su relación con otros argumentos.

Definición 2.1 (Marco Argumentativo Abstracto) *Un* Marco Argumentativo Abstracto (AF) *es una tupla* $\langle \mathbb{A}, \mathbb{R} \rangle$*, donde* $\mathbb{A}$ *es un conjunto finito de argumentos y* $\mathbb{R} \subseteq (\mathbb{A} \times \mathbb{A})$ *es una relación de derrota entre argumentos.*

Como se mencionó anteriormente, los argumentos son entidades abstractas y serán denotados utilizando letras mayúsculas caligráficas (*e. g.*, $\mathcal{A}, \mathcal{B}, \ldots, \mathcal{C}_i, \ldots, \mathcal{C}_n$). La relación de derrota entre dos argumentos $\mathcal{A}$ y $\mathcal{B}$ denota el hecho de que estos argumentos no pueden ser aceptados simultáneamente por hallarse en conflicto. Además, diremos que un argumento $\mathcal{A}$ derrota a un argumento $\mathcal{B}$ cuando $(\mathcal{A}, \mathcal{B}) \in \mathbb{R}$, notado como $\mathcal{A} \longrightarrow \mathcal{B}$. Es posible representar gráficamente un AF utilizando un grafo dirigido, conocido como *grafo de derrotas* asociado al AF. De esta manera, los nodos en el grafo de derrotas corresponden a los argumentos del AF, mientras que los arcos denotan los pares de la relación de derrota entre argumentos. Para ilustrar esto, consideremos el siguiente ejemplo.

Ejemplo 2.1 *Sea* $\mathsf{AF}_{2.1} = \langle \mathbb{A}_{2.1}, \mathbb{R}_{2.1} \rangle$ *un marco argumentativo abstracto, donde:*

$$\mathbb{A}_{2.1} = \{\mathcal{A}, \mathcal{B}, \mathcal{C}, \mathcal{D}\} \qquad \mathbb{R}_{2.1} = \{(\mathcal{A}, \mathcal{B}), (\mathcal{B}, \mathcal{A}), (\mathcal{A}, \mathcal{C}), (\mathcal{B}, \mathcal{C}), (\mathcal{C}, \mathcal{D})\}$$

La Figura 2.2 ilustra el grafo de derrotas de $\mathsf{AF}_{2.1}$*. En este caso, los argumentos* $\mathcal{A}$ *y* $\mathcal{B}$ *se derrotan entre sí. Además,* $\mathcal{A}$ *y* $\mathcal{B}$ *derrotan al argumento* $\mathcal{C}$*, mientras que este último derrota al argumento* $\mathcal{D}$*.*

[1] En [Dun95] se utiliza la terminología 'ataque'. Sin embargo, dado que la relación codifica ataques exitosos, en esta tesis nos referiremos a ella como relación de 'derrota', para así evitar confusiones.

$$\begin{tikzcd}[row sep=small] \mathcal{A} \arrow[dr] \arrow[dd, shift right] & & \\ & \mathcal{C} \arrow[r] & \mathcal{D} \\ \mathcal{B} \arrow[ur] \arrow[uu, shift right] & & \end{tikzcd}$$

Figura 2.2: Grafo de derrotas correspondiente al $\mathsf{AF}_{2.1}$ del Ejemplo 2.1.

2.2.1.2. Semánticas de Aceptabilidad para los Marcos Argumentativos Abstractos

Recordemos que el objetivo de un sistema argumentativo es determinar qué argumentos están aceptados para, posteriormente, arribar a diferentes conclusiones. Por este motivo, es necesario considerar las interacciones existentes entre los argumentos, con el objetivo de evitar la obtención de conclusiones inconsistentes. En particular, para analizar la aceptabilidad de los argumentos de un AF es necesario considerar la relación de derrota entre argumentos, para así evitar que argumentos conflictivos queden simultáneamente aceptados. En consecuencia, es necesario definir un mecanismo para calcular la aceptabilidad de los argumentos de un AF de manera tal que para cada argumento se consideren sus derrotadores, los derrotadores de sus derrotadores, y así sucesivamente.

Como se vio en la Sección 2.1, el cálculo de aceptabilidad de los argumentos puede definirse en forma declarativa o procedural. El formalismo de Dung adopta una aproximación declarativa, definiendo una serie de *semánticas de aceptabilidad* para los argumentos de un AF (para una revisión de diferentes semánticas referirse a [BCG11]). En la literatura es posible identificar dos estilos o enfoques comúnmente empleados para definir las semánticas de aceptabilidad. El *enfoque basado en asignaciones de estado* (en inglés, *status assignment approach* o *labelling-based approach*), propuesto originalmente por Pollock [Pol94] y más recientemente utilizado en [Cam06a, Vre06, Cam07, Ver07], consiste en caracterizar una o varias asignaciones de estado para los argumentos, asociando a cada argumento una etiqueta que denota el estado. Este enfoque típicamente utiliza las etiquetas *In*, *Out* y *Undecided*, donde la primera denota que un argumento se encuentra aceptado de acuerdo a la asignación, la segunda que se encuentra decididamente no aceptado, y la tercera denota indecisión con respecto al estado de aceptación del argumento. Por otra parte, en [Dun95] se propone el *enfoque basado en extensiones* o *enfoque extensional* (en inglés, *extension-based approach*). Esta aproximación consiste en caracterizar declarativamente uno o varios conjuntos de argumentos, denominados *extensiones*, que se consideran simultáneamente aceptados de acuerdo a la semántica adoptada.

En particular, para los desarrollos de esta tesis se adoptará un enfoque extensional considerando las cuatro semánticas originalmente propuestas en [Dun95]: *completa*, *preferida*, *estable*, y *grounded*[2]. Estas semánticas dan lugar, respectivamente, a las extensiones completas, preferidas, estables y grounded, las cuales establecen diferentes condiciones que un conjunto de argumentos aceptados debe cumplir. Si bien en la literatura se han propuesto otras semánticas como la *semi-estable* [Cam06b], *ideal* [DMT06], *CF2* [BGG05] o *prudente* [CMDM05], en esta tesis se optó por considerar únicamente las semánticas propuestas por Dung dado que son ampliamente reconocidas y utilizadas en la literatura de argumentación [RS09].

Previo a la definición de las semánticas de aceptabilidad, en [Dun95] se propone una serie de nociones auxiliares que son luego utilizadas para la caracterización de las extensiones de un AF. Concretamente, se definen las nociones de *conjunto libre de conflictos*, *aceptabilidad* y *admisibilidad* presentadas a continuación.

Definición 2.2 (Conjunto Libre de Conflictos, Aceptabilidad y Admisibilidad) *Sea* $\mathsf{AF} = \langle \mathbb{A}, \mathbb{R} \rangle$ *un marco argumentativo abstracto y* $S \subseteq \mathbb{A}$ *un conjunto de argumentos:*

- S *es un* conjunto libre de conflictos *si y sólo si* $\nexists \mathcal{A}, \mathcal{B} \in S$ *tal que* $(\mathcal{A}, \mathcal{B}) \in \mathbb{R}$.
- $\mathcal{A}$ *es* aceptable con respecto a S *si y sólo si* $\forall \mathcal{B} \in \mathbb{A}$*: si* $(\mathcal{B}, \mathcal{A}) \in \mathbb{R}$*, entonces* $\exists \mathcal{C} \in S$ *tal que* $(\mathcal{C}, \mathcal{B}) \in \mathbb{R}$.
- S *es un* conjunto admisible *si y sólo si es libre de conflictos y* $\forall \mathcal{A} \in S$ *se tiene que* $\mathcal{A}$ *es aceptable con respecto a* S.

La noción de libertad de conflictos captura la intuición de que todo conjunto de argumentos aceptados debe satisfacer la restricción de no aceptar simultáneamente argumentos conflictivos. Por otra parte, la noción de aceptabilidad establece que para considerar a un argumento como aceptable, este debe ser defendido ante todas las derrotas que recibe. Luego, estas dos nociones son combinadas para caracterizar a los conjuntos admisibles, los cuales determinan los requerimientos mínimos que deben satisfacer las extensiones de

[2] Ante la falta de una traducción adecuada para este concepto, en esta tesis se utilizará la terminología correspondiente al idioma Inglés.

un AF. De esta manera, un conjunto admisible puede ser interpretado como una posición coherente que se defiende a sí misma. Por ejemplo, dado el marco argumentativo $\mathsf{AF}_{2.1}$ del Ejemplo 2.1, el argumento $\mathcal{D}$ es aceptable con respecto a los conjuntos $\{\mathcal{A}\}$, $\{\mathcal{B}\}$ y $\{\mathcal{A}, \mathcal{B}\}$. Sin embargo, únicamente los dos primeros son conjuntos admisibles; en contraste, el conjunto $\{\mathcal{A}, \mathcal{B}\}$ no es admisible por no ser libre de conflictos, dado que $(\mathcal{A}, \mathcal{B}) \in \mathbb{R}_{2.1}$ y $(\mathcal{B}, \mathcal{A}) \in \mathbb{R}_{2.1}$.

Finalmente, en [Dun95] se define la semántica *completa* que, partiendo de la noción de admisibilidad, caracteriza a las extensiones completas de un AF. De manera similar, luego se definen las semánticas *preferida*, *estable* y *grounded*, las cuales imponen restricciones adicionales sobre las extensiones completas.

Definición 2.3 (Extensiones de un AF) *Sea $AF = \langle \mathbb{A}, \mathbb{R} \rangle$ un marco argumentativo abstracto y $S \subseteq \mathbb{A}$ un conjunto de argumentos libre de conflictos:*

- *S es una* extensión completa *de AF si y sólo si $\forall \mathcal{A} \in \mathbb{A}$: si $\mathcal{A}$ es aceptable con respecto a S, entonces $\mathcal{A} \in S$.*
- *S es una* extensión preferida *de AF si y sólo si es un conjunto admisible maximal con respecto a la inclusión de conjuntos.*
- *S es una* extensión estable *de AF si y sólo si $\forall \mathcal{A} \in \mathbb{A} \backslash S$: $\exists \mathcal{B} \in S$ tal que $(\mathcal{B}, \mathcal{A}) \in \mathbb{R}$.*
- *S es la* extensión grounded *de AF si y sólo si es la extensión completa minimal con respecto a la inclusión de conjuntos.*

La definición anterior caracteriza diferentes tipos de extensiones, comenzando desde las menos restrictivas para luego añadir mayores restricciones. Las extensiones completas son definidas siguiendo una política osada, dado que incluyen pocas restricciones. De esta manera, las extensiones completas corresponden a conjuntos admisibles que incluyen a todos los argumentos que defienden. Luego, con el objetivo de abarcar la mayor cantidad de argumentos posible, las extensiones preferidas corresponden a conjuntos admisibles maximales o, lo que es equivalente, extensiones completas maximales. Por otra parte, las extensiones estables son definidas de manera tal que incluyen las características de las extensiones preferidas. En consecuencia, una extensión estable es una extensión preferida que derrota a todo argumento que no pertenece a ella. En contraste, no toda extensión preferida es una extensión estable. Por último, la extensión grounded se deriva de la semántica de aceptabilidad más restrictiva. Esta extensión es definida siguiendo una política cauta, asegurando la existencia de un único conjunto de argumentos que cumple con las características establecidas. Es por esto que la extensión grounded se corresponde con la extensión completa minimal de un AF. Cabe destacar que, ante la presencia de conflictos que no se pueden resolver en forma directa (ciclos de derrota no resueltos), la semántica preferida considera diferentes alternativas de extensiones, mientras que la semántica grounded opta por dejar los argumentos conflictivos fuera de la extensión. Para ilustrar esto, consideremos los siguientes ejemplos.

Ejemplo 2.2 *Consideremos el marco argumentativo $\mathsf{AF}_{2.1}$ del Ejemplo 2.1. Las extensiones completas de $\mathsf{AF}_{2.1}$ son $\emptyset$, $\{\mathcal{A}, \mathcal{D}\}$ y $\{\mathcal{B}, \mathcal{D}\}$. A partir de esto, las extensiones preferidas de $\mathsf{AF}_{2.1}$ son $\{\mathcal{A}, \mathcal{D}\}$ y $\{\mathcal{B}, \mathcal{D}\}$, por ser las extensiones completas maximales con respecto a la inclusión de conjuntos. De manera similar, la extensión grounded de $\mathsf{AF}_{2.1}$ es $\emptyset$, es decir, la extensión completa minimal. Por último, las extensiones estables de $\mathsf{AF}_{2.1}$ son $\{\mathcal{A}, \mathcal{D}\}$ y $\{\mathcal{B}, \mathcal{D}\}$ que, en particular, coinciden con las extensiones preferidas de $\mathsf{AF}_{2.1}$.*

El Ejemplo 2.2 permite observar la diferencia entre las extensiones preferidas y la extensión grounded de un AF. En particular, esta diferencia surge a partir del ciclo de derrotas que involucra a los argumentos $\mathcal{A}$ y $\mathcal{B}$ de $\mathsf{AF}_{2.1}$. Dado que este ciclo de derrota no está interrumpido, en primera instancia no sería posible decidir sobre la aceptación de uno de los argumentos involucrados en el ciclo por sobre el otro. Luego, la semántica preferida conduce a la obtención de dos extensiones preferidas ($\{\mathcal{A}, \mathcal{D}\}$ y $\{\mathcal{B}, \mathcal{D}\}$), las cuales consideran por separado la aceptación de cada uno de los argumentos involucrados en el ciclo de derrota. Por otra parte, la semántica grounded toma una política más cauta, dejando fuera de la extensión a los dos argumentos involucrados en el ciclo.

Como se mencionó anteriormente, toda extensión estable de un AF es, además, una extensión preferida del AF. Esto se observa en el Ejemplo 2.2, donde las extensiones estables de $\mathsf{AF}_{2.1}$ coinciden con las extensiones preferidas $\{\mathcal{A}, \mathcal{D}\}$ y $\{\mathcal{B}, \mathcal{D}\}$. No obstante, existen otras situaciones en las que las extensiones estables de un AF son un subconjunto estricto de las extensiones preferidas del marco argumentativo. Para ilustrar esto, consideremos el siguiente ejemplo.

Ejemplo 2.3 *Sea $\mathsf{AF}_{2.3} = \langle \mathbb{A}_{2.3}, \mathbb{R}_{2.3} \rangle$ un marco argumentativo abstracto, donde $\mathbb{A}_{2.3} = \{\mathcal{A}\}$ y $\mathbb{R}_{2.3} = \{(\mathcal{A}, \mathcal{A})\}$. En este caso, la única extensión completa de $\mathsf{AF}_{2.3}$ es*

$\emptyset$. En particular, $\emptyset$ es también la única extensión preferida y la extensión grounded de $AF_{2.3}$. Por otra parte, claramente $\emptyset$ no es una extensión estable de $AF_{2.3}$, dado que $\mathcal{A}$ no pertenece a la extensión y no es derrotado por ningún argumento perteneciente a la extensión.

El ejemplo anterior ilustra otra diferencia importante entre las diferentes semánticas propuestas en [Dun95]. La caracterización de las extensiones de un AF provista en la Definición 2.3 es tal que garantiza la existencia de al menos una extensión completa: el conjunto vacío de argumentos. De esta manera, se garantiza la existencia de al menos una extensión preferida, así como también de la extensión grounded de un AF. En contraste, como se muestra en el Ejemplo 2.3, los marcos argumentativos abstractos no necesariamente poseerán extensiones estables.

Como se mencionó en la Sección 2.1, a partir de la selección de los argumentos aceptables de un sistema es posible clasificar a dichos argumentos en (posiblemente) tres categorías: argumentos aceptados, argumentos rechazados, y argumentos en discusión. Por otra parte, como se vio a partir de los ejemplos anteriores, un marco argumentativo puede poseer múltiples extensiones resultantes de la aplicación de una misma semántica de aceptabilidad. De esta manera, los argumentos de un AF no necesariamente pertenecerán a todas las extensiones resultantes de la aplicación de una determinada semántica. Con el objetivo de capturar esta intuición, desdoblaremos la categoría de argumentos aceptados para identificar aquellos argumentos que se encuentran en discusión. En consecuencia, distinguiremos entre argumentos que se encuentran *escépticamente aceptados* y argumentos que se encuentran *crédulamente aceptados*.

Definición 2.4 *Sea $AF = \langle \mathbb{A}, \mathbb{R} \rangle$ un marco argumentativo abstracto, s una semántica de aceptabilidad tal que $\mathsf{s} \in \{$completa, preferida, estable, grounded$\}$ y $\mathcal{A} \in \mathbb{A}$:*

- *El argumento $\mathcal{A}$ está* escépticamente aceptado *con respecto a s si y sólo si pertenece a toda extensión de AF bajo la semántica s.*

- *El argumento $\mathcal{A}$ está* crédulamente aceptado *con respecto a s si y sólo si pertenece a alguna extensión (pero no a todas) de AF bajo la semántica s.*

- *El argumento $\mathcal{A}$ está* rechazado *con respecto a s si y sólo si no pertenece a ninguna extensión de AF bajo la semántica s.*

Ejemplo 2.4 *Consideremos el marco argumentativo $\mathsf{AF}_{2.1}$ del Ejemplo 2.1. Dada la semántica completa, ningún argumento se encuentra escépticamente aceptado; esto se debe a que, como se observó en el Ejemplo 2.2, entre las extensiones completas de $\mathsf{AF}_{2.1}$ se encuentra el conjunto vacío de argumentos. Luego, los argumentos $\mathcal{A}$, $\mathcal{B}$ y $\mathcal{D}$ están crédulamente aceptados con respecto a la semántica completa, mientras que el argumento $\mathcal{C}$ está rechazado. Dadas las semánticas preferida y estable, el argumento $\mathcal{D}$ está escépticamente aceptado y los argumentos $\mathcal{A}$ y $\mathcal{B}$ están crédulamente aceptados, mientras que el argumento $\mathcal{C}$ está rechazado. Finalmente, todo argumento de $\mathsf{AF}_{2.1}$ está rechazado con respecto a la semántica grounded.*

El Ejemplo 2.4 permite observar que todas las semánticas de la Definición 2.3 coinciden en que el argumento $\mathcal{C}$ está rechazado. Por otra parte, cualquiera sea la forma en que se resuelve el ciclo de derrota entre $\mathcal{A}$ y $\mathcal{B}$, se tendrá que el argumento $\mathcal{D}$ está aceptado, lo cual se ve reflejado en el hecho de que $\mathcal{D}$ está escépticamente aceptado con respecto a las semánticas preferida y estable. Luego, las alternativas que consideran la aceptación de los argumentos $\mathcal{A}$ y $\mathcal{B}$ separadamente conducen al hecho de que $\mathcal{A}$ y $\mathcal{B}$ están crédulamente aceptados con respecto a las semánticas preferida y estable. Finalmente, dado que el conflicto entre $\mathcal{A}$ y $\mathcal{B}$ no puede resolverse en forma directa, al considerar la semántica grounded se tiene que todos los argumentos de $\mathsf{AF}_{2.1}$ están rechazados.

2.2.2. Programación en Lógica Rebatible (DeLP)

En esta sección se presentará el formalismo de *Programación en Lógica Rebatible (DeLP)* (por su nombre en inglés, *Defeasible Logic Programming*) introducido en [Gar00, GS04]. Este formalismo es un SABR que combina resultados de programación en lógica y argumentación rebatible, y que ha sido aplicado exitosamente en diferentes dominios de aplicación (ver *e. g.*, [CCS05, RGS07, GCS08, GGS10]).

DeLP define todos los elementos de la estructura conceptual presentada en la Sección 2.1. En particular, DeLP adopta como lenguaje lógico de representación una extensión de la programación en lógica que incorpora la negación fuerte y permite representar información estricta e información rebatible. Esta información es modelada a través de programas lógicos rebatibles, los cuales constituyen una base de conocimiento. Luego, a partir de estos programas es posible construir argumentos, identificar conflictos entre ellos (mediante el uso de negación fuerte), determinar las derrotas entre ellos (mediante el uso de un criterio de comparación) y utilizar procedimientos de prueba dialécticos para determinar cuáles son los argumentos aceptados del sistema.

Un *programa lógico rebatible* está formado por un conjunto de *hechos y reglas estrictas* y un conjunto de *reglas rebatibles*, permitiendo estas últimas la representación de información tentativa. Toda conclusión del programa deberá estar sustentada por algún *argumento* que pueda construirse utilizando las reglas y los hechos del programa. Cuando se utilicen reglas rebatibles para derivar una conclusión C, esta conclusión será tentativa y podrá ser refutada por información que la contradiga. De esta manera, un argumento que sustenta una conclusión C podrá ser atacado por otros *contra-argumentos* que lo contradigan y que posiblemente resulten ser *derrotadores* para el mismo. Similarmente, estos derrotadores podrán a su vez ser atacados, dando así lugar a nuevos derrotadores. Entonces, para decidir cuándo una conclusión C puede aceptarse a partir de un programa lógico rebatible, se realizará un análisis dialéctico considerando los argumentos a su favor y en su contra. Esto conducirá a la construcción de una estructura arbórea de derrotadores conocida como *árbol de dialéctica*. Finalmente, una conclusión C se hallará *garantizada* si existe un argumento para C que sobreviva a todas las derrotas que recibe en su árbol de dialéctica asociado.

A continuación se introducirán conceptos preliminares de la programación en lógica que son utilizados tanto por DeLP como por el SABR que se presentará en el Capítulo 5. Seguidamente, se indicará cómo DeLP instancia los elementos de la estructura conceptual presentada en la Sección 2.1. Concretamente, se introducirá la sintaxis del lenguaje de representación de los programas lógicos rebatibles, se indicará cómo se efectúa la construcción de argumentos, la identificación de conflictos entre argumentos y la resolución de los mismos para obtener las derrotas correspondientes y, finalmente, se presentará un procedimiento de prueba dialéctico que permite inferir las conclusiones garantizadas a partir de un programa DeLP.

2.2.2.1. Conceptos Preliminares

En esta sección se introducirán conceptos y terminología estándar de la lógica y la programación en lógica que resultan necesarios para la formalización de DeLP y del SABR desarrollado en el Capítulo 5 de esta tesis.

Definición 2.5 (Signatura) *Una* signatura *es una tupla $\sigma = \langle \mathcal{V}, Func, Pred \rangle$, donde $\mathcal{V}$, $Func$, y $Pred$ son conjuntos finitos que representan variables, funciones y predicados respectivamente.*

Dada una signatura σ, una función llamada "*aridad*" le asignará a cada elemento de $Func$ y $Pred$ un número entero positivo. Si $f \in Func$ y $aridad(f) = 0$, entonces f se denominará *constante*. Por otra parte, si $p \in Pred$, y $aridad(p) = 0$, entonces p se denominará *proposición*.

Definición 2.6 (Alfabeto) *El* alfabeto *generado a partir de una signatura σ consiste del conjunto de elementos miembros de la signatura, el símbolo de negación "$\sim$", y los símbolos de puntuación "(", ")", y ",".*

Definición 2.7 (Término) *Sea $\sigma = \langle \mathcal{V}, Func, Pred \rangle$ una signatura. Un* término *T, y respectivamente sus componentes $comp(T)$, se definen inductivamente como sigue:*

- *toda variable $V \in \mathcal{V}$ es un término, y $comp(V) = \{V\}$;*
- *toda constante $c \in Func$ (i. e., $aridad(c) = 0$) es un término, y $comp(c) = \{c\}$; y*
- *si $f \in Func$ tal que $aridad(f) = n$ ($n \geq 1$) y $t_1, \ldots, t_n$ son términos, entonces $f(t_1, \ldots, t_n)$ también es un término.*
  Luego, $comp(f(t_1, \ldots, t_n)) = \{f\} \cup (\bigcup_{1 \leq i \leq n} comp(f_i))$.

Definición 2.8 (Término fijo) *Sea $\sigma = \langle \mathcal{V}, Func, Pred \rangle$ una signatura y T un término. T será un* término fijo *si no contiene variables; es decir, si $comp(T) \cap \mathcal{V} = \emptyset$.*

Definición 2.9 (Átomo) *Sea $\sigma = \langle \mathcal{V}, Func, Pred \rangle$ una signatura y sean $t_1, \ldots, t_n$ términos. Si $p \in Pred$ tal que $aridad(p) = n$, entonces $p(t_1, \ldots, t_n)$ es un* átomo*. Diremos que un* átomo fijo *es un átomo $p(t_1, \ldots, t_n)$ tal que todos sus términos $t_1, \ldots, t_n$ son fijos.*

Definición 2.10 (Literal) *Un* literal *L es un átomo "A" o un átomo negado "$\sim A$", donde "$\sim$" representa la negación fuerte. Un literal L se dirá negativo si es un átomo negado, y positivo en caso contrario. Un* literal fijo *es un átomo fijo o un átomo fijo negado.*

La definición anterior establece que los literales podrán ser átomos negados utilizando el símbolo "∼" de la *negación fuerte* (en inglés, *strong negation*)[3]. Por una parte, este tipo de negación difiere de la negación utilizada en la lógica clásica ya que, por ejemplo, no permite anidamiento (*i. e.*, "∼∼a" no es un literal válido). Por otra parte, la negación fuerte también difiere de la negación por falla (o negación *default*) utilizada en la programación en lógica, la cual es usualmente denotada con el símbolo "not".

Definición 2.11 (Complemento de un Literal) *Sea L un literal y A un átomo. El* complemento *de L con respecto a la negación fuerte, denotado $\overline{L}$, se define de la siguiente manera:*

- *si $L = A$, entonces $\overline{L} = \sim A$;*
- *si $L = \sim A$, entonces $\overline{L} = A$.*

2.2.2.2. Representación de Conocimiento en DeLP: programas lógicos rebatibles

Un programa lógico rebatible estará compuesto por *hechos*, *reglas estrictas* y *reglas rebatibles*, cuya sintaxis se define a continuación.

Definición 2.12 (Hecho) *Un* hecho *es un literal fijo L, es decir, un átomo fijo o un átomo fijo negado.*

Definición 2.13 (Regla Estricta) *Una* regla estricta *es un par ordenado, denotado "Cabeza ← Cuerpo", donde el primer elemento (Cabeza) es un literal fijo y el segundo elemento (Cuerpo) es un conjunto finito y no vacío de literales fijos.*
Una regla estricta con cabeza L_0 y cuerpo $\{L_1, \ldots, L_n\}$ $(n > 0)$ se escribirá también como $L_0 \leftarrow L_1, \ldots, L_n$.

Definición 2.14 (Regla Rebatible) *Una* regla rebatible *es un par ordenado, denotado "Cabeza —≺ Cuerpo", donde el primer elemento (Cabeza) es un literal fijo y el segundo elemento (Cuerpo) es un conjunto finito de literales fijos.*
Una regla rebatible con cabeza L_0 y cuerpo $\{L_1, \ldots, L_n\}$ $(n \geq 0)$ se escribirá también como $L_0 \prec L_1, \ldots, L_n$.

[3]La definición y propiedades de la negación fuerte pueden hallarse en [AP93a, AP93b].

Sintácticamente, el símbolo"$\prec$" es lo único que distingue a una regla rebatible de una regla estricta. Sin embargo, el tipo de conexión expresada por estos dos tipos de reglas es diferente. Por una parte, los hechos y reglas estrictas representan conocimiento seguro y libre de excepciones. En particular, dada una regla estricta, siempre que se crea en los literales que conforman el cuerpo se podrá creer con la misma seguridad en el literal correspondiente a la cabeza. Por otra parte, las reglas rebatibles se utilizan para representar conocimiento rebatible, es decir, información tentativa que puede utilizarse en la medida que no exista información que la contradiga. De esta manera, una regla rebatible "$Cabeza \prec Cuerpo$" expresa que "razones para creer en $Cuerpo$ proveen razones para creer en $Cabeza$". Teniendo en cuenta esto, conocimiento como "*las gallinas son aves*" se representará utilizando la regla estricta:

$$ave \leftarrow gallina$$

mientras que la información "*las aves usualmente vuelan*" se representará mediante la regla rebatible:

$$vuela \prec ave$$

Por último, otra diferencia entre las reglas estrictas y las reglas rebatibles es que estas últimas pueden tener un conjunto vacío de literales en el cuerpo. Una regla rebatible de este tipo es denotada como "$Cabeza \prec$" y recibe el nombre de *presuposición* [Nut88]. De esta manera, una presuposición "$P \prec$" expresa que "existen razones tentativas para creer en P".

A partir de estos elementos, se define un *programa lógico rebatible* de la siguiente manera:

Definición 2.15 (Programa Lógico Rebatible) *Un* programa lógico rebatible *es un conjunto $\mathcal{P}$ de hechos, reglas estrictas y reglas rebatibles. Dado un programa $\mathcal{P}$, se identificará con Θ al conjunto de hechos, con Ω al conjunto de reglas estrictas, y con Δ al conjunto de reglas rebatibles. Por conveniencia, también se denotará con Π al conjunto $\Theta \cup \Omega$. Por lo tanto, en ocasiones se denotará un programa lógico rebatible $\mathcal{P}$ con el par $\mathcal{P} = (\Pi, \Delta)$, mientras que cuando se quieran distinguir los hechos y reglas estrictas del programa se utilizará la terna $\mathcal{P} = (\Theta, \Omega, \Delta)$.*

La definición anterior establece que los programas lógicos rebatibles utilizan únicamente literales fijos. Por este motivo, salvo que se indique lo contrario, cuando hablemos de un "literal" asumiremos que es un literal fijo. No obstante esto, en algunos ejemplos se utilizarán variables en las reglas estrictas o rebatibles, denotando *esquemas de reglas*. Dado un esquema de una regla R, se define $Ground(R)$ como el conjunto de todas las instancias de R con literales fijos. Una instancia fija para R es una versión de la regla tal que las variables de R son reemplazadas por términos fijos, asumiendo que las variables con el mismo nombre corresponden al mismo elemento dentro de la regla. De esta manera, dado un programa lógico rebatible $\mathcal{P}$ se define:

$$Ground(\mathcal{P}) = \bigcup_{R \in \mathcal{P}} Ground(R)$$

En consecuencia, dado un programa lógico rebatible $\mathcal{P}$ con esquemas de reglas, el operador $Ground(\mathcal{P})$ permite obtener todas las instancias fijas de reglas asociadas a $\mathcal{P}$. Luego, para diferenciar las variables de los demás elementos de un programa lógico rebatible, estas serán denotadas con una letra mayúscula inicial. Para ilustrar estas nociones, consideremos el siguiente ejemplo.

Ejemplo 2.5 *Consideremos el programa lógico rebatible* $\mathcal{P}_{2.5} = (\Pi_{2.5}, \Delta_{2.5})$ *que expresa información acerca de aves, donde*

$$\Pi_{2.5} = \left\{ \begin{array}{l} gallina(tina) \\ asustada(tina) \\ paloma(petete) \\ \sim vuela(petete) \\ ave(X) \leftarrow gallina(X) \\ ave(X) \leftarrow paloma(X) \end{array} \right\} \Delta_{2.5} = \left\{ \begin{array}{l} vuela(X) \prec ave(X) \\ \sim vuela(X) \prec gallina(X) \\ vuela(X) \prec gallina(X), asustada(X) \end{array} \right\}$$

En este caso, los hechos gallina(tina), asustada(tina), paloma(petete) y ∼vuela(petete) representan, respectivamente, que "Tina" es una gallina que está asustada y que "Petete" es una paloma que no vuela. De manera similar, las reglas estrictas expresan que tanto las gallinas como las palomas son aves. Por otra parte, las reglas rebatibles expresan las siguientes conexiones tentativas: si X *es un ave, entonces existen razones para creer que* X *vuela; si* X *es una gallina, entonces existen razones para creer que* X *no vuela; y si* X *es una gallina que está asustada, entonces existen razones para creer que* X *vuela. En particular, podemos obtener las siguientes instancias fijas de reglas de* $\mathcal{P}_{2.5}$*, entre otras: "*$ave(tina) \leftarrow gallina(tina)$*", "*$ave(petete) \leftarrow gallina(petete)$*", "*$vuela(petete) \prec ave(petete)$*" y "*$\sim vuela(tina) \prec gallina(tina)$*".*

A continuación se introduce la noción de *derivación rebatible*, la cual define qué literales pueden ser derivados utilizando las reglas de un programa lógico rebatible.

Definición 2.16 (Derivación Rebatible de un Literal) *Sea $\mathcal{P} = (\Pi, \Delta)$ un programa lógico rebatible y L un literal. Una* derivación rebatible *para L a partir de $\mathcal{P}$ es una secuencia finita de literales fijos $L_1, L_2, \ldots, L_n = L$ tal que cada literal L_i pertenece a la secuencia porque:*

(a) *L_i es un hecho en Π;*

(b) *existe en Π una regla estricta R con cabeza L_i y cuerpo $B_1, \ldots, B_k$, donde todo literal B_j del cuerpo $(1 \leq j \leq k)$ es un elemento de la secuencia que precede a L_i; o*

(c) *existe en Δ una regla rebatible R con cabeza L_i y cuerpo $B_1, \ldots, B_k$, donde todo literal B_j del cuerpo $(0 \leq j \leq k)$ es un elemento de la secuencia que precede a L_i.*

La derivación se dice rebatible porque aunque un literal L pueda ser derivado a partir de un programa lógico rebatible, puede existir en el programa información que contradiga a L. Para ilustrar esto consideremos el siguiente ejemplo.

Ejemplo 2.6 *Sea $\mathcal{P}_{2.5}$ el programa lógico rebatible del Ejemplo 2.5. A partir de este programa es posible obtener, entre otras, las siguientes derivaciones rebatibles:*

- *La secuencia $paloma(petete)$, $ave(petete)$, $vuela(petete)$ es una derivación rebatible para el literal "$vuela(petete)$".*
- *La secuencia $gallina(tina)$, $ave(tina)$, $vuela(tina)$ es una derivación rebatible para el literal "$vuela(tina)$".*
- *La secuencia $gallina(tina)$, $\sim vuela(tina)$ es una derivación rebatible para el literal "$\sim vuela(tina)$".*
- *La secuencia $gallina(tina)$, $asustada(tina)$, $vuela(tina)$ es una derivación rebatible para el literal "$vuela(tina)$".*

El Ejemplo 2.6 muestra que a partir de un programa lógico rebatible es posible obtener derivaciones rebatibles para un literal y su complemento, como en el caso de los literales "$vuela(tina)$" y "$\sim vuela(tina)$". Más aún, es posible obtener más de una derivación rebatible para un literal. Tal es el caso de la segunda y cuarta derivación del Ejemplo 2.6, las cuales constituyen derivaciones rebatibles para el literal "$vuela(tina)$". Por otra parte, también pueden existir derivaciones que utilizan sólo reglas rebatibles o sólo reglas estrictas. En particular, como se verá más adelante, este último tipo de derivación resulta de especial interés, y se encuentra caracterizada por la siguiente definición.

Definición 2.17 (Derivación Estricta de un Literal) *Sea $\mathcal{P} = (\Pi, \Delta)$ un programa lógico rebatible y L un literal para el cual existe una derivación rebatible $L_1, L_2, \ldots, L_n = L$. Diremos que L tiene una* derivación estricta *si todas las reglas de $\mathcal{P}$ utilizadas para obtener la secuencia $L_1, L_2, \ldots, L_n = L$ son reglas estrictas.*

Ejemplo 2.7 *Dado el programa lógico rebatible $\mathcal{P}_{2.5}$ del Ejemplo 2.5 podemos obtener, entre otras, las siguientes derivaciones estrictas:*

- *La secuencia $paloma(petete)$, $ave(petete)$ es una derivación estricta para el literal "$ave(petete)$".*
- *La secuencia $gallina(tina)$ es una derivación estricta para el literal "$gallina(tina)$".*

El ejemplo anterior ilustra que, en particular, el conjunto de reglas estrictas utilizado en una derivación rebatible puede ser vacío. Por lo tanto, para todo literal que constituye un hecho del programa existirá una derivación estricta. Como se verá más adelante, las derivaciones estrictas sustentan conclusiones que no podrán ser refutadas, ya que se basan en información segura.

Dado que, como se mencionó anteriormente, a partir de un programa lógico rebatible es posible obtener derivaciones para literales complementarios, a continuación identificaremos los *conjuntos contradictorios.*

Definición 2.18 (Conjunto Contradictorio) *Dado un programa lógico rebatible $\mathcal{P}$, diremos que un subconjunto C de $\mathcal{P}$ es* contradictorio *si y sólo si a partir de C es posible obtener derivaciones (estrictas o rebatibles) para un literal L y su complemento $\overline{L}$.*

Por ejemplo, si consideramos el programa lógico rebatible $\mathcal{P}_{2.5}$ del Ejemplo 2.5, el conjunto $\Pi_{2.5}$ no es contradictorio, mientras que el conjunto $\Pi_{2.5} \cup \Delta_{2.5}$ es contradictorio ya que, como se observó en el Ejemplo 2.6, permite obtener derivaciones rebatibles para los literales "$vuela(tina)$" y "$\sim vuela(tina)$".

Observación 2.2 *El uso de la negación fuerte enriquece la expresividad del lenguaje y permite la representación de información contradictoria. Como las reglas rebatibles permiten expresar información tentativa, dado un programa lógico rebatible* $\mathcal{P} = (\Pi, \Delta)$*, generalmente ocurrirá que el conjunto* $\Pi \cup \Delta$ *es contradictorio. Sin embargo, el conjunto* Π *de hechos y reglas estrictas es utilizado para representar información segura y, por lo tanto, debe poseer cierta coherencia interna. En consecuencia, para todo programa lógico rebatible* $\mathcal{P} = (\Pi, \Delta)$ *se asumirá que el conjunto* Π *es no contradictorio.*

La observación anterior establece que los programas lógicos rebatibles permiten expresar información potencialmente contradictoria mediante el uso de reglas rebatibles. Esto es, al añadir conocimiento rebatible al conjunto de hechos y reglas estrictas de un programa, es posible obtener derivaciones rebatibles para un literal y su complemento. Esto sugiere que la noción de derivación rebatible no resulta conveniente como procedimiento de prueba para un literal de un programa lógico rebatible. Por lo tanto, para decidir si un literal L está aceptado a partir de un programa, se realizará un análisis global más profundo considerando toda la información relevante del mismo.

En particular, DeLP define un procedimiento de prueba para decidir qué literales serán aceptados a partir de un programa lógico rebatible. Este procedimiento de prueba está basado en el formalismo de argumentación rebatible y permitirá definir *argumentos* para un literal, hallar los *contra-argumentos* para esos argumentos, y luego determinar las *derrotas* entre argumentos. Finalmente, un análisis dialéctico permitirá determinar cuándo un literal de un programa lógico rebatible está garantizado.

2.2.2.3. Construcción de Argumentos en DeLP

En DeLP se define la noción de *argumento* para un literal de un programa lógico rebatible, la cual se corresponde con el segundo elemento de la estructura conceptual presentada en la Sección 2.1.

Definición 2.19 (Argumento) *Sea $\mathcal{P} = (\Pi, \Delta)$ un programa lógico rebatible y h un literal. Un* argumento *para h es un par $\langle \mathcal{A}, h \rangle$, donde $\mathcal{A}$ es un conjunto de reglas rebatibles de Δ tal que:*

1. *existe una derivación rebatible para h a partir de $\Pi \cup \mathcal{A}$;*

2. *$\Pi \cup \mathcal{A}$ es un conjunto no contradictorio; y*

3. *$\mathcal{A}$ es minimal: $\nexists \mathcal{A}' \subset \mathcal{A}$ tal que $\mathcal{A}'$ satisface las condiciones (1) y (2).*

Intuitivamente, un argumento $\mathcal{A}$ para un literal h se caracteriza por un conjunto minimal y no contradictorio de reglas rebatibles que permiten obtener una derivación rebatible para h. Además, el literal h es identificado como la conclusión del argumento.

Ejemplo 2.8 *A partir del programa lógico rebatible $\mathcal{P}_{2.5}$ del Ejemplo 2.5 es posible construir los siguientes argumentos:*

- $\langle \mathcal{A}, vuela(tina) \rangle$, *con* $\mathcal{A} = \{vuela(tina) \prec ave(tina)\}$.

- $\langle \mathcal{B}, {\sim}vuela(tina) \rangle$, *con* $\mathcal{B} = \{{\sim}vuela(tina) \prec gallina(tina)\}$.

- $\langle \mathcal{C}, vuela(tina) \rangle$, *con* $\mathcal{C} = \{vuela(tina) \prec gallina(tina), asustada(tina)\}$.

Además, es posible construir otros argumentos como $\langle \emptyset, gallina(tina) \rangle$, $\langle \emptyset, ave(petete) \rangle$ y $\langle \emptyset, {\sim}vuela(petete) \rangle$. Por otra parte, el Ejemplo 2.6 muestra que existe una derivación rebatible para el literal "vuela(petete)". Sin embargo, no es posible obtener un argumento para dicho literal ya que la unión del conjunto de reglas rebatibles utilizadas en su derivación ($\{vuela(petete) \prec ave(petete)\}$) y el conjunto $\Pi_{2.5}$ conduce a un conjunto contradictorio dado que permite derivar los literales "vuela(petete)" y "∼vuela(petete)".

El ejemplo anterior permite observar que en DeLP es posible construir argumentos cuyo conjunto de reglas rebatibles es vacío. En particular, la conclusión de este tipo de argumentos es obtenida a partir de una derivación estricta, es decir, una derivación que utiliza solamente hechos y reglas estrictas del programa. Por lo tanto, en esta tesis nos referiremos a dichos argumentos como *argumentos estrictos*. De esta manera, DeLP asegura que siempre existirá un argumento estricto para cada literal obtenido a partir de una derivación estricta. En consecuencia, si existe un argumento estricto para un literal h, entonces no existirán argumentos para $\overline{h}$. Esta situación ocurre en el Ejemplo 2.8, donde existe un argumento estricto para el literal "$\sim vuela(petete)$", mientras que no existen argumentos para el literal "$vuela(petete)$".

Teniendo en cuenta que un argumento constituye un conjunto de reglas rebatibles, es posible definir la noción de *sub-argumento*. Básicamente, un sub-argumento es un argumento cuyo conjunto de reglas rebatibles está contenido en otro argumento.

Definición 2.20 (Sub-argumento) *Sea $\mathcal{P}$ un programa lógico rebatible y sean $\langle \mathcal{A}, h \rangle$, $\langle \mathcal{B}, q \rangle$ dos argumentos construidos a partir de $\mathcal{P}$. Diremos que $\langle \mathcal{B}, q \rangle$ es un* sub-argumento *de $\langle \mathcal{A}, h \rangle$ (análogamente, $\langle \mathcal{A}, h \rangle$ es un* super-argumento *de $\langle \mathcal{B}, q \rangle$) si y sólo si $\mathcal{B} \subseteq \mathcal{A}$.*

La definición anterior establece que todo argumento es, en particular, sub-argumento de sí mismo. Además, dado que los argumentos estrictos poseen un conjunto vacío de reglas rebatibles, estos serán sub-argumentos de cualquier otro argumento construido a partir de un programa DeLP. Si consideramos el Ejemplo 2.8, tenemos que, por ejemplo, los argumentos $\langle \emptyset, gallina(tina) \rangle$ y $\langle \emptyset, \sim vuela(petete) \rangle$ son sub-argumentos de $\langle \mathcal{A}, vuela(tina) \rangle$, $\langle \mathcal{B}, \sim vuela(tina) \rangle$ y $\langle \mathcal{C}, vuela(tina) \rangle$. En contraste, los argumentos $\langle \mathcal{A}, vuela(tina) \rangle$, $\langle \mathcal{B}, \sim vuela(tina) \rangle$ y $\langle \mathcal{C}, vuela(tina) \rangle$ no poseen super-argumentos. Para ilustrar otra situación que refleja la existencia de sub-argumentos, consideremos el siguiente ejemplo.

Ejemplo 2.9 *Sea $\mathcal{P}_{2.9} = (\Pi_{2.9}, \Delta_{2.9})$ el siguiente programa lógico rebatible:*

$$\Pi_{2.9} = \left\{ \begin{array}{l} buenPrecio(acme) \\ enFusion(acme) \end{array} \right\}$$

$$\Delta_{2.9} = \left\{ \begin{array}{l} comprarAcciones(E) \prec buenPrecio(E) \\ \sim comprarAcciones(E) \prec buenPrecio(E), empresaRiesgosa(E) \\ empresaRiesgosa(E) \prec enFusion(E) \end{array} \right\}$$

A partir de $\mathcal{P}_{2.9}$ es posible construir los argumentos $\langle \mathcal{D}, comprarAcciones(acme) \rangle$, $\langle \mathcal{E}, empresaRiesgosa(acme) \rangle$ y $\langle \mathcal{F}, \sim comprarAcciones(acme) \rangle$, donde

$$\mathcal{D} = \left\{ \; comprarAcciones(acme) \prec buenPrecio(acme) \; \right\}$$

$$\mathcal{E} = \left\{ \; empresaRiesgosa(acme) \prec enFusion(acme) \; \right\}$$

$$\mathcal{F} = \left\{ \begin{array}{l} \sim comprarAcciones(acme) \prec buenPrecio(acme), empresaRiesgosa(acme) \\ empresaRiesgosa(acme) \prec enFusion(acme) \end{array} \right\}$$

En este caso, podemos observar que el argumento $\langle \mathcal{E}, empresaRiesgosa(acme) \rangle$ *es un sub-argumento de* $\langle \mathcal{F}, \sim comprarAcciones(acme) \rangle$.

2.2.2.4. Identificación de Conflictos: ataques entre argumentos de DeLP

Como se mencionó anteriormente, el tercer elemento de la estructura conceptual introducida en la Sección 2.1 corresponde a la identificación de conflictos entre los argumentos del sistema. Dada la posibilidad de utilizar la negación fuerte "∼", es posible representar literales complementarios en un programa DeLP. Más aún, es posible obtener derivaciones para un par de literales complementarios, e incluso construir argumentos para estos literales. A continuación introduciremos la noción de *desacuerdo*, la cual generaliza la noción de literales complementarios para un programa DeLP.

Definición 2.21 (Literales en Desacuerdo) *Sea* $\mathcal{P} = (\Pi, \Delta)$ *un programa lógico rebatible. Diremos que dos literales* h_1 *y* h_2 *están en* desacuerdo *si y sólo si el conjunto* $\Pi \cup \{h_1, h_2\}$ *es contradictorio.*

Claramente, cualquier par de literales complementarios estará en desacuerdo. Por ejemplo, los literales "p" y "$\sim p$" están en desacuerdo ya que el conjunto $\{p, \sim p\}$ es contradictorio de por sí, sin necesidad de tomar en cuenta el conjunto Π de hechos y reglas estrictas de un programa lógico rebatible. Sin embargo, como muestra el siguiente ejemplo, dos literales no complementarios también pueden estar en desacuerdo.

Ejemplo 2.10 *Sea* $\mathcal{P}_{2.10} = (\Pi_{2.10}, \emptyset)$ *un programa lógico rebatible, donde* $\Pi_{2.10} = \{(b), (a \leftarrow b, p), (\sim a \leftarrow c), (c \leftarrow q)\}$[4]. *Los literales "p" y "q" están en desacuerdo dado que* $\{p, q\} \cup \Pi_{2.10}$ *es un conjunto contradictorio que, en particular, permite obtener derivaciones estrictas para los literales complementarios "a" y "∼a":*

- *la secuencia p, b, a constituye una derivación para el literal "a"; y*

[4]Para una mayor claridad en la notación, en algunos casos se utilizarán paréntesis para distinguir los elementos dentro de un conjunto.

- *la secuencia q, c, $\sim a$ constituye una derivación para el literal "$\sim a$".*

A partir de la noción de desacuerdo es posible caracterizar los conflictos entre argumentos. De esta manera, dos argumentos se atacarán cuando sus conclusiones estén en desacuerdo. Más aún, un argumento atacará a otro si sustenta una conclusión que está en desacuerdo con la conclusión de alguno de los sub-argumentos del otro argumento. Esta intuición es capturada por la siguiente definición.

Definición 2.22 (Contra-argumento o Ataque) *Sean $\langle \mathcal{A}, p\rangle$ y $\langle \mathcal{B}, q\rangle$ dos argumentos construidos a partir de un programa lógico rebatible $\mathcal{P}$. El argumento $\langle \mathcal{A}, p\rangle$* contra-argumenta *o* ataca *al argumento $\langle \mathcal{B}, q\rangle$ en el literal h si y sólo si existe un sub-argumento $\langle \mathcal{C}, h\rangle$ de $\langle \mathcal{B}, q\rangle$ tal que los literales p y h están en desacuerdo. El argumento $\langle \mathcal{C}, h\rangle$ se llama* sub-argumento de desacuerdo*, y el literal h es considerado el* punto de contra-argumentación*.*

Ejemplo 2.11 *Consideremos los argumentos ilustrados en el Ejemplo 2.8. En tal caso, el argumento $\langle \mathcal{B}, \sim vuela(tina)\rangle$ es un contra-argumento para los argumentos $\langle \mathcal{A}, vuela(tina)\rangle$ y $\langle \mathcal{C}, vuela(tina)\rangle$. En particular, los sub-argumentos de desacuerdo son, respectivamente, $\langle \mathcal{A}, vuela(tina)\rangle$ y $\langle \mathcal{C}, vuela(tina)\rangle$. Además, el punto de contra-argumentación en ambos casos es el literal "vuela(tina)". Por otra parte, recíprocamente, $\langle \mathcal{A}, vuela(tina)\rangle$ y $\langle \mathcal{C}, vuela(tina)\rangle$ son contra-argumentos para $\langle \mathcal{B}, \sim vuela(tina)\rangle$.*

A partir de la Definición 2.22 puede ocurrir que, como se muestra en el Ejemplo 2.11, el sub-argumento $\langle \mathcal{C}, h\rangle$ sea el propio $\langle \mathcal{B}, q\rangle$. En tal caso diremos que $\langle \mathcal{A}, p\rangle$ contra-argumenta o ataca *directamente* a $\langle \mathcal{B}, q\rangle$. Por otra parte, cuando $\langle \mathcal{C}, h\rangle$ es un sub-argumento propio de $\langle \mathcal{B}, q\rangle$, el ataque se produce en un punto interno de $\langle \mathcal{B}, q\rangle$ (el punto de contra-argumentación h), en cuyo caso diremos que $\langle \mathcal{A}, p\rangle$ contra-argumenta o ataca *internamente* a $\langle \mathcal{B}, q\rangle$.

Nótese que los argumentos estrictos nunca tendrán contra-argumentos. Esto se debe a que, como se mencionó anteriormente, dado un argumento estricto para un literal h no será posible construir argumentos para literales en desacuerdo con h, ya que estos últimos estarían en contradicción con la parte estricta del programa que deriva h.

2.2.2.5. Resolución de Conflictos: derrotas entre argumentos de DeLP

La noción de contra-argumentación establece cuándo un argumento es atacado por otro, capturando la noción de conflicto entre argumentos a través del uso de la negación fuerte. Sin embargo, siempre que un argumento $\mathcal{A}$ ataca a otro argumento $\mathcal{B}$, existe un sub-argumento $\mathcal{C}$ de $\mathcal{B}$ que es un contra-argumento para $\mathcal{A}$. Dada esta simetría, la noción de ataque no permite identificar cuándo un argumento se impone por sobre el otro. Por este motivo es necesario contar con un mecanismo de evaluación sobre los argumentos que establezca cuándo un argumento es mejor que otro, para así determinar cuál de ellos prevalece ante un conflicto. Este mecanismo involucra el uso de un *criterio de comparación entre argumentos* que permite determinar el cuarto elemento de la estructura conceptual presentada en la Sección 2.1: las *derrotas* entre argumentos.

DeLP utiliza un criterio de comparación modular[5] para determinar cuándo un argumento es tanto o más preferido que otro argumento. Por lo tanto, como la comparación entre argumentos puede definirse de diversas formas, en lo que resta de esta sección nos abstraeremos de la especificación de este criterio, asumiendo que se trata de un orden parcial entre argumentos denotado "$\preceq$". De esta manera, utilizaremos $\mathcal{A} \preceq \mathcal{B}$ para denotar que $\mathcal{A}$ es al menos tan preferido como $\mathcal{B}$. De esta manera, siguiendo la notación usualmente adoptada en la literatura, escribiremos $\mathcal{A} \prec \mathcal{B}$ cuando $\mathcal{A} \preceq \mathcal{B}$ y $\mathcal{B} \not\preceq \mathcal{A}$.

Si $\langle \mathcal{A}, p \rangle$ es un contra-argumento para $\langle \mathcal{B}, q \rangle$, es necesario analizar si el ataque de $\langle \mathcal{A}, p \rangle$ sobre $\langle \mathcal{B}, q \rangle$ es lo suficientemente fuerte como para derrotar a $\langle \mathcal{B}, q \rangle$. En consecuencia, es necesario efectuar una comparación entre el argumento atacante y el sub-argumento atacado. Concretamente, si $\langle \mathcal{A}, p \rangle$ ataca a $\langle \mathcal{B}, q \rangle$ en el punto h, siendo $\langle \mathcal{C}, h \rangle$ el sub-argumento de desacuerdo de $\langle \mathcal{B}, q \rangle$, se debe comparar el argumento atacante $\langle \mathcal{A}, p \rangle$ y el sub-argumento de desacuerdo $\langle \mathcal{C}, h \rangle$.

Definición 2.23 (Derrota) *Sean $\langle \mathcal{A}, p \rangle$ y $\langle \mathcal{B}, q \rangle$ dos argumentos construidos a partir de un programa lógico rebatible $\mathcal{P}$. El argumento $\langle \mathcal{A}, p \rangle$ es un* derrotador *para el argumento $\langle \mathcal{B}, q \rangle$ si y sólo si existe un sub-argumento $\langle \mathcal{C}, h \rangle$ de $\langle \mathcal{B}, q \rangle$ tal que $\langle \mathcal{A}, p \rangle$ ataca a $\langle \mathcal{B}, q \rangle$ en el literal h y vale que:*

- *$\langle \mathcal{A}, p \rangle$ es mejor que $\langle \mathcal{C}, h \rangle$ de acuerdo al criterio de comparación utilizado (i. e., $\langle \mathcal{C}, h \rangle \prec \langle \mathcal{A}, p \rangle$), denotando que $\langle \mathcal{A}, p \rangle$ es un* derrotador propio *de $\langle \mathcal{B}, q \rangle$; o*

[5]El criterio es especificado por el usuario, permitiendo así emplear aquel que mejor se adapte al dominio de aplicación.

- *$\langle \mathcal{A}, p\rangle$ no es mejor que $\langle \mathcal{C}, h\rangle$, y $\langle \mathcal{C}, h\rangle$ no es mejor que $\langle \mathcal{A}, p\rangle$ de acuerdo al criterio de comparación utilizado (i. e., $\langle \mathcal{C}, h\rangle \not\succ \langle \mathcal{A}, p\rangle$ y $\langle \mathcal{A}, p\rangle \not\succ \langle \mathcal{C}, h\rangle$), denotando que $\langle \mathcal{A}, p\rangle$ es un* derrotador por bloqueo *de $\langle \mathcal{B}, q\rangle$.*

La definición anterior identifica dos situaciones en las que un ataque entre argumentos resulta exitoso. La primera de ellas, correspondiente a la derrota propia, describe la situación en que el argumento atacante $\langle \mathcal{A}, p\rangle$ es estrictamente preferido al sub-argumento atacado $\langle \mathcal{C}, h\rangle$. En contraste, el segundo caso describe la derrota por bloqueo, la cual ocurre cuando el argumento atacante $\langle \mathcal{A}, p\rangle$ y el sub-argumento atacado $\langle \mathcal{C}, h\rangle$ son igualmente preferidos o resultan incomparables de acuerdo al criterio de comparación adoptado.

Si bien DeLP utiliza un criterio de comparación modular, existen criterios concretos para comparar argumentos en este formalismo (ver *e. g.*, [FEGS07, GS04]). En particular, en los ejemplos de esta tesis utilizaremos el criterio de comparación *especificidad generalizada* [Gar00, GS04], el cual se basa en la noción de especificidad definida en [Poo85, SL92]. Resumidamente, este criterio prefiere aquellos argumentos con información más precisa o argumentos más directos. Por ejemplo, el argumento $\langle \{\sim a \prec b, c\}, \sim a\rangle$ es considerado más específico que el argumento $\langle \{a \prec b\}, a\rangle$ porque utiliza más información para derivar su conclusión. De manera similar, el argumento $\langle \{a \prec b\}, a\rangle$ es más específico que el argumento $\langle \{(\sim a \prec d), (d \prec b)\}, \sim a\rangle$ por ser más directo.

Definición 2.24 (Especificidad Generalizada) *Sea $\mathcal{P} = (\Theta, \Omega, \Delta)$ un programa lógico rebatible y sea $Lits(\mathcal{P})$ el conjunto de literales que tienen una derivación (estricta o rebatible) a partir de $\mathcal{P}$. Un argumento $\langle \mathcal{A}, p\rangle$ es* más específico *que otro argumento $\langle \mathcal{B}, q\rangle$, denotado $\langle \mathcal{B}, q\rangle \prec \langle \mathcal{A}, p\rangle$, si se verifican las siguientes condiciones:*

1. *para todo conjunto $H \subseteq Lits(\mathcal{P})$ vale que:*
   si existe una derivación rebatible para p a partir de $\Omega \cup H \cup \mathcal{A}$ (i. e., H activa a $\mathcal{A}$)
   y no existe una derivación estricta para p a partir de $\Omega \cup H$,
   entonces existe una derivación rebatible para q a partir de $\Omega \cup H \cup \mathcal{B}$ (i. e., H activa a $\mathcal{B}$); y

2. *existe al menos un conjunto $H' \subseteq Lits(\mathcal{P})$ tal que:*
   existe una derivación rebatible para q a partir de $\Omega \cup H' \cup \mathcal{B}$ (i. e., H' activa a $\mathcal{B}$),
   no existe una derivación estricta para q a partir de $\Omega \cup H'$, y
   no existe una derivación rebatible para p a partir de $\Omega \cup H' \cup \mathcal{A}$ (i. e., H' no activa a $\mathcal{A}$).

Dada la definición anterior, los conjuntos H y H' contienen literales, los cuales son tomados como hechos para obtener las derivaciones. La condición (1) establece que para todo conjunto de literales $H \subseteq Lits(\mathcal{P})$, si H permite derivar p utilizando reglas de $\Omega \cup \mathcal{A}$, donde al menos una regla de $\mathcal{A}$ es utilizada (ya que no existe una derivación estricta de p a partir de $\Omega \cup H$), entonces ese mismo conjunto H permite derivar rebatiblemente a q utilizando las reglas de $\Omega \cup \mathcal{B}$. Luego, la condición (2) busca que exista al menos un conjunto $H' \subseteq Lits(\mathcal{P})$ que permita derivar rebatiblemente a q pero no a p. Para ilustrar el uso de este criterio de comparación, consideremos el siguiente ejemplo.

Ejemplo 2.12 *Como se observó en el Ejemplo 2.11, existen ataques de los argumentos $\langle \mathcal{A}, vuela(tina) \rangle$ y $\langle \mathcal{C}, vuela(tina) \rangle$ hacia el argumento $\langle \mathcal{B}, \sim vuela(tina) \rangle$ y viceversa, donde los sub-argumentos de desacuerdo son los mismos argumentos atacados. Veamos ahora cómo se resuelven estos ataques mediante el uso de especificidad generalizada como criterio de comparación. En este caso debemos comparar el argumento $\langle \mathcal{B}, \sim vuela(tina) \rangle$ respectivamente con los argumentos $\langle \mathcal{A}, vuela(tina) \rangle$ y $\langle \mathcal{C}, vuela(tina) \rangle$.*

Todo conjunto H_1 que active al argumento

$$\langle \mathcal{B}, \sim vuela(tina) \rangle = \langle \{\sim vuela(tina) \prec gallina(tina)\}, \sim vuela(tina) \rangle$$

deberá contener al literal "gallina(tina)", por lo que H también activará al argumento

$$\langle \mathcal{A}, vuela(tina) \rangle = \langle \{vuela(tina) \prec ave(tina)\}, vuela(tina) \rangle$$

dada la existencia de la regla estricta "ave(tina) $\leftarrow$ gallina(tina)" en el programa $\mathcal{P}_{2.5}$.

En contraste, el conjunto $H_1' = \{ave(tina)\}$ activa a $\langle \mathcal{A}, vuela(tina) \rangle$ pero no a $\langle \mathcal{B}, \sim vuela(tina) \rangle$. De esta manera, $\langle \mathcal{A}, vuela(tina) \rangle$ es más específico que $\langle \mathcal{B}, \sim vuela(tina) \rangle$, conduciendo a una derrota propia de $\langle \mathcal{A}, vuela(tina) \rangle$ sobre $\langle \mathcal{B}, \sim vuela(tina) \rangle$.

Por otra parte, todo conjunto H_2 que active al argumento

$$\langle \mathcal{C}, vuela(tina) \rangle = \langle \{vuela(tina) \prec gallina(tina), asustada(tina)\}, vuela(tina) \rangle$$

deberá contener los literales "gallina(tina)" y "asustada(tina)", por lo que el argumento

$$\langle \mathcal{B}, \sim vuela(tina) \rangle = \langle \{\sim vuela(tina) \prec gallina(tina)\}, \sim vuela(tina) \rangle$$

también será activado por H_2.

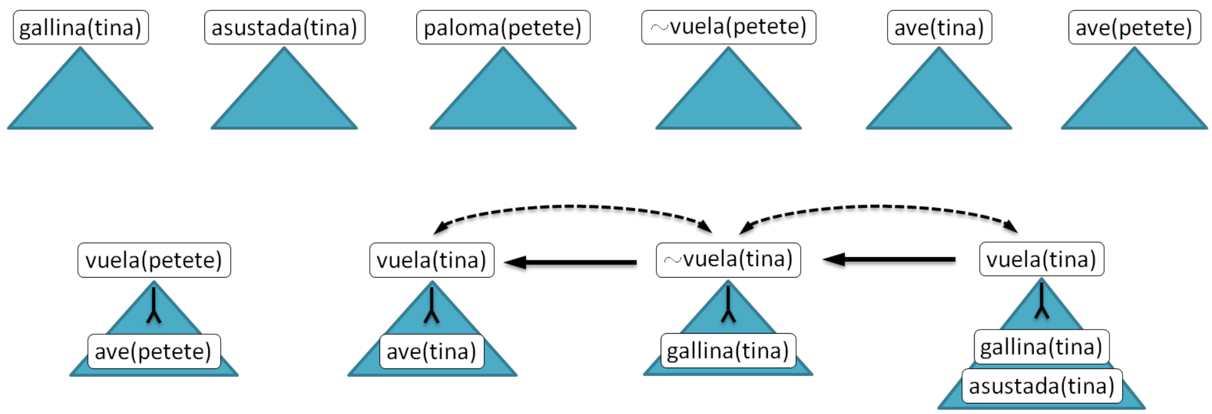


Figura 2.3: Representación gráfica de los argumentos, ataques y derrotas obtenidos a partir del programa lógico rebatible $\mathcal{P}_{2.5}$ del Ejemplo 2.5.

Luego, el conjunto $H'_2 = \{gallina(tina)\}$ *activa a* $\langle \mathcal{B}, {\sim}vuela(tina) \rangle$ *pero no a* $\langle \mathcal{C}, vuela(tina) \rangle$*. En consecuencia,* $\langle \mathcal{C}, vuela(tina) \rangle$ *es más específico que* $\langle \mathcal{B}, {\sim}vuela(tina) \rangle$*, siendo entonces un derrotador propio de* $\langle \mathcal{B}, {\sim}vuela(tina) \rangle$*.*

A continuación introduciremos una representación gráfica para los argumentos de DeLP, los ataques entre argumentos y las correspondientes derrotas, la cual será adaptada posteriormente para los desarrollos presentados en el Capítulo 5. En la notación gráfica propuesta, los argumentos serán representados mediante triángulos. Además, los triángulos dentro de triángulos mayores denotarán sub-argumentos. La información ubicada en el vértice superior de un triángulo corresponde a la conclusión del argumento (sub-argumento) respectivo, mientras que la información situada dentro del triángulo corresponde a los literales utilizados para construir el argumento. El símbolo " $\prec$ " representa la conexión establecida entre estos literales a través de las reglas rebatibles pertenecientes al argumento. En particular, los sub-argumentos correspondientes a argumentos estrictos se denotarán identificando únicamente el literal que concluyen. Por último, los ataques entre argumentos se denotarán mediante flechas punteadas, mientras que las derrotas estarán identificadas mediante flechas sólidas. Esta notación se halla ilustrada por el siguiente ejemplo.

Ejemplo 2.13 *Consideremos el programa lógico rebatible* $\mathcal{P}_{2.5}$ *del Ejemplo 2.5. La Figura 2.3 ilustra todos los argumentos construibles a partir de* $\mathcal{P}_{2.5}$ *(algunos de los cuales* *fueron previamente introducidos en el Ejemplo 2.8), así como también los vínculos de ataque y derrota entre ellos (respectivamente introducidos en los ejemplos 2.11 y 2.12).*

2.2.2.6. Cálculo de Aceptabilidad: argumentos aceptados y conclusiones garantizadas en DeLP

Como se explicó al comienzo de este capítulo, el objetivo de un sistema argumentativo está determinado por el quinto elemento de la estructura conceptual presentada en la Sección 2.1: identificación de los argumentos aceptados. Para esto, es necesario considerar las derrotas entre los argumentos del sistema. En el caso de DeLP, además, los argumentos aceptados determinarán qué literales están *garantizados* a partir de un programa lógico rebatible. Dado el conjunto de argumentos que pueden construirse a partir de un programa lógico rebatible $\mathcal{P}$, la relación de derrota entre argumentos sólo puede establecer un orden de preferencia sobre los argumentos en conflicto. Sin embargo, el estado de aceptabilidad de un argumento $\langle \mathcal{A}, h \rangle$ con respecto al programa $\mathcal{P}$ dependerá de la interacción de $\langle \mathcal{A}, h \rangle$ con el resto de los argumentos del sistema.

Por ejemplo, si a partir de un programa $\mathcal{P}$ se obtiene un argumento $\langle \mathcal{A}_0, h_0 \rangle$ que no posee derrotadores, entonces un individuo que disponga de $\mathcal{P}$ para razonar podrá "creer" en el literal h_0. Sin embargo, podría ocurrir que $\langle \mathcal{A}_0, h_0 \rangle$ posea un derrotador $\langle \mathcal{A}_1, h_1 \rangle$, con lo cual existirán razones para invalidar la creencia del individuo en h_0. Tal es el caso del Ejemplo 2.12, donde la derrota de $\langle \mathcal{B}, {\sim}vuela(tina) \rangle$ sobre $\langle \mathcal{A}, vuela(tina) \rangle$ invalida la creencia de que Tina vuela. Sin embargo, también puede ocurrir que el argumento $\langle \mathcal{A}_1, h_1 \rangle$ tenga a su vez un derrotador $\langle \mathcal{A}_2, h_2 \rangle$, reinstalando así la creencia del individuo en h_0. Nuevamente, esto ocurre en el Ejemplo 2.12 mediante la derrota de $\langle \mathcal{C}, vuela(tina) \rangle$ sobre $\langle \mathcal{B}, {\sim}vuela(tina) \rangle$, reinstalando la creencia de que Tina vuela.

La secuencia de interacciones arriba mencionada puede ir más lejos, existiendo a su vez un derrotador $\langle \mathcal{A}_3, h_3 \rangle$ para $\langle \mathcal{A}_2, h_2 \rangle$ que invalida nuevamente la creencia del individuo en h_0 y así siguiendo. Esto da origen a una secuencia de argumentos $[\langle \mathcal{A}_0, h_0 \rangle, \langle \mathcal{A}_1, h_1 \rangle, \langle \mathcal{A}_2, h_2 \rangle, \langle \mathcal{A}_3, h_3 \rangle, \ldots]$ conocida como *línea argumentativa*, donde cada elemento en la secuencia es un derrotador de su predecesor.

Definición 2.25 (Línea Argumentativa) *Sea $\mathcal{P}$ un programa lógico rebatible y $\langle \mathcal{A}_0, h_0 \rangle$ un argumento construido a partir de $\mathcal{P}$. Una* línea argumentativa *para $\langle \mathcal{A}_0, h_0 \rangle$ es una secuencia de argumentos de $\mathcal{P}$, denotada*

$\Lambda = [\langle \mathcal{A}_0, h_0 \rangle, \langle \mathcal{A}_1, h_1 \rangle, \langle \mathcal{A}_2, h_2 \rangle, \langle \mathcal{A}_3, h_3 \rangle, \ldots]$, *donde cada argumento* $\langle \mathcal{A}_i, h_i \rangle$ *perteneciente a la secuencia es un derrotador del argumento predecesor* $\langle \mathcal{A}_{i-1}, h_{i-1} \rangle$.

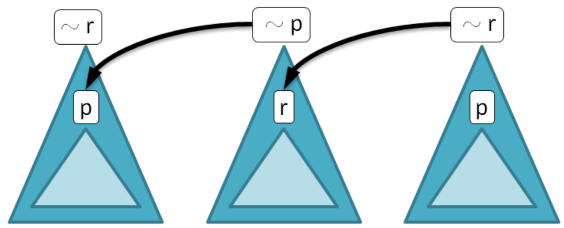
Dada una línea argumentativa para el argumento $\langle \mathcal{A}_0, h_0 \rangle$, los argumentos en posiciones impares conforman el conjunto de *argumentos de apoyo*[6] (es decir, argumentos que de alguna manera reinstalan la posición de $\langle \mathcal{A}_0, h_0 \rangle$) y se denotan como Λ_A , mientras que los argumentos en posiciones pares componen el conjunto de *argumentos de interferencia* (es decir, argumentos en contra de la posición de $\langle \mathcal{A}_0, h_0 \rangle$) denotado como Λ_I.

Para evitar secuencias de argumentos indeseables que pueden representar cadenas de razonamiento falaz, DeLP impone el requerimiento de que una línea argumentativa debe ser *aceptable*. De esta manera, se define una serie de características que este tipo de líneas argumentativas debe satisfacer. Resumidamente, una línea argumentativa será aceptable si es una secuencia finita, no contiene argumentos o sub-argumentos repetidos, no posee dos derrotas por bloqueo consecutivas, y los conjuntos de argumentos de apoyo e interferencia son consistentes con el conocimiento estricto.

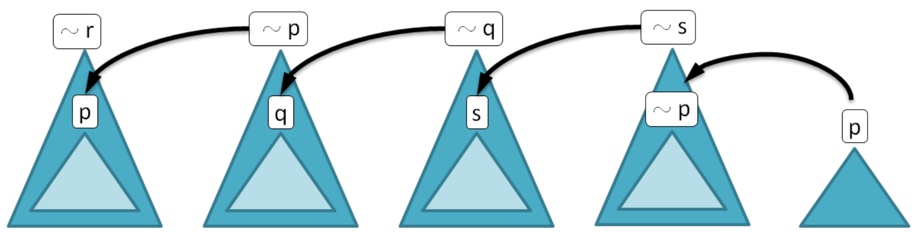
El requisito de que una línea argumentativa sea finita está directamente relacionado con el requerimiento de no introducir argumentos repetidos en la línea, con el objetivo de evitar lo que se denomina como *argumentación circular* [GS04]. La Figura 2.4 ilustra dos situaciones que conducen a argumentación circular. La primera de ellas, correspondiente a la Figura 2.4(a), ocurre cuando un argumento es reintroducido para defenderse a sí mismo. El otro caso, ilustrado en la Figura 2.4(b), corresponde a una generalización del caso anterior, en el que un sub-argumento de desacuerdo es reintroducido en la línea para defender al argumento que lo contiene. Claramente, las situaciones ilustradas en la Figura 2.4 conducen a procesos argumentativos falaces y posiblemente infinitos, motivo por el cual debe evitarse que ocurran en una línea argumentativa aceptable.

El requisito de impedir dos derrotas por bloqueo consecutivas en una línea argumentativa aceptable es razonable por el siguiente motivo. Una derrota por bloqueo de un argumento $\langle \mathcal{A}_1, h_1 \rangle$ sobre un argumento $\langle \mathcal{A}_0, h_0 \rangle$ se desprende del hecho de que los argumentos $\langle \mathcal{A}_1, h_1 \rangle$ y $\langle \mathcal{A}_0, h_0 \rangle$ resultaron equivalentes o incomparables de acuerdo al criterio de comparación utilizado. Es decir, es una derrota que corresponde a un conflicto que

[6]En [GS04] los autores denominan a estos argumentos como "argumentos de soporte". Sin embargo, para evitar confusiones con el uso de la noción de soporte en esta tesis, se decidió utilizar la terminología "argumentos de apoyo".



(a)



(b)

Figura 2.4: Ejemplos de argumentación circular.

no pudo ser resuelto mediante la asignación de pesos a los argumentos. Luego, si existe un argumento $\langle \mathcal{A}_2, h_2 \rangle$ tal que derrota por bloqueo a $\langle \mathcal{A}_1, h_1 \rangle$, se tiene que el criterio de comparación tampoco fue capaz de decidir entre los argumentos $\langle \mathcal{A}_2, h_2 \rangle$ y $\langle \mathcal{A}_1, h_1 \rangle$. De esta manera, si se permitiera incluir dos derrotas por bloqueo seguidas en una línea argumentativa aceptable, se estaría determinando en cierta medida que el argumento $\langle \mathcal{A}_0, h_0 \rangle$ prevalece ante la derrota de $\langle \mathcal{A}_1, h_1 \rangle$, dado que es defendido por $\langle \mathcal{A}_2, h_2 \rangle$. Claramente esta es una situación indeseable, ya que el literal h_0 pertenecería a las inferencias del sistema simplemente por poseer más argumentos que lo apoyan.

Por último, la condición que indica que los argumentos de apoyo y los de interferencia deben ser consistentes con el conocimiento estricto intenta evitar la situación falaz en la que un argumento es defendido por otro argumento que se encuentra en conflicto con él. Este requerimiento establece que al considerar conjuntamente estos argumentos y el conocimiento estricto, el conjunto resultante no debe ser contradictorio. Teniendo en cuenta todas estas condiciones, la siguiente definición formaliza la noción de línea argumentativa aceptable.

Definición 2.26 (Línea Argumentativa Aceptable) *Sea $\mathcal{P} = (\Pi, \Delta)$ un programa lógico rebatible y $\Lambda = [\langle \mathcal{A}_0, h_0 \rangle, \ldots, \langle \mathcal{A}_i, h_i \rangle, \ldots, \langle \mathcal{A}_n, h_n \rangle]$ una línea argumentativa para $\langle \mathcal{A}_0, h_0 \rangle$ obtenida a partir de $\mathcal{P}$. Diremos que Λ es una* línea argumentativa aceptable *si satisface las siguientes condiciones:*

1. *Λ es una secuencia finita;*
2. *el conjunto Λ_A de argumentos de apoyo de Λ y el conjunto Λ_I de argumentos de interferencia de Λ son tal que $\Pi \cup \{\mathcal{A} \mid \langle \mathcal{A}, p \rangle \in \Lambda_A\}$ y $\Pi \cup \{\mathcal{B} \mid \langle \mathcal{B}, q \rangle \in \Lambda_I\}$ son conjuntos no contradictorios;*
3. *ningún argumento $\langle \mathcal{A}_k, h_k \rangle$ de Λ es un sub-argumento de un argumento $\langle \mathcal{A}_j, h_j \rangle$ que aparece previamente en Λ $(j < k)$; y*
4. *para todo argumento $\langle \mathcal{A}_i, h_i \rangle$ de Λ tal que $\langle \mathcal{A}_i, h_i \rangle$ es un derrotador por bloqueo de $\langle \mathcal{A}_{i-1}, h_{i-1} \rangle$ y no es un derrotador propio de $\langle \mathcal{A}_{i-1}, h_{i-1} \rangle$, si existe $\langle \mathcal{A}_{i+1}, h_{i+1} \rangle$ en Λ, entonces $\langle \mathcal{A}_{i+1}, h_{i+1} \rangle$ es un derrotador propio de $\langle \mathcal{A}_i, h_i \rangle$.*

Dada la Definición 2.26, pareciera ser que la noción de línea argumentativa aceptable provee un mecanismo para decidir sobre la garantía de un literal h_0 a partir de un programa $\mathcal{P}$. Siguiendo esta definición, podríamos establecer lo siguiente: si la línea argumentativa para $\langle \mathcal{A}_0, h_0 \rangle$ tiene un número impar de argumentos, entonces el literal h_0 está garantizado; en caso contrario, si la línea argumentativa posee una cantidad par de argumentos o no existe argumento alguno para h_0, entonces no lo está. No obstante, como se muestra en el siguiente ejemplo, una línea argumentativa no es suficiente para analizar la interacción entre argumentos, ya que para cada argumento puede existir más de un derrotador.

Ejemplo 2.14 *Sea $\mathcal{P}$ un programa lógico rebatible a partir del cual se construyen los argumentos $\langle \mathcal{A}_0, h_0 \rangle$, $\langle \mathcal{A}_1, h_1 \rangle$, $\langle \mathcal{A}_2, h_2 \rangle$ y $\langle \mathcal{A}_3, h_3 \rangle$. Supongamos que $\langle \mathcal{A}_1, h_1 \rangle$ y $\langle \mathcal{A}_2, h_2 \rangle$ son derrotadores propios de $\langle \mathcal{A}_0, h_0 \rangle$, y que $\langle \mathcal{A}_3, h_3 \rangle$ es un derrotador propio de $\langle \mathcal{A}_2, h_2 \rangle$. A partir de esto, se obtienen las siguientes líneas argumentativas aceptables para $\langle \mathcal{A}_0, h_0 \rangle$:*

- $\Lambda_1 = [\langle \mathcal{A}_0, h_0 \rangle]$
- $\Lambda_2 = [\langle \mathcal{A}_0, h_0 \rangle, \langle \mathcal{A}_1, h_1 \rangle]$
- $\Lambda_3 = [\langle \mathcal{A}_0, h_0 \rangle, \langle \mathcal{A}_2, h_2 \rangle]$

- $\Lambda_4 = [\langle \mathcal{A}_0, h_0 \rangle, \langle \mathcal{A}_2, h_2 \rangle, \langle \mathcal{A}_3, h_3 \rangle]$

En este caso, no resulta adecuado considerar aisladamente las líneas argumentativas aceptables para decidir sobre la garantía de h_0 a partir de $\mathcal{P}$. Esto se debe a que, por ejemplo, la línea argumentativa Λ_4 posee una cantidad impar de argumentos, sugiriendo que h_0 estaría garantizado a partir de $\mathcal{P}$. En contraste, por ejemplo, Λ_2 sugiere lo contrario, dado que posee una cantidad par de argumentos.

El Ejemplo 2.14 muestra que la noción de línea argumentativa aceptable no es suficiente como mecanismo de prueba para los literales inferidos a partir de un programa lógico rebatible. Dado que para cada argumento puede existir más de un derrotador, esto conduce a una ramificación de líneas argumentativas, dando origen a un árbol de derrotadores llamado *árbol de dialéctica*. Además, nótese que la noción de línea argumentativa aceptable no asegura la exploración completa de los derrotadores para un argumento. Si consideramos el Ejemplo 2.14, la línea argumentativa $[\langle \mathcal{A}_0, h_0 \rangle]$ satisface las condiciones impuestas por la Definición 2.26. De esta manera, para asegurar el análisis exhaustivo de derrotadores, la noción de árbol de dialéctica agrupa todas las líneas argumentativas aceptables maximales.

Definición 2.27 (Árbol de Dialéctica) *Sea $\mathcal{P}$ un programa lógico rebatible y $\langle \mathcal{A}_0, h_0 \rangle$ un argumento construido a partir de $\mathcal{P}$. Un* árbol de dialéctica para $\langle \mathcal{A}_0, h_0 \rangle$*, denotado $\mathcal{T}_{\langle \mathcal{A}_0, h_0 \rangle}$, es un árbol tal que los nodos son argumentos construidos a partir de $\mathcal{P}$, los arcos representan derrotas entre los nodos, y se verifican las siguientes condiciones:*

1. *$\langle \mathcal{A}_0, h_0 \rangle$ es la raíz de $\mathcal{T}_{\langle \mathcal{A}_0, h_0 \rangle}$; y*
2. *si $\Lambda = [\langle \mathcal{A}_0, h_0 \rangle, \langle \mathcal{A}_1, h_1 \rangle, \ldots, \langle \mathcal{A}_n, h_n \rangle]$ es una línea argumentativa aceptable para $\langle \mathcal{A}_0, h_0 \rangle$ y no existe $\langle \mathcal{A}_m, h_m \rangle$ tal que $\Lambda' = [\langle \mathcal{A}_0, h_0 \rangle, \langle \mathcal{A}_1, h_1 \rangle, \ldots, \langle \mathcal{A}_n, h_n \rangle, \langle \mathcal{A}_m, h_m \rangle]$ es una línea argumentativa aceptable para $\langle \mathcal{A}_0, h_0 \rangle$, entonces $\langle \mathcal{A}_n, h_n \rangle$ es una hoja de $\mathcal{T}_{\langle \mathcal{A}_0, h_0 \rangle}$ y Λ es un camino desde la raíz $\langle \mathcal{A}_0, h_0 \rangle$ hasta la hoja $\langle \mathcal{A}_n, h_n \rangle$.*

Ejemplo 2.15 *Consideremos el programa lógico rebatible $\mathcal{P}_{2.5}$ del Ejemplo 2.5 y el argumento $\langle \mathcal{A}, vuela(tina) \rangle$ construido a partir de $\mathcal{P}_{2.5}$ (ver Ejemplo 2.8). De acuerdo al Ejemplo 2.12, el argumento $\langle \mathcal{B}, \sim vuela(tina) \rangle$ es un derrotador propio de $\langle \mathcal{A}, vuela(tina) \rangle$, y el argumento $\langle \mathcal{C}, vuela(tina) \rangle$ es un derrotador propio de $\langle \mathcal{B}, \sim vuela(tina) \rangle$, siendo estas las*

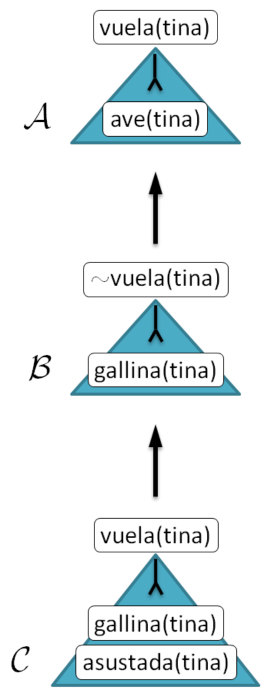


Figura 2.5: Árbol de dialéctica $\mathcal{T}_{\langle \mathcal{A}, vuela(tina)\rangle}$ para el argumento $\langle \mathcal{A}, vuela(tina)\rangle$, correspondiente al Ejemplo 2.15.

únicas derrotas existentes. Por lo tanto, la única línea argumentativa aceptable (maximal) para $\langle \mathcal{A}, vuela(tina)\rangle$ *es* $\Lambda_{2.15} = [\langle \mathcal{A}, vuela(tina)\rangle, \langle \mathcal{B}, {\sim}vuela(tina)\rangle, \langle \mathcal{C}, vuela(tina)\rangle]$*. En particular,* $\Lambda_{2.15}$ *constituye el árbol de dialéctica* $\mathcal{T}_{\langle \mathcal{A}, vuela(tina)\rangle}$ *para el argumento* $\langle \mathcal{A}, vuela(tina)\rangle$*, y se encuentra ilustrado en la Figura 2.5.*

Como puede observarse en el Ejemplo 2.15, los nodos hoja del árbol de dialéctica corresponden a argumentos no derrotados. En contraste, un nodo interno que tiene como hijo a un nodo hoja corresponderá a un argumento derrotado. Siguiendo con este análisis desde las hojas hacia la raíz, los nodos en un árbol de dialéctica pueden marcarse como *derrotados* o *no derrotados*, respectivamente denotados como **D** y **U** (por sus nombres en inglés, *defeated* y *undefeated*). La siguiente definición determina cómo se efectúa el procedimiento de marcado para un árbol de dialéctica de DeLP.

Definición 2.28 (Árbol de Dialéctica Marcado) *Sea* $\mathcal{P}$ *un programa lógico rebatible y* $\mathcal{T}_{\langle \mathcal{A}, h\rangle}$ *un árbol de dialéctica para un argumento* $\langle \mathcal{A}, h\rangle$ *construido a partir de* $\mathcal{P}$*. El* árbol de dialéctica marcado *para* $\langle \mathcal{A}, h\rangle$*, denotado* $\mathcal{T}^*_{\langle \mathcal{A}, h\rangle}$*, se obtiene marcando cada nodo* $\mathcal{N}$ *de* $\mathcal{T}_{\langle \mathcal{A}, h\rangle}$ *como* **D** *(derrotado) o* **U** *(no derrotado) de acuerdo al siguiente criterio:*

- *Si $\mathcal{N}$ es un nodo hoja de $\mathcal{T}_{\langle\mathcal{A},h\rangle}$, entonces $\mathcal{N}$ es marcado como* **U** *en $\mathcal{T}^*_{\langle\mathcal{A},h\rangle}$.*
- *Si $\mathcal{N}$ es un nodo interno de $\mathcal{T}_{\langle\mathcal{A},h\rangle}$ y posee algún hijo marcado como* **U** *en $\mathcal{T}^*_{\langle\mathcal{A},h\rangle}$, entonces $\mathcal{N}$ es marcado como* **D** *en $\mathcal{T}^*_{\langle\mathcal{A},h\rangle}$.*
- *Si $\mathcal{N}$ es un nodo interno de $\mathcal{T}_{\langle\mathcal{A},h\rangle}$ y todos sus hijos son marcados como* **D** *en $\mathcal{T}^*_{\langle\mathcal{A},h\rangle}$, entonces $\mathcal{N}$ es marcado como* **U** *en $\mathcal{T}^*_{\langle\mathcal{A},h\rangle}$.*

Ejemplo 2.16 *Dado el árbol de dialéctica $\mathcal{T}_{\langle\mathcal{A},\, vuela(tina)\rangle}$ del Ejemplo 2.15 (ilustrado en la Figura 2.5), el proceso de marcado determina que el árbol de dialéctica marcado para $\langle\mathcal{A}, vuela(tina)\rangle$ es $\mathcal{T}^*_{\langle\mathcal{A},\, vuela(tina)\rangle}$, donde el argumento $\langle\mathcal{C}, vuela(tina)\rangle$ está marcado como* **U***, el argumento $\langle\mathcal{B}, \sim vuela(tina)\rangle$ está marcado como* **D***, y el argumento $\langle\mathcal{A}, vuela(tina)\rangle$ está marcado como* **U***.*

El siguiente ejemplo ilustra un caso más general del proceso de marcado de árboles de dialéctica en DeLP.

Ejemplo 2.17 *Consideremos un programa lógico rebatible a partir del cual se construye un argumento $\langle\mathcal{A}, a\rangle$, cuyo árbol de dialéctica $\mathcal{T}_{\langle\mathcal{A},a\rangle}$ se ilustra en la Figura 2.6. El árbol de dialéctica marcado $\mathcal{T}^*_{\langle\mathcal{A},a\rangle}$, correspondiente al argumento $\langle\mathcal{A}, a\rangle$, se ilustra en la Figura 2.7. En este caso, podemos observar que los argumentos correspondientes a nodos hoja del árbol son marcados como* **U***. Luego, aquellos argumentos correspondientes a nodos padre de los nodos hoja son marcados como* ***D****. Finalmente, y continuando con el proceso establecido por la Definición 2.28, tenemos que el argumento $\langle\mathcal{A}, a\rangle$, correspondiente a la raíz del árbol de dialéctica marcado $\mathcal{T}^*_{\langle\mathcal{A},a\rangle}$, está marcado como* **D***.*

Un árbol de dialéctica marcado $\mathcal{T}^*_{\langle\mathcal{A},h\rangle}$ representa el análisis dialéctico que considera todos los argumentos relevantes de un programa lógico rebatible $\mathcal{P}$ a fin de decidir el estado de aceptabilidad de un argumento $\langle\mathcal{A}, h\rangle$. Es decir, el análisis dialéctico toma en cuenta todos los argumentos de $\mathcal{P}$ que influyen, por pertenecer a una línea argumentativa aceptable para $\langle\mathcal{A}, h\rangle$, en la determinación del estado de aceptabilidad del argumento $\langle\mathcal{A}, h\rangle$. Por lo tanto, el estado de un argumento $\langle\mathcal{A}, h\rangle$ será *rechazado* si la raíz de $\mathcal{T}^*_{\langle\mathcal{A},h\rangle}$ está marcada como **D**, o *aceptado* si la raíz de $\mathcal{T}^*_{\langle\mathcal{A},h\rangle}$ está marcada como **U**.

Definición 2.29 (Argumento Aceptado) *Sea $\mathcal{P}$ un programa lógico rebatible y $\langle\mathcal{A}, h\rangle$ un argumento construido a partir de $\mathcal{P}$. Diremos que $\langle\mathcal{A}, h\rangle$ es un* argumento aceptado

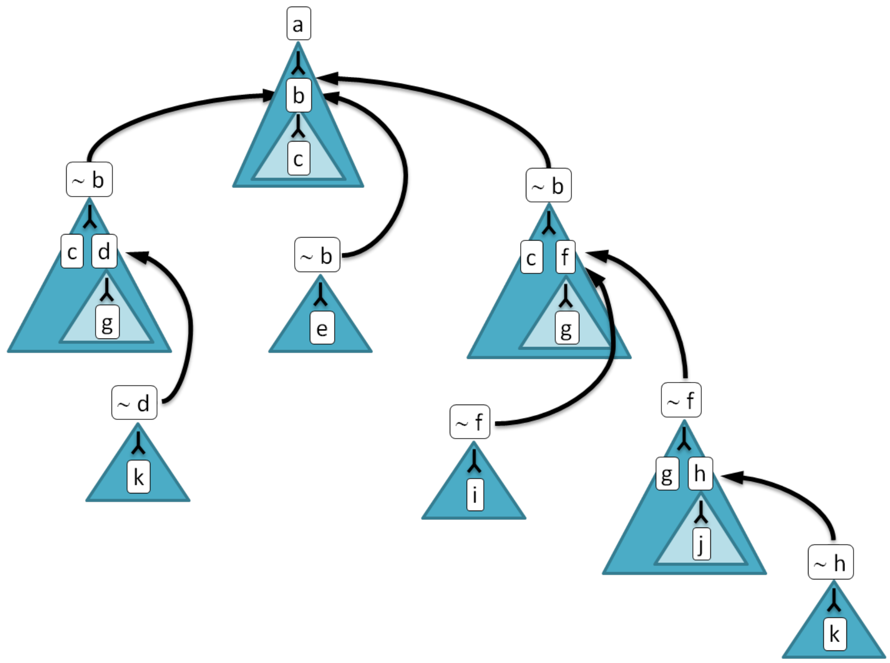


Figura 2.6: Árbol de dialéctica $\mathcal{T}_{\langle \mathcal{A}, a \rangle}$, correspondiente al Ejemplo 2.17.

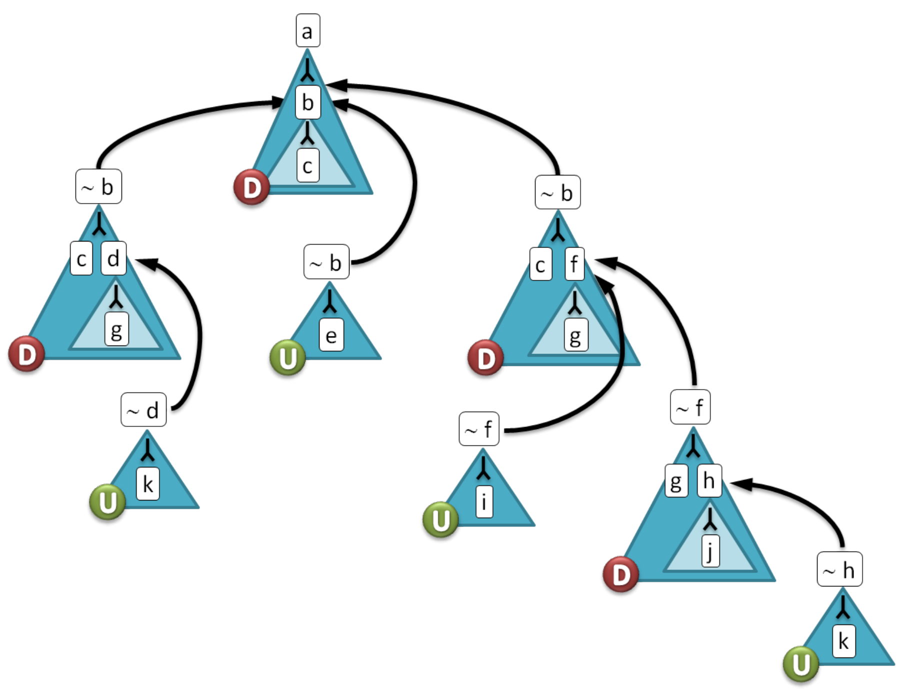


Figura 2.7: Árbol de dialéctica marcado $\mathcal{T}^*_{\langle \mathcal{A}, a \rangle}$, correspondiente al Ejemplo 2.17.

de $\mathcal{P}$ si y sólo si la raíz del árbol de dialéctica marcado $\mathcal{T}^*_{\langle \mathcal{A}, h\rangle}$ está marcada como U*. En caso contrario, diremos que $\langle \mathcal{A}, h\rangle$ es un* argumento rechazado *de $\mathcal{P}$.*

Claramente, si hay un argumento aceptado a partir de un programa lógico rebatible, esto implica que hay una garantía para inferir su conclusión. Siguiendo este razonamiento, DeLP define la noción de *literal garantizado*, correspondiente a aquellos literales que son las conclusiones de los argumentos aceptados. En consecuencia, los literales garantizados a partir de un programa lógico rebatible serán las inferencias obtenidas bajo el procedimiento de prueba dialéctico de DeLP.

Definición 2.30 (Literal Garantizado) *Sea $\mathcal{P}$ un programa lógico rebatible y h un literal. Diremos que h es un* literal garantizado *a partir de $\mathcal{P}$ si y sólo si existe un argumento $\langle \mathcal{A}, h\rangle$ tal que es un argumento aceptado de $\mathcal{P}$.*

Por ejemplo, consideremos el programa lógico rebatible $\mathcal{P}_{2.5}$ del Ejemplo 2.5 y el argumento $\langle \mathcal{A}, vuela(tina)\rangle$ construido a partir de dicho programa. Como se ilustra en el Ejemplo 2.16, la raíz del árbol de dialéctica marcado $\mathcal{T}^*_{\langle \mathcal{A}, vuela(tina)\rangle}$ está etiquetada como U, implicando que $\langle \mathcal{A}, vuela(tina)\rangle$ es un argumento aceptado de $\mathcal{P}_{2.5}$. En consecuencia, podemos concluir que "$vuela(tina)$" pertenece a las inferencias del programa por ser un literal garantizado a partir de $\mathcal{P}_{2.5}$.

2.3. La Noción de Soporte en Argumentación

Como se mencionó al comienzo de este capítulo, las bases para el estudio de argumentación residen en áreas como el razonamiento legal y la filosofía [Tou58, Pol87]. Uno de los trabajos más influyentes es el libro "*The Uses of Argument*" escrito por Stephen Toulmin en el año 1958. En su libro Toulmin promulgó la idea de que es necesario analizar los argumentos utilizando un formato más rico que la dicotomía de premisas y conclusiones empleada en la lógica formal. Por lo tanto, propuso un modelo para el diseño de argumentos que, además de *datos* (*data*) y *conclusión* (*claim*), distingue cuatro elementos adicionales: *garantía* (*warrant*), *respaldo* (*backing*)[7], *derrotador* (*rebuttal*), y *calificador*

[7]Teniendo en cuenta que las nociones de "*warrant*" y "*backing*" propuestas por Toulmin serán utilizadas con frecuencia a lo largo de esta tesis, y dada la falta de una traducción fiel para estos conceptos, de aquí en adelante utilizaremos los términos correspondientes al idioma inglés.

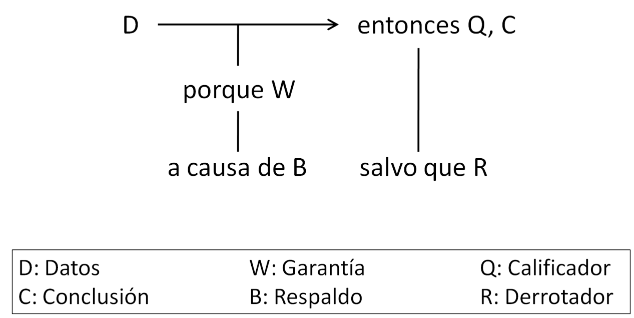


Figura 2.8: El modelo para el diseño de argumentos propuesto por Toulmin [Tou58].

(*qualifier*). La Figura 2.8 provee una representación gráfica del esquema propuesto por Toulmin.

De acuerdo a lo establecido por Toulmin, la *conclusión* es la aserción original a la que se compromete el individuo que propone el argumento, la cual debe ser justificada en caso de ser cuestionada. Los *datos* constituyen la base que se produce como soporte para la aserción original; en otras palabras, representan la información en la que se basa la conclusión. El *warrant* es una sentencia general en forma de regla que permite arribar a la conclusión del argumento; es decir, es una regla de inferencia que provee la conexión entre los datos y la conclusión. El *calificador* representa el grado de fuerza con el que los datos permiten arribar a la conclusión en virtud de la regla de inferencia expresada por el warrant. Al cuestionar una conclusión puede preguntarse por qué el warrant correspondiente es válido. Por lo tanto, el cuestionar una conclusión de esta manera puede dar lugar a cuestionar, de forma más general, la legitimidad de toda una familia de argumentos asociada a dicho warrant. Para justificar el warrant que es cuestionado surge el *backing*, el cual muestra por qué el warrant asociado es válido. Finalmente, la presencia de un *derrotador* indica una condición de excepción para el argumento propuesto, la cual puede conducir a la derrota del mismo.

La Figura 2.9 ejemplifica los elementos presentes en el esquema ilustrado por la Figura 2.8, utilizando un ejemplo propuesto por Toulmin que versa acerca de la nacionalidad de un sujeto llamado Harry. En este caso, el dato se halla representado por la información que establece que *Harry nació en las Bermudas*, mientras que la conclusión calificada del argumento indica que *presumiblemente, Harry es Británico*. En particular, el paso de inferencia entre los datos y conclusión del argumento está dado por el warrant, el cual enuncia la regla de que *un hombre nacido en las Bermudas generalmente es Británico*.

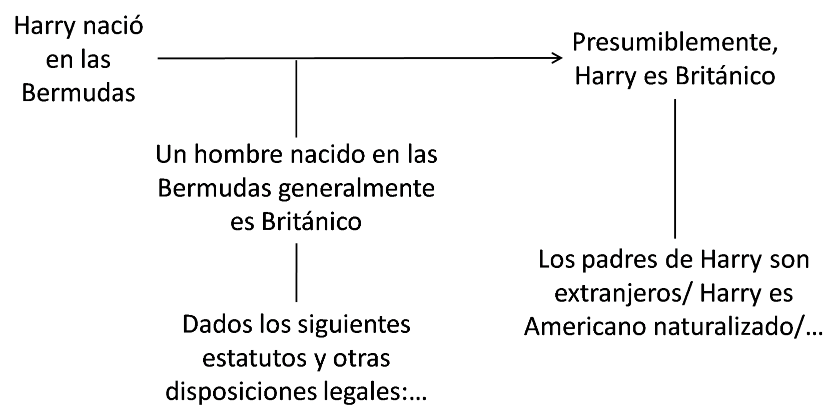


Figura 2.9: Un ejemplo de argumento acorde al modelo propuesto por Toulmin.

Luego, el backing expresa la *existencia de estatutos y otras disposiciones legales* que avalan la regla establecida por el warrant, brindándole así el correspondiente soporte. Por último, la información de que *los padres de Harry son extranjeros* o que *Harry es un Americano naturalizado* provee condiciones de excepción para el argumento, dando lugar a posibles derrotadores para el mismo.

A partir del esquema propuesto por Toulmin se pueden distinguir dos tipos de interacciones entre sus elementos. En primer lugar, además de los datos que soportan la conclusión, el backing provee soporte para el warrant. En segundo lugar, la presencia de un derrotador conlleva el rechazo de la conclusión mediante la derrota del argumento. Esto muestra que la noción de *soporte* ha estado presente en la literatura de argumentación desde sus trabajos fundacionales. Sin embargo, estudios subsiguientes en el área dejaron de lado la noción de soporte para concentrarse en la noción de derrota, la cual constituye un elemento clave en el cálculo de aceptabilidad de los argumentos de un sistema. Por ejemplo, en [Dun95] (ver Sección 2.2.1) se considera únicamente la relación de derrota entre argumentos para efectuar el cálculo de aceptabilidad, sin contemplar de forma alguna la existencia de soporte.

Siguiendo la línea de trabajo de Dung, gran parte de los trabajos en la literatura se han dedicado a formalismos argumentativos cuyo foco se encuentra en la relación de derrota. No obstante esto, en la última década, el estudio de la noción de soporte recobró atención entre los investigadores del área. El contar con una noción de soporte entre argumentos es de gran utilidad en sistemas que utilizan argumentación como herramienta para la toma de decisiones humanas en contextos con incertidumbre, tales como CAPSULE [WGM+97] y RAGs [CGFE01]. En particular, esto resulta de especial interés cuando se emplea argumentación para construir herramientas que combinan información de diferentes fuentes que no son totalmente confiables, etiquetando a las conclusiones con el grado de confianza que puede depositarse en ellas [PSM11, TCM+12, TGG+12].

En trabajos como [WGM+97, CGFE01] se observa que al evaluar la evidencia disponible, los seres humanos no sólo identifican conflictos entre argumentos, sino que también identifican situaciones en que los argumentos se *soportan* entre sí. Por lo tanto, en entornos como estos, resulta conveniente contar con una noción de soporte entre argumentos. Para ilustrar esto, consideremos el siguiente ejemplo inspirado en [PV02] y [CLS09]:

Ejemplo 2.18 *Consideremos el siguiente conjunto de argumentos intercambiados durante una reunión de la junta editorial de un periódico:*

$\mathcal{P}$: *Asumiendo que hay libertad de expresión, la información I acerca de la persona X debe ser publicada.*

$\mathcal{NP}$: *La persona X considera que la información I es privada y que, por lo tanto, no debe publicarse.*

$\mathcal{VP}$: *La información I concierne la salud de la persona X. Información relacionada con la salud de una persona es información acerca de su vida privada. Por lo tanto, I es información acerca de la vida privada de la persona X.*

$\mathcal{JG}$: *X es el jefe de gobierno de la ciudad. Por lo tanto, cualquier información relacionada con X es de interés público.*

Claramente, durante la discusión surgen algunos conflictos entre los argumentos expuestos. Tal es el caso del conflicto entre los argumentos $\mathcal{P}$ y $\mathcal{NP}$, y entre los argumentos $\mathcal{JG}$ y $\mathcal{NP}$. Por otra parte, existe una relación entre los argumentos $\mathcal{NP}$ y $\mathcal{VP}$ la cual, claramente, no es de conflicto. Más aún, el argumento $\mathcal{VP}$ provee nueva información, reforzando el argumento $\mathcal{NP}$.

Si quisiéramos representar el escenario ilustrado en el Ejemplo 2.18 utilizando un marco argumentativo como el propuesto por Dung, necesitaríamos encontrar una forma de representar la relación entre los argumentos $\mathcal{NP}$ y $\mathcal{VP}$. Dado que el formalismo de Dung simplemente considera un tipo de interacción entre argumentos (la relación de derrota), el soporte que el argumento $\mathcal{VP}$ brinda al argumento $\mathcal{NP}$ sólo podría ser representado mediante la noción de reinstalación (ver Sección 2.1). Sin embargo, el hecho de que $\mathcal{VP}$ reinstale al argumento $\mathcal{NP}$ a través de una derrota sobre el argumento $\mathcal{P}$ o sobre el argumento $\mathcal{JG}$ no resulta intuitivo, dado que no hay razones para que el argumento $\mathcal{VP}$ esté en conflicto con los argumentos $\mathcal{P}$ o $\mathcal{JG}$. En consecuencia, la consideración del argumento $\mathcal{VP}$ conduciría a modificar el argumento $\mathcal{NP}$ o encontrar una solución más intuitiva para representar la interacción entre los argumentos $\mathcal{VP}$ y $\mathcal{NP}$. Una alternativa para lograr esto consiste en considerar una nueva relación entre argumentos, independiente de la relación de derrota: una *relación de soporte*. Esto se debe a que la idea en un sistema argumentativo no es modificar los argumentos ya existentes sino, en la medida de lo posible, tratar de representar todos los tipos de interacciones entre los argumentos disponibles.

Recientemente, diversos trabajos en la literatura han abordado el estudio de la noción de soporte, proponiendo diferentes interpretaciones para la misma [CGGS13b, CLS13]. Por ejemplo, en [CLS05] los autores consideran una relación de *soporte general* que simplemente representa una interacción positiva entre argumentos, es decir, que no impone restricciones adicionales sobre los argumentos que relaciona. Otra interpretación posible es la de *soporte evidencial* [ON08], la cual permite distinguir entre argumentos *prima facie* y argumentos estándar. Los argumentos *prima facie* representan la noción de evidencia y no requieren del soporte de otros argumentos para subsistir, mientras que los argumentos estándar no pueden ser aceptados a menos que estén soportados por evidencia. En [BGvdTV10] los autores proponen una interpretación deductiva para la relación de soporte. De esta manera, el *soporte deductivo* intenta capturar la siguiente intuición: si un argumento $\mathcal{A}$ soporta a un argumento $\mathcal{B}$, entonces la aceptación de $\mathcal{A}$ implica la aceptación de $\mathcal{B}$ y, recíprocamente, la no aceptación de $\mathcal{B}$ implica la no aceptación de $\mathcal{A}$. Por otra parte, en [NR10] se introdujo por primera vez una relación de *soporte por necesidad*. El soporte por necesidad refuerza la siguiente restricción: si un argumento $\mathcal{A}$ soporta a un argumento $\mathcal{B}$, esto significa que $\mathcal{A}$ es necesario para $\mathcal{B}$. Por lo tanto, la aceptación de $\mathcal{B}$ implica la aceptación de $\mathcal{A}$ y, de manera similar, la no aceptación de $\mathcal{A}$ implica la no aceptación de $\mathcal{B}$.

Como puede observarse, la mayor parte de la literatura aborda el estudio de la noción de soporte en un contexto de argumentación abstracta, considerando explícitamente una relación de soporte entre argumentos. Sin embargo, otras aproximaciones como [Ver03a] y [MGS06] consideran la existencia de soporte en un entorno diferente. En particular, DefLog [Ver02, Ver03a] constituye una aproximación a argumentación dialéctica que permite la representación de los elementos del esquema de Toulmin, así como también de los vínculos de soporte entre ellos [Ver05, Ver09]. Por otra parte, en [MGS06] se provee a los argumentos de suficiente estructura interna como para definir explícitamente una relación de sub-argumento entre ellos. De esta manera, la relación de sub-argumento representa el soporte que un argumento provee a sus super-argumentos.

En el Capítulo 3 se estudiarán diferentes propuestas existentes en la literatura que contemplan el uso de soporte en sistemas argumentativos. En particular, se analizarán las diferentes interpretaciones de soporte arriba mencionadas, efectuando un estudio comparativo entre ellas. Por otra parte, en los capítulos 4 y 5 se desarrollarán otras propuestas que contemplan una forma alternativa de soporte entre argumentos. En particular, estos capítulos introducirán sistemas argumentativos que permiten modelar el soporte existente entre las nociones de backing y warrant propuestas por Toulmin [Tou58]. Por una parte, el formalismo desarrollado en el Capítulo 4 abordará esta cuestión en un contexto de argumentación abstracta mientras que, por otra parte, en el Capítulo 5 se introducirá un SABR que captura estas nociones.

2.4. Resumen

En este capítulo se introdujeron nociones correspondientes a argumentación como proceso de razonamiento argumentativo. En primer lugar se brindó una perspectiva histórica del estudio de argumentación, resaltando la utilidad de los sistemas argumentativos en diferentes dominios de aplicación. Luego, en la Sección 2.1 se presentó la estructura conceptual introducida en [PV02], la cual caracteriza a los sistemas argumentativos mediante la definición de cinco elementos: un lenguaje lógico subyacente, la definición de la noción de argumento, la identificación de conflictos entre argumentos, la obtención de las derrotas entre argumentos y, finalmente, la determinación del estado de aceptabilidad de los argumentos. A partir de esta caracterización se analizó cómo la definición de estos elementos toma lugar durante los diferentes pasos del proceso argumentativo.

En la Sección 2.2 se presentaron dos alternativas para la formalización de sistemas argumentativos: sistemas argumentativos abstractos y sistemas argumentativos basados en reglas. En particular, los sistemas del primer tipo se abstraen de la mayoría de los elementos de la estructura conceptual de [PV02], para concentrarse en el análisis del estado de aceptación de los argumentos del sistema. En contraste, los sistemas del segundo tipo definen todos los elementos de la estructura conceptual, por lo que usualmente son identificados como sistemas argumentativos concretos. A partir de esta clasificación, se prosiguió a presentar ejemplos particulares de estos tipos de sistemas: los Marcos Argumentativos Abstractos (AF) [Dun95] y la Programación en Lógica Rebatible (DeLP) [GS04]. Cabe destacar que, particularmente, se optó por presentar estos sistemas dado que serán utilizados para los desarrollos de los capítulos 4 y 5 de esta tesis.

En la Sección 2.2.1 se introdujeron las nociones básicas correspondientes a los AF de Dung. Concretamente, se caracterizó la propuesta de Dung identificando qué elementos de la estructura conceptual de [PV02] son instanciados por este formalismo. Como se mencionó anteriormente, los AF se abstraen del lenguaje lógico de representación, del proceso de construcción de argumentos y de la identificación de conflictos entre ellos, para concentrarse en el cálculo de aceptabilidad de los argumentos. De esta manera, Dung define a un AF como un par ordenado donde el primer elemento es un conjunto de argumentos y el segundo elemento es una relación de derrota entre ellos. A partir de esta definición se introdujeron diferentes nociones como libertad de conflictos, aceptabilidad y admisibilidad, las cuales son utilizadas para definir las semánticas de aceptabilidad de los AF. De esta manera, el formalismo de Dung sigue una aproximación declarativa para el cómputo de aceptabilidad de los argumentos del sistema, adoptando un enfoque extensional. Finalmente, a partir de una semántica es posible identificar los conjuntos de argumentos aceptados del AF, los cuales corresponden a una extensión bajo esa semántica.

Por otra parte, en la Sección 2.2.2 se presentó DeLP, el formalismo propuesto en [GS04]. DeLP es un SABR que combina resultados de programación en lógica y argumentación rebatible. Dada su caracterización como un SABR, se mostró cómo DeLP instancia los diferentes elementos de la estructura conceptual presentada en la Sección 2.1. En primer lugar, se introdujo la sintaxis del lenguaje de representación a partir del cual se especifican los programas lógicos rebatibles. Luego, a partir de estos programas, se mostró cómo se efectúa la construcción de argumentos a partir de la noción de derivación. Dado el uso de la negación fuerte en la especificación de los programas, es posible obtener argumentos que se hallen en conflicto por sustentar conclusiones contradictorias. Para resolver estos conflictos DeLP utiliza un criterio de comparación modular, el cual asigna pesos a los argumentos involucrados en el conflicto. De esta manera, si un argumento ataca a otro y es más fuerte que el argumento atacado se produce una derrota. A partir de la consideración de todos los argumentos y las relaciones de derrota entre ellos, DeLP define un procedimiento de prueba dialéctico que permite determinar cuáles son los argumentos aceptados del sistema. Por último, esto permite identificar cuáles son las inferencias del sistema (literales garantizados a partir de un programa lógico rebatible), las cuales corresponden a conclusiones sustentadas por argumentos aceptados.

Finalmente, en la Sección 2.3 se presentó el modelo para el diseño de argumentos propuesto por Toulmin [Tou58] como motivador para la existencia de una relación de soporte entre argumentos. El trabajo de Toulmin fue uno de los más influyentes en los comienzos de las investigaciones en el área de argumentación. En su libro *"The Uses of Argument"* [Tou58], Toulmin promulgó la idea de que los argumentos debían ser analizados utilizando un formato más rico que la dicotomía de premisas y conclusión empleada en la lógica formal. De esta manera, propuso un modelo para el diseño de argumentos que, además de datos y conclusión, distingue cuatro elementos: warrant, backing, derrotador y calificador. A partir de este esquema es posible distinguir dos tipos de interacciones entre sus elementos. Por una parte, además de los datos que soportan a la conclusión, el backing provee soporte para el warrant. Por otra parte, la presencia de un derrotador conduce al rechazo de la conclusión mediante la derrota del argumento. Esto sugiere que la noción de soporte ha estado presente en la literatura de argumentación desde sus trabajos fundacionales. A pesar de esto, estudios subsiguientes en el área dejaron de lado la noción de soporte para concentrarse en la noción de derrota, siguiendo así la línea de trabajo de [Dun95]. No obstante esto, como se verá en el Capítulo 3, en la última década el estudio de la noción de soporte recobró atención entre los investigadores del área de argumentación, surgiendo así diversas propuestas que adoptan interpretaciones alternativas para este concepto.

Capítulo 3

Formalismos Argumentativos con Soporte: un análisis de la literatura

En las últimas décadas, gran parte de los trabajos en la literatura se ha dedicado al estudio de formalismos argumentativos que se enfocan en la existencia de una relación de derrota entre argumentos. Como se mencionó en el Capítulo 2, recientemente, el estudio de una relación de soporte entre argumentos recobró atención entre los investigadores del área. A partir de esto, la mayor parte de las investigaciones ha estado centrada en el estudio de la noción soporte en un contexto de argumentación abstracta, mediante la consideración explícita de una relación de soporte entre argumentos. Sin embargo, existen otras aproximaciones que consideran la noción de soporte en un entorno diferente.

Partiendo del trabajo realizado en [CGGS13b, CLS13], en este capítulo se realizará una revisión de diferentes interpretaciones para la noción de soporte propuestas en la literatura tales como *soporte deductivo*, *soporte por necesidad*, *soporte evidencial*, y *sub-argumento*, entre otras; las primeras correspondientes a aproximaciones de argumentación abstracta, mientras que la última se encuentra asociada a un formalismo argumentativo que brinda parcialmente una estructura interna para los argumentos. Se efectuará un estudio comparativo, destacando similitudes y diferencias entre estas interpretaciones, y se discutirá cómo son abordadas por diferentes formalismos argumentativos. En particular, el análisis se enfocará en el significado asociado a cada interpretación de soporte, teniendo en cuenta las restricciones de aceptabilidad impuestas por cada una de ellas.

3.1. Marcos Argumentativos Bipolares (BAF)

La noción de bipolaridad ha sido ampliamente estudiada en diferentes dominios tales como la representación de conocimiento y preferencias [BDKP02, DP05], y la toma de decisiones [ABP05, DFB08]. En [ACLS04] se introdujeron por primera vez los *Marcos Argumentativos Bipolares (BAF)* (por su nombre en inglés, *Bipolar Argumentation Frameworks*). En dicho trabajo, profundizado en [ACLSL08], los autores estudiaron el uso de bipolaridad en argumentación, analizando cómo esta noción se manifiesta de diferentes formas en cada paso del proceso de razonamiento argumentativo. En particular, como parte del análisis de la existencia de bipolaridad a nivel de interacción entre argumentos, los autores presentaron una primera aproximación a los BAF. Básicamente, los BAF constituyen una extensión a los Marcos Argumentativos Abstractos de [Dun95] que consideran dos interacciones independientes entre argumentos, las cuales poseen una naturaleza opuesta: una relación de derrota y una relación de soporte. Una posterior formalización de los BAF fue desarrollada en [CLS05], donde la relación de soporte es tenida en cuenta para el análisis de aceptabilidad de los argumentos del sistema.

Definición 3.1 (Marco Argumentativo Bipolar) *Un* Marco Argumentativo Bipolar (BAF) *es una tupla* $\langle \mathbb{A}, \mathbb{R}_d, \mathbb{R}_s \rangle$*, donde* $\mathbb{A}$ *es un conjunto finito no vacío de argumentos,* $\mathbb{R}_d \subseteq \mathbb{A} \times \mathbb{A}$ *es una relación de derrota entre argumentos, y* $\mathbb{R}_s \subseteq \mathbb{A} \times \mathbb{A}$ *es una relación de soporte entre argumentos.*

Nótese que la relación $\mathbb{R}_d$ de los BAF es la misma que en los Marcos Argumentativos Abstractos propuestos por Dung. Por lo tanto, como se mencionó anteriormente, nos referiremos a ella como relación de derrota.[1] Un BAF puede representarse gráficamente mediante un grafo dirigido llamado *grafo de interacción bipolar*, donde los nodos son argumentos y hay dos tipos de arcos. En [CLS05] los autores utilizan $\not\longrightarrow$ y $\longrightarrow$ para denotar, respectivamente, derrota y soporte entre argumentos. No obstante esto, con el objetivo de proveer una representación unificada, en esta tesis seguiremos la convención usual de notación para los marcos argumentativos de Dung, en la cual la flecha sólida $\longrightarrow$ denota la relación de derrota. Por otra parte, denotaremos la relación de soporte entre argumentos

[1]Los autores utilizan la terminología 'ataque'. Sin embargo, al igual que lo expresado anteriormente para los Marcos Argumentativos Abstractos de Dung, para evitar confusiones nos referiremos a ella como relación de 'derrota'.

$$\begin{tikzcd} \mathcal{VP} \arrow[r, Rightarrow, "s"] & \mathcal{NP} \arrow[r, shift left] & \mathcal{P} \arrow[l, shift left] \\ & \mathcal{JG} \arrow[u] & \end{tikzcd}$$

Figura 3.1: Grafo de interacción bipolar para el $\mathsf{BAF}_{3.1}$ del Ejemplo 3.1.

utilizando una flecha doble $\Longrightarrow$. De aquí en adelante utilizaremos esta notación para distinguir las relaciones de derrota y soporte. Adicionalmente, incorporaremos una etiqueta sobre la flecha doble $\Longrightarrow$, la cual identificará la interpretación adoptada para la relación de soporte.

Dado que la relación de soporte de un BAF es simplemente una interacción positiva entre argumentos, sin imponer restricciones particulares, utilizaremos la etiqueta 's' para indicar que se trata de una relación de soporte general. De esta manera, la relación de soporte de un BAF será denotada mediante $\overset{s}{\Longrightarrow}$. Como se mostrará más adelante, otros formalismos adoptan interpretaciones alternativas de soporte, con lo cual se empleará una etiqueta diferente para distinguir cada una de ellas.

Ejemplo 3.1 *Consideremos la discusión durante la junta editorial de un periódico presentada en el Ejemplo 2.18 (ver pág. 52). Los argumentos allí presentados y sus interacciones pueden representarse mediante el* $\mathsf{BAF}_{3.1} = \langle \mathbb{A}_{3.1}, \mathbb{R}_{d_{3.1}}, \mathbb{R}_{s_{3.1}} \rangle$, *donde:*

$$\begin{aligned} \mathbb{A}_{3.1} &= \{\mathcal{P}, \mathcal{NP}, \mathcal{VP}, \mathcal{JG}\} \\ \mathbb{R}_{d_{3.1}} &= \{(\mathcal{P}, \mathcal{NP}), (\mathcal{NP}, \mathcal{P}), (\mathcal{JG}, \mathcal{NP})\} \\ \mathbb{R}_{s_{3.1}} &= \{(\mathcal{VP}, \mathcal{NP})\} \end{aligned}$$

La Figura 3.1 ilustra el grafo de interacción bipolar asociado a $\mathsf{BAF}_{3.1}$, *donde el argumento* $\mathcal{VP}$ *soporta al argumento* $\mathcal{NP}$, *el argumento* $\mathcal{JG}$ *derrota al argumento* $\mathcal{NP}$, *y los argumentos* $\mathcal{P}$ *y* $\mathcal{NP}$ *se derrotan entre sí.*

Con el objeto de considerar la interacción entre argumentos relacionados que soportan y atacan a otros argumentos, en [CLS05] se introdujeron las nociones de *derrota soportada* y *derrota secundaria*[2] (en inglés, *supported defeat* y *secondary defeat* respectivamente),

[2]En [CLS05] los autores emplean el término derrota 'indirecta' (indirect defeat). Sin embargo, en trabajos posteriores como [CLS11] adoptaron el término derrota 'secundaria'.

las cuales combinan una secuencia de soportes con una derrota directa[3]. En adelante, siguiendo la terminología introducida en [CLS11, CLS13] nos referiremos a estas derrotas como *derrotas complejas*.

Definición 3.2 (Derrota Soportada y Derrota Secundaria) *Sea* $\langle \mathbb{A}, \mathbb{R}_d, \mathbb{R}_s \rangle$ *un* *BAF y* $\mathcal{A}, \mathcal{B} \in \mathbb{A}$.

- *Una* derrota soportada *de* $\mathcal{A}$ *a* $\mathcal{B}$ *es una secuencia* $\mathcal{A}_1\mathcal{R}_1 \ldots \mathcal{R}_{n-1}\mathcal{A}_n, n \geq 3$*, donde* $\mathcal{A}_1 = \mathcal{A}$*,* $\mathcal{A}_n = \mathcal{B}$*, tal que* $\forall i = 1 \ldots n-2$*,* $\mathcal{R}_i = \mathbb{R}_s$ *y* $\mathcal{R}_{n-1} = \mathbb{R}_d$.

- *Una* derrota secundaria *de* $\mathcal{A}$ *a* $\mathcal{B}$ *es una secuencia* $\mathcal{A}_1\mathcal{R}_1 \ldots \mathcal{R}_{n-1}\mathcal{A}_n, n \geq 3$*, donde* $\mathcal{A}_1 = \mathcal{A}$*,* $\mathcal{A}_n = \mathcal{B}$*, tal que* $\mathcal{R}_1 = \mathbb{R}_d$ *y* $\forall i = 2 \ldots n-1$*,* $\mathcal{R}_i = \mathbb{R}_s$.

En [CLS05] los autores enuncian que, por extensión, una secuencia $\mathcal{A}\ \mathbb{R}_d\ \mathcal{B}$ reducida a dos argumentos (*i. e.*, una derrota directa $\mathcal{A} \longrightarrow \mathcal{B}$) es también considerada como una derrota soportada de $\mathcal{A}$ sobre $\mathcal{B}$. Por lo tanto, dado un BAF $\langle \mathbb{A}, \mathbb{R}_d, \mathbb{R}_s \rangle$ y los argumentos $\mathcal{A}_1, \ldots, \mathcal{A}_k, \mathcal{B} \in \mathbb{A}$, un camino $\mathcal{A}_1 \overset{s}{\Longrightarrow} \ldots \overset{s}{\Longrightarrow} \mathcal{A}_k \longrightarrow \mathcal{B}$ en el grafo de interacción bipolar conduce a k derrotas soportadas de $\mathcal{A}_i$ sobre $\mathcal{B}$ respectivamente ($1 \leq i \leq k$). De manera similar, un camino $\mathcal{B} \longrightarrow \mathcal{A}_1 \overset{s}{\Longrightarrow} \ldots \overset{s}{\Longrightarrow} \mathcal{A}_k$ en el grafo de interacción bipolar conduce a $k-1$ derrotas secundarias de $\mathcal{B}$ sobre $\mathcal{A}_j$ respectivamente ($2 \leq j \leq k$). Por ejemplo, en el $\mathsf{BAF}_{3.1}$ del Ejemplo 3.1 existe una derrota soportada del argumento $\mathcal{VP}$ sobre el argumento $\mathcal{P}$, determinada por el camino $\mathcal{VP} \overset{s}{\Longrightarrow} \mathcal{NP} \longrightarrow \mathcal{P}$ en el grafo de interacción bipolar de $\mathsf{BAF}_{3.1}$. Por otra parte, las derrotas directas $\mathcal{JG} \longrightarrow \mathcal{NP}$, $\mathcal{NP} \longrightarrow \mathcal{P}$ y $\mathcal{P} \longrightarrow \mathcal{NP}$ son también derrotas soportadas.

Ejemplo 3.2 *Consideremos el marco argumentativo bipolar* $\mathsf{BAF}_{3.2}$ *caracterizado por el grafo de interacción bipolar que se ilustra en la Figura 3.2(a). Las derrotas directas del* $\mathsf{BAF}_{3.2}$ *se ilustran en la Figura 3.2(b) utilizando flechas sólidas (*$\longrightarrow$*). Por otra parte, las derrotas secundarias y las derrotas soportadas (aquellas que no son además derrotas directas) se encuentran ilustradas en la Figura 3.2(b) mediante flechas punteadas (*$\dashrightarrow$*). En este caso, dado que* $\mathcal{A}$ *derrota a* $\mathcal{B}$*,* $\mathcal{B}$ *soporta a* $\mathcal{C}$*, y* $\mathcal{C}$ *soporta a* $\mathcal{D}$ *y* $\mathcal{E}$*, existen derrotas secundarias de* $\mathcal{A}$ *sobre* $\mathcal{C}, \mathcal{D}$ *y* $\mathcal{E}$*. Además, como* $\mathcal{E}$ *derrota a* $\mathcal{F}$*, existen derrotas soportadas de* $\mathcal{B}$ *y* $\mathcal{C}$ *sobre* $\mathcal{F}$*. Finalmente, nótese que, como* $\mathcal{F}$ *derrota a* $\mathcal{G}$ *y* $\mathcal{G}$ *soporta*

[3]Dado un BAF $= \langle \mathbb{A}, \mathbb{R}_d, \mathbb{R}_s \rangle$, una derrota directa está determinada por un par perteneciente a la relación de derrota $\mathbb{R}_d$.

(a) (b)

Figura 3.2: (a) Grafo de interacción bipolar de $\mathsf{BAF}_{3.2}$ y (b) las derrotas entre sus argumentos, correspondientes al Ejemplo 3.2.

a $\mathcal{F}$, existe una derrota soportada de $\mathcal{G}$ sobre sí mismo y una derrota secundaria de $\mathcal{F}$ sobre sí mismo.

Como se puede observar en el Ejemplo 3.2 si $\mathcal{A} \overset{s}{\Longrightarrow} \mathcal{B}$ y $\mathcal{B} \longrightarrow \mathcal{A}$ (es decir, si los argumentos $\mathcal{A}$ y $\mathcal{B}$ se soportan y derrotan mutuamente), entonces $\mathcal{A}$ y $\mathcal{B}$ son argumentos auto-derrotados (en inglés, *self-defeating arguments*) ya que existe una derrota soportada de $\mathcal{A}$ sobre sí mismo y una derrota secundaria de $\mathcal{B}$ sobre sí mismo. Esto ocurre porque en la especificación de un BAF no se impone restricción alguna que prevenga la existencia de argumentos que se derrotan y soportan mutuamente. No obstante esto, los BAF introducen algunas restricciones con respecto al soporte y derrota entre argumentos, explicadas a continuación.

Siguiendo la aproximación de [Dun95], en [CLS05] los autores definen las características un conjunto de argumentos aceptados de un BAF debería satisfacer. Por una parte, consideran que un conjunto aceptable de argumentos debe ser internamente coherente, es decir, no debe contener argumentos que se derroten entre sí. Para lograr esto extienden la noción de libertad de conflictos propuesta por [Dun95], dando lugar a la noción de *libertad de conflictos*$^{+}$ que considera la existencia de derrotas soportadas y derrotas secundarias. Por otra parte, dado que los BAF permiten la representación de soporte entre argumentos además de derrotas, consideran que un conjunto aceptable de argumentos debe ser también externamente coherente: un conjunto de argumentos que derrota y soporta al mismo argumento no puede ser aceptable. Esta noción de coherencia externa es capturada por los BAF mediante la noción de *seguridad* (en inglés, *safety*), como se indica a continuación.

Definición 3.3 (Libertad de Conflictos^{+} y Seguridad) *Sea $\langle \mathbb{A}, \mathbb{R}_d, \mathbb{R}_s \rangle$ un BAF y $S \subseteq \mathbb{A}$.*

- *S es un* conjunto libre de conflictos$^+$ *si y sólo si* $\nexists \mathcal{A}, \mathcal{B} \in S$ *tal que existe una derrota soportada o una derrota secundaria de* $\mathcal{A}$ *sobre* $\mathcal{B}$*.*

- *S es un* conjunto seguro *si y sólo si* $\nexists \mathcal{A} \in \mathbb{A}$*,* $\nexists \mathcal{B}, \mathcal{C} \in S$ *tal que existe una derrota soportada o una derrota secundaria de* $\mathcal{B}$ *sobre* $\mathcal{A}$*, y o bien existe una secuencia de soporte de* $\mathcal{C}$ *hacia* $\mathcal{A}$ *o* $\mathcal{A} \in S$*.*

El símbolo '+' en la definición anterior expresa que para verificar si un conjunto es *libre de conflictos*$^+$ deben tenerse en cuenta las derrotas soportadas y las derrotas secundarias, además de las derrotas directas consideradas por la noción clásica de *libertad de conflictos* definida en [Dun95]. De esta manera, la noción de *libertad de conflictos*$^+$ es más restrictiva que aquella propuesta por Dung. Para ilustrar esto, consideremos el $\mathsf{BAF}_{3.2}$ del Ejemplo 3.2. El conjunto $\{\mathcal{A}, \mathcal{F}\}$ es tanto *libre de conflictos* como *libre de conflictos*$^+$. Por el contrario, el conjunto $\{\mathcal{A}, \mathcal{C}\}$ es *libre de conflictos* pero no *libre de conflictos*$^+$, ya que existe una derrota secundaria de $\mathcal{A}$ sobre $\mathcal{C}$. De manera similar, el conjunto $\{\mathcal{G}\}$ es *libre de conflictos* pero no es *libre de conflictos*$^+$, dado que $\mathcal{G}$ se derrota a sí mismo a través de una derrota soportada. Por último, el conjunto $\{\mathcal{B}, \mathcal{C}, \mathcal{D}, \mathcal{E}, \mathcal{H}\}$ es *libre de conflictos*$^+$ pero no es *seguro*, ya que $\mathcal{E}$ derrota a $\mathcal{F}$ y $\mathcal{H}$ soporta a $\mathcal{F}$. Luego, si consideramos los argumentos $\mathcal{E}$ y $\mathcal{H}$ por separado, los conjuntos $\{\mathcal{B}, \mathcal{C}, \mathcal{D}, \mathcal{E}\}$ y $\{\mathcal{B}, \mathcal{C}, \mathcal{D}, \mathcal{H}\}$ son *seguros*. En [CLS05] los autores muestran que la noción de *seguridad* abarca la noción de *libertad de conflictos*$^+$ (*i. e.*, si un conjunto es *seguro*, entonces también es *libre de conflictos*$^+$). Claramente la inversa no vale, como se puede observar en el ejemplo anterior. Sin embargo, en [CLS05] los autores también muestran que si un conjunto S es *libre de conflictos*$^+$ y cerrado bajo $\mathbb{R}_s$ (*i. e.*, si $\mathcal{A}\mathbb{R}_s\mathcal{B}$ y $\mathcal{A} \in S$, entonces $\mathcal{B} \in S$), entonces S es también un conjunto *seguro*.

Siguiendo la aproximación de Dung, en [CLS05] se define la *aceptabilidad* de un argumento $\mathcal{A}$ con respecto a un conjunto de argumentos S requiriendo que los todos los derrotadores de $\mathcal{A}$ sean derrotados por S. Luego, las extensiones preferidas de un BAF se definen como conjuntos admisibles maximales con respecto a la inclusión de conjuntos. Luego, con la idea de reforzar la coherencia de los conjuntos admisibles, en [CLS05] se definen tres nociones de admisibilidad, desde la más general hacia la más específica. Primero, al requerir que un conjunto admisible sea libre de conflictos$^+$ y que defienda a todos sus elementos, introducen la noción de *d-admisibilidad* ('*d*' por admisibilidad *a la Dung*). Segundo, la noción de *s-admisibilidad* ('*s*' por *segura*) refuerza la d-admisibilidad al requerir que los conjuntos admisibles sean seguros. Finalmente, se refuerza la coherencia externa al requerir que los conjuntos admisibles sean cerrados bajo la relación de soporte $\mathbb{R}_s$, dando lugar a la noción de *c-admisibilidad* ('*c*' por *cerrada*).

Diferentes nociones de admisibilidad pueden dar lugar a diferentes extensiones de un BAF. Por ejemplo, dado el $\mathsf{BAF}_{3.2}$ del Ejemplo 3.2, el conjunto $\{\mathcal{B}, \mathcal{C}, \mathcal{D}, \mathcal{E}, \mathcal{H}\}$ no es d-admisible ya que, si bien es libre de conflictos$^+$, no defiende a $\mathcal{B}, \mathcal{C}, \mathcal{D}$ y $\mathcal{E}$ de las derrotas de $\mathcal{A}$. Luego, la única extensión d-preferida de $\mathsf{BAF}_{3.2}$ es $\{\mathcal{A}, \mathcal{H}\}$. Dado que $\{\mathcal{A}, \mathcal{H}\}$ es un conjunto seguro, $\{\mathcal{A}, \mathcal{H}\}$ es también la única extensión s-preferida de $\mathsf{BAF}_{3.2}$. Finalmente, el conjunto $\{\mathcal{A}, \mathcal{H}\}$ no es c-admisible ya que $\mathcal{H} \stackrel{s}{\Longrightarrow} \mathcal{F}$, y $\mathcal{F}$ no pertenece al conjunto. En consecuencia, la única extensión c-preferida de $\mathsf{BAF}_{3.2}$ es $\{\mathcal{A}\}$.

3.1.1. Una Alternativa al Cálculo de Aceptabilidad para los Marcos Argumentativos Bipolares utilizando Coaliciones

En [CLS07, CLS10] los autores presentan una aproximación alternativa para el manejo de bipolaridad en marcos argumentativos abstractos. La idea es transformar un BAF en un marco argumentativo al estilo de [Dun95] compuesto de un conjunto de coaliciones y una relación de derrota entre ellas. Las derrotas originalmente consideradas por el BAF solamente aparecerán a nivel de coaliciones (*i. e.*, no habrá derrotas entre argumentos sino derrotas entre coaliciones); como consecuencia, una coalición agrupará argumentos no conflictivos. Por lo tanto, la relación de soporte del BAF inicial no aparecerá a nivel de coaliciones, sino que será utilizada para agrupar argumentos en coaliciones. Como se menciona en [CLS10], los dos principios fundamentales que gobiernan la definición de una coalición son el *principio de coherencia* (las coaliciones deben ser libre de conflictos) y el *principio de soporte* (los argumentos en una coalición deben estar directa o indirectamente relacionados por la relación de soporte).

Sea $\mathsf{BAF} = \langle \mathbb{A}, \mathbb{R}_d, \mathbb{R}_s \rangle$ un marco argumentativo bipolar cuyo grafo de interacción bipolar es $\mathcal{G}_{\mathsf{BAF}}$. El grafo que representa el marco argumentativo parcial $\langle \mathbb{A}, \mathbb{R}_s \rangle$ (*i. e.*, el marco compuesto por el conjunto de argumentos y la relación de soporte del BAF) es denotado por $\mathcal{G}_s$. Similarmente, $\mathsf{AF}_{\mathsf{BAF}}$ denota el marco argumentativo abstracto $\langle \mathbb{A}, \mathbb{R}_d \rangle$ asociado al BAF (*i. e.*, el marco argumentativo compuesto por el conjunto de argumentos y la relación de derrota del BAF).

Definición 3.4 (Coalición) *Sea* $\mathsf{BAF} = \langle \mathbb{A}, \mathbb{R}_d, \mathbb{R}_s \rangle$ *un marco argumentativo bipolar.* $C \subseteq \mathbb{A}$ *es una* coalición *de* BAF *si y sólo si es un conjunto maximal libre de conflictos para* $\mathsf{AF}_{\mathsf{BAF}}$ *tal que el sub-grafo de* $\mathcal{G}_s$ *inducido por* C *es conexo.*

Resumidamente, el sub-grafo de $\mathcal{G}_s$ inducido por C consiste de los argumentos de C y los arcos de $\mathcal{G}_s$ que conectan a argumentos de C. Adicionalmente, este sub-grafo inducido se dice conexo si y sólo si existe un camino (dirigido o no) de soporte entre cualquier par de argumentos del sub-grafo [Ber01]. Por ejemplo, dado el $\mathsf{BAF}_{3.2}$ del Ejemplo 3.2, se obtienen las coaliciones $C_1 = \{\mathcal{A}\}$, $C_2 = \{\mathcal{B}, \mathcal{C}, \mathcal{D}, \mathcal{E}, \mathcal{G}\}$, y $C_3 = \{\mathcal{F}, \mathcal{H}\}$.

La definición 3.4 exige que las coaliciones sean libre de conflictos para $\mathsf{AF}_{\mathsf{BAF}}$. Esto significa que no existen argumentos $\mathcal{A}$ y $\mathcal{B}$ en una coalición C tales que $\mathcal{A}$ derrota directamente a $\mathcal{B}$. Es decir, la especificación de las coaliciones emplea la noción de libertad de conflictos clásica propuesta por [Dun95]. En general, esto evita que argumentos conflictivos pertenezcan a la misma coalición. Sin embargo, si se considerasen las derrotas soportadas y las derrotas secundaras, una coalición podría no ser *libre de conflictos*$^+$. Tal es el caso de las coaliciones C_2 y C_3 obtenidas a partir del $\mathsf{BAF}_{3.2}$ del Ejemplo 3.2, las cuales no son *libre de conflictos*$^+$ ya que $\mathcal{G}$ y $\mathcal{F}$ son argumentos que se derrotan a sí mismos.

Sea $\mathsf{BAF} = \langle \mathbb{A}, \mathbb{R}_d, \mathbb{R}_s \rangle$ un marco argumentativo bipolar y $C(\mathbb{A})$ el conjunto de coaliciones de BAF. En [CLS07] se definen las coaliciones conflictivas como sigue: dos coaliciones C_1 y C_2 se encuentran en conflicto, denotado como C_1 *c-derrota* C_2, si $\exists \mathcal{A}_1 \in C_1, \exists \mathcal{A}_2 \in C_2$ tal que $\mathcal{A}_1 \mathbb{R}_d \mathcal{A}_2$. Dado el conjunto de coaliciones $C(\mathbb{A})$ y la relación *c-derrota*, se define un marco argumentativo $\mathsf{CAF} = \langle C(\mathbb{A}), \textit{c-derrota} \rangle$, llamado *Marco Argumentativo de Coaliciones asociado al* BAF. Finalmente, en [CLS07] se propone una aproximación para computar las extensiones de un BAF en términos de las extensiones de su CAF asociado. En primer lugar, se obtienen las extensiones de un CAF siguiendo la metodología de [Dun95]. Luego, las diferentes coaliciones en una extensión del CAF son combinadas para obtener la correspondiente extensión del BAF. De esta manera, por ejemplo, al agrupar todos los argumentos pertenecientes a las coaliciones de una extensión preferida del CAF se obtiene una *cp-extensión* (extensión preferida de coaliciones) del BAF.

Ejemplo 3.3 *Dado el* $\mathsf{BAF}_{3.2}$ *del Ejemplo 3.2 es posible obtener el marco argumentativo de coaliciones asociado* $\mathsf{CAF}_{3.3}$. *La Figura 3.3(a) ilustra las coaliciones obtenidas a partir de* $\mathsf{BAF}_{3.2}$, *mientras que el* $\mathsf{CAF}_{3.3}$ *se halla ilustrado en la Figura 3.3(b). En este caso, la*

(a) (b)

Figura 3.3: (a) Coaliciones obtenidas a partir del $\mathsf{BAF}_{3.2}$ del Ejemplo 3.2 y (b) su marco argumentativo de coaliciones asociado $\mathsf{CAF}_{3.3}$.

única extensión preferida de $\mathsf{CAF}_{3.3}$ *es* $\{C_1, C_3\}$. *Por lo tanto, la única cp-extensión de* $\mathsf{BAF}_{3.2}$ *es* $\{\mathcal{A}, \mathcal{F}, \mathcal{H}\}$.

Puede observarse que la cp-extensión del $\mathsf{BAF}_{3.2}$ en el Ejemplo 3.3 difiere de las d/s/c-extensiones preferidas obtenidas directamente a partir del $\mathsf{BAF}_{3.2}$. Por otra parte, como la relación *c-derrota* de un CAF no tiene en cuenta las derrotas soportadas y las derrotas secundarias del BAF asociado, las extensiones resultantes del CAF pueden contener argumentos que, en caso de considerar estas derrotas adicionales, estarían auto-derrotados. Tal es el caso del argumento $\mathcal{F}$ de $\mathsf{BAF}_{3.2}$, el cual pertenece a la extensión preferida de $\mathsf{CAF}_{3.3}$ (y correspondientemente a la cp-extensión de $\mathsf{BAF}_{3.2}$); sin embargo, como se mostró en el Ejemplo 3.2, el argumento $\mathcal{F}$ se derrota a sí mismo. En contraste, si las derrotas soportadas y las derrotas secundarias fueran consideradas al construir la relación de *c-derrota*, los argumentos auto-derrotados darían lugar a coaliciones auto-derrotadas, previniendo de esta manera que dichos argumentos pertenezcan en las extensiones del marco argumentativo.

En [CLS07, CLS10] los autores remarcan que los CAF preservan algunas propiedades demostradas para los marcos argumentativos de Dung. No obstante esto, también mencionan que otras propiedades de los marcos argumentativos de Dung no son verificadas por el formalismo que ellos proponen. Por ejemplo, como se muestra en el siguiente ejemplo, puede darse el caso de que una cp-extensión de un BAF no sea admisible.

Ejemplo 3.4 *Consideremos el* $\mathsf{BAF}_{3.4}$ *representado por el grafo de interacción bipolar que se ilustra en la Figura 3.4*(a). *A partir de* $\mathsf{BAF}_{3.4}$ *se obtienen las coaliciones* $C_1 = \{\mathcal{I}\}$, $C_2 = \{\mathcal{J}, \mathcal{K}, \mathcal{L}\}$ *y* $C_3 = \{\mathcal{M}\}$. *De esta manera, el marco argumentativo de coaliciones asociado es* $\mathsf{CAF}_{3.4} = \langle\{C_1, C_2, C_3\}, \{(C_1, C_2), (C_2, C_3)\}\rangle$, *ilustrado en la Figura 3.4*(b). *La*

$$\mathcal{I} \longrightarrow \mathcal{J} \overset{s}{\Longleftarrow} \mathcal{K} \overset{s}{\Longrightarrow} \mathcal{L} \longrightarrow \mathcal{M}$$

(a)

$$\begin{array}{ccccc} C_1 & \longmapsto & C_2 & \longmapsto & C_3 \\ \{\mathcal{I}\} & & \{\mathcal{J}, \mathcal{K}, \mathcal{L}\} & & \{\mathcal{M}\} \end{array}$$

(b)

Figura 3.4: (a) Grafo de interacción bipolar del $\mathsf{BAF}_{3.4}$ correspondiente al Ejemplo 3.4 y (b) su marco argumentativo de coaliciones asociado $\mathsf{CAF}_{3.4}$.

única extensión preferida de $\mathsf{CAF}_{3.4}$ *es* $\{C_1, C_3\}$ *y, por lo tanto, la única cp-extensión de* $\mathsf{BAF}_{3.4}$ *es* $\{\mathcal{I}, \mathcal{M}\}$. *Sin embargo, tenemos que* $\mathcal{L} \longrightarrow \mathcal{M}$ *pero* $\{\mathcal{I}, \mathcal{M}\}$ *no defiende a* $\mathcal{M}$ *de la derrota de* $\mathcal{L}$ *(ya sea a través de una derrota directa, soportada o secundaria). En contraste, tenemos que el argumento* $\mathcal{I}$ *derrota al argumento* $\mathcal{J}$, *el cual pertenece a la misma coalición que derrota a* $\mathcal{M}$. *Por lo tanto,* $\{\mathcal{I}, \mathcal{M}\}$ *no es admisible en el sentido propuesto por Dung.*

Los autores en [CLS07, CLS10] enuncian que las coaliciones deben ser consideradas como un todo y, en consecuencia, sus elementos no deben utilizarse por separado en el proceso de derrota. De esta manera, establecen que la pérdida de admisibilidad, como ocurre en el Ejemplo 3.4, no es un problema ya que surge de considerar derrotas entre argumentos y no entre coaliciones. En todo caso, consideran que la pérdida de admisibilidad en el Ejemplo 3.4 está asociada al tamaño de la coalición $C_2 = \{\mathcal{J}, \mathcal{K}, \mathcal{L}\}$, la cual fue construida sin tener en cuenta la dirección de los caminos de soporte en el grafo de interacción bipolar. Es por esto que proponen la noción de *coalición elemental*, la cual es obtenida a partir de caminos de soporte maximales y libres de conflicto. En consecuencia, dado el $\mathsf{BAF}_{3.4}$ del Ejemplo 3.4, se obtendrían dos coaliciones elementales independientes $\{\mathcal{J}, \mathcal{K}\}$ y $\{\mathcal{K}, \mathcal{L}\}$ en lugar de la coalición original $\{\mathcal{J}, \mathcal{K}, \mathcal{L}\}$. Sin embargo, como fue notado por los autores, y como puede observarse en el siguiente ejemplo, las coaliciones elementales no siempre permiten recuperar las propiedades perdidas de los marcos argumentativos de Dung.

Ejemplo 3.5 *Sea* $\mathsf{BAF}_{3.5}$ *el marco argumentativo bipolar caracterizado por el grafo de interacción bipolar ilustrado en la Figura 3.5*(a). *Las coaliciones elementales obtenidas a partir de* $\mathsf{BAF}_{3.5}$ *son* $EC_1 = \{\mathcal{I}\}$, $EC_2 = \{\mathcal{J}, \mathcal{K}, \mathcal{L}\}$, *y* $EC_3 = \{\mathcal{M}\}$. *El marco argumentativo de coaliciones asociado es* $\mathsf{CAF}_{3.5} = \langle \{EC_1, EC_2, EC_3\}, \{(EC_1, EC_2), (EC_2, EC_3)\} \rangle$,

$$\mathcal{I} \longrightarrow \mathcal{J} \overset{s}{\Longleftarrow} \mathcal{K} \overset{s}{\Longleftarrow} \mathcal{L} \longrightarrow \mathcal{M}$$

(a)

$$\underset{\{\mathcal{I}\}}{EC_1} \longmapsto \underset{\{\mathcal{J},\mathcal{K},\mathcal{L}\}}{EC_2} \longmapsto \underset{\{\mathcal{M}\}}{EC_3}$$

(b)

Figura 3.5: (a) Grafo de interacción bipolar del $\mathsf{BAF}_{3.5}$ correspondiente al Ejemplo 3.5 y (b) su marco argumentativo de coaliciones asociado $\mathsf{CAF}_{3.5}$.

ilustrado en la Figura 3.5(b)*. En este caso, la única extensión preferida de* $\mathsf{CAF}_{3.5}$ *es* $\{EC_1, EC_3\}$*, por lo que la única cp-extensión de* $\mathsf{BAF}_{3.5}$ *es* $\{\mathcal{I}, \mathcal{M}\}$*. Sin embargo, al igual que lo ocurrido en el Ejemplo 3.4,* $\{\mathcal{I}, \mathcal{M}\}$ *no es admisible ya que no defiende a* $\mathcal{M}$ *ante la derrota de* $\mathcal{L}$*.*

Finalmente, en [CLS07, CLS10] los autores argumentan que la pérdida de admisibilidad como se muestra en los ejemplos 3.4 and 3.5 no es problemática. Como se mencionó anteriormente, ellos justifican su postura basándose en el hecho de que la noción de admisibilidad propuesta por Dung tiene en cuenta derrota y defensa "individual" mientras que, en su aproximación de coaliciones, las nociones de derrota y defensa deben considerarse de manera "colectiva". Como se mostrará a continuación, estas cuestiones han sido abordadas y criticadas por otras propuestas de la literatura que contemplan el uso de soporte en sistemas argumentativos.

3.2. Marcos Argumentativos Bipolares con Soporte Deductivo (d-BAF)

En [BGvdTV10] los autores introducen una aproximación que permite la consideración de soporte entre argumentos utilizando meta-argumentación. Resumidamente, meta-argumentación (o argumentación a meta-nivel) se refiere a que los argumentos en este tipo de sistemas versan sobre la aceptabilidad de argumentos pertenecientes a un sistema argumentativo de un nivel inferior [BGvdTV09, MBC11]. A diferencia de la relación de soporte de los BAF, en [BGvdTV09] se introduce la noción de *soporte deductivo* que brinda una interpretación particular para la relación de soporte. Brevemente, el soporte deductivo establece las siguientes restricciones sobre la aceptabilidad de los argumentos que relaciona: si un argumento $\mathcal{A}$ soporta a un argumento $\mathcal{B}$ y $\mathcal{A}$ está aceptado, entonces $\mathcal{B}$ también está aceptado; y si $\mathcal{B}$ no está aceptado, entonces $\mathcal{A}$ tampoco está aceptado.

Como es mencionado por los autores en [BGvdTV10], su propuesta aborda dos inconvenientes presentes en la aproximación de los CAF presentada en la Sección 3.1. Por una parte, atacan la pérdida de admisibilidad en el sentido de Dung mientras que, por otra parte, se refieren a la caracterización de derrotas en el contexto de una relación de soporte. Adicionalmente, enuncian que es posible introducir una relación de soporte entre argumentos sin necesidad de extender el formalismo propuesto por Dung en [Dun95]. De esta manera, proponen el uso de meta-argumentación para instanciar un marco argumentativo abstracto con el objetivo de representar soporte deductivo entre argumentos.

Como se mencionó en la Sección 3.1.1, los autores en [CLS07, CLS10] enuncian que la pérdida de admisibilidad, como ocurre en los ejemplos 3.4 y 3.5, no es problemática ya que toma en cuenta derrotas individuales. En contraste, los autores de [BGvdTV10] remarcan que su objetivo de utilizar meta-argumentación es preservar todas las propiedades y principios existentes para los marcos argumentativos de Dung, motivo por el cual eligen utilizar meta-argumentos en lugar de agrupar argumentos en coaliciones.

Continuando con la notación gráfica introducida en la sección anterior, $\mathcal{A} \longrightarrow \mathcal{B}$ denotará que el argumento $\mathcal{A}$ derrota al argumento $\mathcal{B}$. Al igual que en los BAF, la relación de derrota en la aproximación de [BGvdTV10] corresponde a la relación de derrota de los marcos argumentativos de Dung. Por otra parte, utilizaremos $\mathcal{A} \overset{d}{\Longrightarrow} \mathcal{B}$ para denotar que el argumento $\mathcal{A}$ *soporta deductivamente* al argumento $\mathcal{B}$. En este caso, la etiqueta 'd' sobre la flecha doble indica que la relación de soporte entre argumentos posee una interpretación deductiva.

Dada la coexistencia de argumentos que se derrotan y soportan entre sí, en [BGvdTV10] se consideran derrotas adicionales que resultan de combinar una secuencia de soportes con una derrota directa. Continuando con la terminología introducida en la Sección 3.1, identificaremos a estas derrotas como *derrotas complejas*. En particular, en [BGvdTV10] los autores proponen las *derrotas mediadas* (en inglés, *mediated defeats*), las cuales refuerzan las restricciones impuestas por el soporte deductivo. Si $\mathcal{A} \overset{d}{\Longrightarrow} \mathcal{B}$ y $\mathcal{C} \longrightarrow \mathcal{B}$, entonces ocurre una *derrota mediada* de $\mathcal{C}$ sobre $\mathcal{A}$. En esta tesis se utilizará el acrónimo d-BAF para denotar un marco argumentativo bipolar con soporte deductivo.

$$\begin{tikzcd}
\mathcal{I} \arrow[r] \arrow[rr, dashed, bend left] \arrow[rrr, dashed, bend right] & \mathcal{J} & \mathcal{K} \arrow[l, Rightarrow, "d"'] & \mathcal{L} \arrow[l, Rightarrow, "d"'] \arrow[r] & \mathcal{M}
\end{tikzcd}$$

Figura 3.6: Marco argumentativo bipolar con soporte deductivo $\textsf{d-BAF}_{3.6}$, correspondiente al Ejemplo 3.6.

Definición 3.5 (Derrota Mediada) *Sea* $\langle \mathbb{A}, \longrightarrow, \stackrel{d}{\Longrightarrow} \rangle$ *un d-BAF y sean* $\mathcal{A}, \mathcal{B} \in \mathbb{A}$. *Una* derrota mediada *de* $\mathcal{A}$ *sobre* $\mathcal{B}$ *es una secuencia* $\mathcal{A}_1 \stackrel{d}{\Longrightarrow} \ldots \stackrel{d}{\Longrightarrow} \mathcal{A}_{n-1}$ *y* $\mathcal{A}_n \longrightarrow \mathcal{A}_{n-1}$ $(n \geq 3)$*, donde* $\mathcal{A}_1 = \mathcal{B}$ *y* $\mathcal{A}_n = \mathcal{A}$.

El siguiente ejemplo ilustra cómo las derrotas mediadas permiten recuperar la pérdida de admisibilidad ocurrida en el Ejemplo 3.5.

Ejemplo 3.6 *Sea* $\textsf{BAF}_{3.5}$ *el marco argumentativo bipolar del Ejemplo 3.5. Adoptando una interpretación deductiva para la relación de soporte de* $\textsf{BAF}_{3.5}$*, obtenemos el* $\textsf{d-BAF}_{3.6}$ *que se ilustra en la Figura 3.6. En este caso, ocurre una derrota mediada de* $\mathcal{I}$ *sobre* $\mathcal{K}$ *dado que* $\mathcal{I}$ *derrota a* $\mathcal{J}$*, el cual es soportado por* $\mathcal{K}$*. Además, existe una derrota mediada de* $\mathcal{I}$ *sobre* $\mathcal{L}$*. Las derrotas mediadas de* $\textsf{d-BAF}_{3.6}$ *son denotadas en la Figura 3.6 utilizando flechas punteadas. Como se mostró en el Ejemplo 3.5, la cp-extensión de* $\textsf{BAF}_{3.5}$ *es* $\{\mathcal{I}, \mathcal{M}\}$*, la cual no es admisible en el sentido de Dung. Sin embargo, al considerar el* $\textsf{d-BAF}_{3.6}$*, el conjunto* $\{\mathcal{I}, \mathcal{M}\}$ *es admisible. Esto se debe a que, dada la derrota mediada de* $\mathcal{I}$ *sobre* $\mathcal{L}$*, el argumento* $\mathcal{M}$ *es ahora defendido por* $\mathcal{I}$ *ante la derrota del argumento* $\mathcal{L}$*.*

El otro inconveniente señalado por los autores de [BGvdTV10] es que, utilizando la aproximación presentada en la Sección 3.1, las derrotas secundarias pueden dar lugar a resultados no deseados cuando se adopta una interpretación deductiva de soporte. Para ilustrar esto, consideremos el siguiente ejemplo inspirado en [BGvdTV10].

Ejemplo 3.7 *Consideremos un escenario en el que dos equipos de fútbol, Liverpool* $(\mathcal{L})$ *y Manchester United* $(\mathcal{MU})$*, se encuentran en la puja final para ganar la Premier League* $(\mathcal{PL})$*. Supongamos que Liverpool será el campeón de la Premiere League* $(\mathcal{LCPL})$ *si gana su último partido* $(\mathcal{LG})$ *o Manchester United no gana su último partido* $(\mathcal{MUNG})$*. Esto se debe a que ambos equipos no juegan entre sí su último partido, por lo que el resultado de sus partidos es independiente. De esta manera, tenemos que "Liverpool gana su último*

$$\begin{tikzcd} \mathcal{LG} \arrow[r, Rightarrow, "d"] & \mathcal{LCPL} & \mathcal{MUNG} \arrow[l, Rightarrow, "d"'] \\ \mathcal{LP} \arrow[u] \arrow[ur, dashed, bend right] & & \end{tikzcd}$$

Figura 3.7: Representación gráfica correspondiente al escenario del Ejemplo 3.7.

partido" soporta a "Liverpool es el campeón de la Premier League" ($\mathcal{LG} \overset{d}{\Longrightarrow} \mathcal{LCPL}$)*, y "Manchester United no gana su último partido" también soporta a "Liverpool es el campeón de la Premier League"* ($\mathcal{MUNG} \overset{d}{\Longrightarrow} \mathcal{LCPL}$)*. Supongamos ahora que Liverpool pierde su último partido* ($\mathcal{LP}$) *y que Manchester United no gana el suyo* ($\mathcal{MUNG}$)*. En consecuencia, "Liverpool pierde su último partido" derrota a "Liverpool gana su último partido"* ($\mathcal{LP} \longrightarrow \mathcal{LG}$)*. Por lo tanto, al considerar las derrotas secundarias, tenemos que "Liverpool pierde su último partido" derrota a "Liverpool es el campeón de la Premier League"* ($\mathcal{LP} \longrightarrow \mathcal{LCPL}$)*. La Figura 3.7 provee una representación gráfica de esta situación utilizando soporte deductivo, donde la derrota secundaria es denotada con una flecha punteada. Luego, a partir de esta representación ocurriría que* $\mathcal{LCPL}$ *(que expresa "Liverpool es el campeón de la Premier League") no es un argumento aceptado del sistema ya que no es defendido ante la derrota de* $\mathcal{LP}$*.*

El resultado obtenido en el Ejemplo 3.7 resulta contra-intuitivo porque el argumento $\mathcal{LCPL}$ también está soportado por el argumento $\mathcal{MUNG}$, para el cual no hay derrotadores. Es decir, si ocurre que Manchester United no gana su último partido, entonces debería ocurrir que Liverpool es el campeón de la Premier League. Para evitar resultados contra-intuitivos, en [BGvdTV10] los autores proponen no utilizar las derrotas secundarias en su aproximación de meta-argumentación, e incorporar las derrotas mediadas en su lugar. No obstante esto, si bien las derrotas secundarias pueden dar lugar a resultados no deseados cuando se considera una interpretación deductiva de soporte, como se verá más adelante en la Sección 3.3, estas resultan útiles en contextos donde se adoptan otras interpretaciones de soporte.

Como se mencionó al comienzo de esta sección, el soporte deductivo impone las siguientes restricciones sobre los argumentos que relaciona: si $\mathcal{A}$ soporta a $\mathcal{B}$ y $\mathcal{A}$ está aceptado, entonces $\mathcal{B}$ también está aceptado; y si $\mathcal{A}$ soporta a $\mathcal{B}$, y $\mathcal{B}$ no está aceptado, entonces $\mathcal{A}$ tampoco está aceptado. A diferencia de la aproximación de coaliciones presentada en la Sección 3.1.1, en [BGvdTV10] no se agrupan argumentos, sino que se incorporan *meta-argumentos*. La idea es representar el soporte deductivo de un argumento $\mathcal{A}$ para un argumento $\mathcal{B}$ a través de la derrota del argumento $\mathcal{B}$ sobre un argumento auxiliar llamado $\mathcal{Z}_{\mathcal{A},\mathcal{B}}$, junto con la derrota del argumento $\mathcal{Z}_{\mathcal{A},\mathcal{B}}$ sobre el argumento $\mathcal{A}$.

Dado un conjunto $\mathbb{U}$, llamado universo de argumentos, para cada argumento $\mathcal{A} \in \mathbb{U}$ se introduce el meta-argumento $acc(\mathcal{A})$ que expresa "*el argumento $\mathcal{A}$ está aceptado*". Adicionalmente, se incorporan meta-argumentos $X_{\mathcal{A},\mathcal{B}}$ y $Y_{\mathcal{A},\mathcal{B}}$ por cada derrota $\mathcal{A} \longrightarrow \mathcal{B}$. El meta-argumento $X_{\mathcal{A},\mathcal{B}}$ expresa "*la derrota de $\mathcal{A}$ sobre $\mathcal{B}$ no está activa*", mientras que el meta-argumento $Y_{\mathcal{A},\mathcal{B}}$ expresa "*la derrota de $\mathcal{A}$ sobre $\mathcal{B}$ está activa*". Por cada par de argumentos $\mathcal{C}, \mathcal{D} \in \mathbb{U}$ tal que $\mathcal{C} \overset{d}{\Longrightarrow} \mathcal{D}$, se introduce un meta-argumento $Z_{\mathcal{C},\mathcal{D}}$ que expresa "*el argumento $\mathcal{C}$ no soporta al argumento $\mathcal{D}$*". Luego, el universo de meta-argumentos es caracterizado como $\mathbb{UM} = \{acc(\mathcal{A}) \mid \mathcal{A} \in \mathbb{U}\} \cup \{X_{\mathcal{A},\mathcal{B}}, Y_{\mathcal{A},\mathcal{B}}, Z_{\mathcal{A},\mathcal{B}} \mid \mathcal{A}, \mathcal{B} \in \mathbb{U}\}$. Finalmente, se define una relación de meta-derrota $\longmapsto$ que expresa los conflictos entre los argumentos originales en un meta-nivel.

La siguiente definición caracteriza el *marco meta-argumentativo* (EAF) que permite representar soporte deductivo entre argumentos. Dado un d-BAF, un argumento $\mathcal{A}$ pertenecerá a una extensión de d-BAF si y sólo si su meta-argumento asociado $acc(\mathcal{A})$ pertenece a la extensión correspondiente del marco meta-argumentativo EAF.

Definición 3.6 (Marco Meta-argumentativo asociado a d-BAF) *Sea d-BAF $= \langle \mathbb{A}, \longrightarrow, \overset{d}{\Longrightarrow} \rangle$ un marco argumentativo bipolar con soporte deductivo. El* marco meta-argumentativo asociado a d-BAF *es EAF $= \langle \mathbb{MA}, \longmapsto \rangle$, donde $\mathbb{MA} = \{acc(\mathcal{A}) \mid \mathcal{A} \in \mathbb{A}\} \cup \{X_{\mathcal{A},\mathcal{B}}, Y_{\mathcal{A},\mathcal{B}}, Z_{\mathcal{A},\mathcal{B}} \mid \mathcal{A}, \mathcal{B} \in \mathbb{A}\}$ y $\longmapsto \subseteq \mathbb{MA} \times \mathbb{MA}$ es una relación de meta-derrota tal que:*

- *Si $\mathcal{A} \longrightarrow \mathcal{B}$, entonces $acc(\mathcal{A}) \longmapsto X_{\mathcal{A},\mathcal{B}}$, $X_{\mathcal{A},\mathcal{B}} \longmapsto Y_{\mathcal{A},\mathcal{B}}$, y $Y_{\mathcal{A},\mathcal{B}} \longmapsto acc(\mathcal{B})$.*
- *Si $\mathcal{A} \overset{d}{\Longrightarrow} \mathcal{B}$, entonces $acc(\mathcal{B}) \longmapsto Z_{\mathcal{A},\mathcal{B}}$, y $Z_{\mathcal{A},\mathcal{B}} \longmapsto acc(\mathcal{A})$.*

Ejemplo 3.8 *Consideremos el d-BAF$_{3.8}$ caracterizado por el grafo de interacción ilustrado en la Figura 3.8. El marco argumentativo de coaliciones asociado a d-BAF$_{3.8}$ es CAF$_{3.8}$, ilustrado en la Figura 3.9(a). Luego, el marco meta-argumentativo asociado a d-BAF$_{3.8}$ es EAF$_{3.8}$, ilustrado en la Figura 3.9(b). La única extensión preferida de CAF$_{3.8}$ es $\{C_1\}$, por lo que la cp-extensión de d-BAF$_{3.8}$ es $\{\mathcal{P}, \mathcal{Q}, \mathcal{S}\}$. En contraste, la única extensión*

$$\begin{tikzcd}
\mathcal{P} \arrow[r, Rightarrow, "d"] \arrow[d, Rightarrow, "d"'] & \mathcal{Q} \arrow[r, Rightarrow, "d"] & \mathcal{R} \\
\mathcal{S} \arrow[urr] & &
\end{tikzcd}$$

Figura 3.8: Marco argumentativo bipolar con soporte deductivo $\mathsf{d\text{-}BAF}_{3.8}$, correspondiente al Ejemplo 3.8.

$$\begin{tikzcd}
\underset{\{\mathcal{P},\,\mathcal{Q},\,\mathcal{S}\}}{C_1} \arrow[r, mapsto] & \underset{\{\mathcal{P},\,\mathcal{Q},\,\mathcal{R}\}}{C_2}
\end{tikzcd}$$

(a)

$$\begin{tikzcd}
& Z_{\mathcal{P},\mathcal{Q}} \arrow[dl, mapsto] & acc(\mathcal{Q}) \arrow[l, mapsto] & Z_{\mathcal{Q},\mathcal{R}} \arrow[l, mapsto] & \\
acc(\mathcal{P}) & & & & acc(\mathcal{R}) \arrow[ul, mapsto] \\
Z_{\mathcal{P},\mathcal{S}} \arrow[u, mapsto] & acc(\mathcal{S}) \arrow[l, mapsto] \arrow[r, mapsto] & X_{\mathcal{S},\mathcal{R}} \arrow[rr, mapsto] & & Y_{\mathcal{S},\mathcal{R}} \arrow[u, mapsto]
\end{tikzcd}$$

(b)

Figura 3.9: (a) Marco argumentativo de coaliciones $\mathsf{CAF}_{3.8}$ y (b) marco meta-argumentativo $\mathsf{EAF}_{3.8}$ asociados al $\mathsf{d\text{-}BAF}_{3.8}$ del Ejemplo 3.8.

preferida de $\mathsf{EAF}_{3.8}$ *es* $\{acc(\mathcal{S}), Y_{\mathcal{S},\mathcal{R}}, Z_{\mathcal{Q},\mathcal{R}}, Z_{\mathcal{P},\mathcal{Q}}\}$*, por lo que la única extensión preferida de* $\mathsf{d\text{-}BAF}_{3.8}$ *sería* $\{\mathcal{S}\}$*.*

Nótese que las aproximaciones de [CLS07, CLS10] y [BGvdTV10] conducen a diferentes resultados en el Ejemplo 3.8. Dado que [BGvdTV10] considera soporte deductivo entre argumentos, si $\mathcal{A} \stackrel{d}{\Longrightarrow} \mathcal{B}$ y $\mathcal{B}$ no está aceptado, entonces $\mathcal{A}$ tampoco puede estar aceptado. Por lo tanto, en el Ejemplo 3.8, como $\mathcal{S} \longrightarrow \mathcal{R}$ y $\mathcal{Q} \stackrel{d}{\Longrightarrow} \mathcal{R}$, $\mathcal{Q}$ no puede estar aceptado ya que $\mathcal{R}$ no está aceptado. De manera similar, $\mathcal{P}$ tampoco está aceptado ya que soporta deductivamente a $\mathcal{Q}$.

Adicionalmente, es importante destacar que la relación de meta-derrota $\longmapsto$ (ver Definición 3.6) no tiene en cuenta las derrotas mediadas y las derrotas soportadas. Como se menciona en [BGvdTV10], dado $\mathcal{A} \stackrel{d}{\Longrightarrow} \mathcal{B}$, se condensan todas las derrotas que involucran a los argumentos $\mathcal{A}$ y $\mathcal{B}$ utilizando únicamente el meta-argumento $Z_{\mathcal{A},\mathcal{B}}$. De esta manera, el meta-argumento $Z_{\mathcal{A},\mathcal{B}}$ permite capturar el comportamiento modelado por las derrotas mediadas y las derrotas soportadas. En consecuencia, se simplifica la representación los

$$
\begin{tikzcd}
\mathcal{T} \arrow[r] \arrow[dr, dashed, bend right] & \mathcal{O} \arrow[r] & \mathcal{V} \\
& \mathcal{N} \arrow[u, Rightarrow, "d"] \arrow[ur, dashed, bend right] &
\end{tikzcd}
$$

Figura 3.10: Marco argumentativo bipolar con soporte deductivo $\mathsf{d\text{-}BAF}_{3.9}$, correspondiente al Ejemplo 3.9.

$$
\begin{tikzcd}
acc(\mathcal{T}) \arrow[r, mapsto] \arrow[dr, mapsto] & X_{\mathcal{T},\mathcal{O}} \arrow[r, mapsto] & Y_{\mathcal{T},\mathcal{O}} \arrow[r, mapsto] & acc(\mathcal{O}) \arrow[r, mapsto] \arrow[d, mapsto] & X_{\mathcal{O},\mathcal{V}} \arrow[r, mapsto] & Y_{\mathcal{O},\mathcal{V}} \arrow[r, mapsto] & acc(\mathcal{V}) \\
& X_{\mathcal{T},\mathcal{N}} \arrow[dr, mapsto] & & Z_{\mathcal{N},\mathcal{O}} \arrow[d, mapsto] & & Y_{\mathcal{N},\mathcal{V}} \arrow[ur, mapsto] & \\
& & Y_{\mathcal{T},\mathcal{N}} \arrow[r, mapsto] & acc(\mathcal{N}) \arrow[r, mapsto] & X_{\mathcal{N},\mathcal{V}} \arrow[ur, mapsto] & &
\end{tikzcd}
$$

Figura 3.11: Grafo de derrotas asociado al marco meta-argumentativo $\mathsf{EAF}_{3.9}$ del Ejemplo 3.9.

marcos meta-argumentativos en que ocurren derrotas mediadas y derrotas soportadas. Para ilustrar esto, consideremos el siguiente ejemplo.

Ejemplo 3.9 *Consideremos el marco argumentativo bipolar con soporte deductivo $\mathsf{d\text{-}BAF}_{3.9}$ cuyo grafo de interacción se ilustra en la Figura 3.10, donde las derrotas mediadas y las derrotas soportadas se denotan utilizando flechas punteadas.*

Supongamos que las derrotas mediadas y las derrotas soportadas de $\mathsf{d\text{-}BAF}_{3.9}$ son consideradas en la construcción del marco meta-argumentativo asociado $\mathsf{EAF}_{3.9}$. En tal caso, el grafo de derrotas correspondiente a $\mathsf{EAF}_{3.9}$ sería el ilustrado en la Figura 3.11. La derrota mediada de $\mathcal{T}$ sobre $\mathcal{N}$ expresa que si $\mathcal{T}$ está aceptado, entonces $\mathcal{N}$ no debe estar aceptado; esto es capturado por dos caminos en el grafo de derrotas de $\mathsf{EAF}_{3.9}$. En particular, como uno de estos caminos incluye al meta-argumento $Z_{\mathcal{N},\mathcal{O}}$, es posible desestimar el otro sin perder el efecto de la derrota mediada. Por otra parte, la derrota soportada de $\mathcal{N}$ sobre $\mathcal{V}$ también puede capturarse mediante diferentes caminos en el grafo de derrotas de $\mathsf{EAF}_{3.9}$. En la Figura 3.11 se observa el camino de $acc(\mathcal{N})$ a $acc(\mathcal{V})$, el cual corresponde a la derrota soportada en sí. Sin embargo, también es posible capturar el comportamiento de la derrota soportada mediante la combinación de otros dos caminos en el grafo de derrotas.

$$
\begin{tikzcd}
acc(\mathcal{T}) \arrow[r, mapsto] & X_{\mathcal{T},\mathcal{O}} \arrow[r, mapsto] & Y_{\mathcal{T},\mathcal{O}} \arrow[r, mapsto] & acc(\mathcal{O}) \arrow[r, mapsto] \arrow[d, mapsto] & X_{\mathcal{O},\mathcal{V}} \arrow[r, mapsto] & Y_{\mathcal{O},\mathcal{V}} \arrow[r, mapsto] & acc(\mathcal{V}) \\
 & & & Z_{\mathcal{N},\mathcal{O}} \arrow[d, mapsto] & & & \\
 & & & acc(\mathcal{N}) & & &
\end{tikzcd}
$$

Figura 3.12: Versión simplificada del grafo de derrotas asociado al marco meta-argumentativo $\mathsf{EAF}_{3.9}$ del Ejemplo 3.9.

La derrota soportada de $\mathcal{N}$ sobre $\mathcal{V}$ expresa que si $\mathcal{N}$ está aceptado, entonces $\mathcal{V}$ no debe estar aceptado. Entonces, si el meta-argumento $acc(\mathcal{N})$ está aceptado, esto implica que el meta-argumento $Z_{\mathcal{N},\mathcal{O}}$ no está aceptado. Más aún, como el único derrotador de $Z_{\mathcal{N},\mathcal{O}}$ es $acc(\mathcal{O})$, esto también implica que $acc(\mathcal{O})$ está aceptado y, como consecuencia, $acc(\mathcal{V})$ no está aceptado. Puede observarse que, como en el caso de la derrota mediada, esta alternativa considera un camino que involucra el meta-argumento $Z_{\mathcal{N},\mathcal{O}}$, posibilitando entonces la desestimación del camino de $acc(\mathcal{N})$ a $acc(\mathcal{V})$. Finalmente, la única extensión preferida de $\mathsf{EAF}_{3.9}$ es $\{acc(\mathcal{T}), Y_{\mathcal{T},\mathcal{O}}, Y_{\mathcal{T},\mathcal{N}}, Z_{\mathcal{N},\mathcal{O}}, X_{\mathcal{O},\mathcal{V}}, X_{\mathcal{N},\mathcal{V}}, acc(\mathcal{V})\}$, con lo cual la única extensión preferida de $\mathsf{d\text{-}BAF}_{3.9}$ es $\{\mathcal{T}, \mathcal{V}\}$.

Supongamos ahora que las derrotas mediadas y las derrotas soportadas no son consideradas en la relación de meta-derrota de $\mathsf{EAF}_{3.9}$. En este caso, obtenemos una versión simplificada del grafo de derrotas, ilustrada en la Figura 3.12. Dada esta representación simplificada, puede observarse que la única extensión preferida de $\mathsf{EAF}_{3.9}$ es $\{acc(\mathcal{T}), Y_{\mathcal{T},\mathcal{O}}, Z_{\mathcal{N},\mathcal{O}}, X_{\mathcal{O},\mathcal{V}}, acc(\mathcal{V})\}$, a diferencia de lo ocurrido para la versión extendida del grafo de derrotas. No obstante esto, la única extensión preferida de $\mathsf{d\text{-}BAF}_{3.9}$ sigue siendo $\{\mathcal{T}, \mathcal{V}\}$.

La representación simplificada mostrada en el Ejemplo 3.9 es posible ya que en [BGvdTV10] los autores prueban que, dado un d-BAF, la incorporación de las derrotas soportadas y las derrotas mediadas al EAF asociado no cambia las extensiones del d-BAF empleando las semánticas de [Dun95]. En particular, esto último puede observarse en el Ejemplo 3.9, para el cual ambas aproximaciones dan como resultado que la única extensión preferida de $\mathsf{d\text{-}BAF}_{3.9}$ es $\{\mathcal{T}, \mathcal{V}\}$.

3.2.1. Soporte Rebatible: anulando las restricciones del soporte deductivo

En [BGvdTV10] los autores presentan una aproximación extendida que considera dos tipos de derrotas de segundo orden. Por una parte, consideran derrotas originadas en un argumento o una derrota y cuyo destino es otra derrota. Por otra parte, consideran derrotas cuyo origen es un argumento y el destino es un soporte. En la literatura existen otras aproximaciones que abordan el primer tipo de derrota de segundo orden, tales como [Mod09] y [BCGG11]). Sin embargo, en [BGvdTV10] los autores remarcan que la diferencia entre su propuesta y las otras existentes en la literatura es que ellos además consideran el caso en que una derrota de segundo orden tiene como origen y destino a una derrota. Además, destacan que el segundo tipo de derrota de segundo orden que proponen permite el modelado de *soporte rebatible* (en inglés, *defeasible support*), el cual anula las restricciones impuestas por el soporte deductivo.

Con respecto al segundo tipo de derrota de segundo orden propuesta en [BGvdTV10], el trabajo de Pollock [Pol87] proveyó las bases para futuras investigaciones en este tópico. En su trabajo, Pollock enunció que las razones rebatibles (las cuales pueden ensamblarse para construir argumentos) poseen derrotadores, e identificó dos tipos de derrotadores: *derrotadores por refutación* (en inglés, *rebutting defeaters*) y *derrotadores por socavamiento* (en inglés,*undercutting defeaters*). Los derrotadores del primer tipo derrotan la conclusión de una inferencia proveyendo un argumento para la conclusión contraria, mientras que los del segundo tipo derrotan la conexión entre las premisas y la conclusión (*i. e.*, no derrotan a la conclusión directamente). El trabajo de Boella *et al.* [BGvdTV10] fue el primero en considerar las ideas de Pollock en el contexto de marcos argumentativos bipolares. Por otra parte, en el Capítulo 4 se desarrollará un formalismo que contempla las ideas de Pollock, proponiendo un marco argumentativo abstracto con una relación de soporte explícita [CGS11a, CGS12b].

La siguiente definición provee un mecanismo para instanciar un marco argumentativo de Dung como un marco argumentativo bipolar de segundo orden utilizando meta-argumentación, de acuerdo a lo propuesto en [BGvdTV10]. Como se mencionó anteriormente, los conjuntos $\mathbb{U}$ y $\mathbb{UM}$ representan, respectivamente, el universo de argumentos y el universo de meta-argumentos. Adicionalmente, el acrónimo d-EBAF es utilizado para denotar un marco argumentativo bipolar con soporte deductivo y derrotas de segundo orden, al que llamaremos *marco argumentativo bipolar extendido con soporte deductivo*.

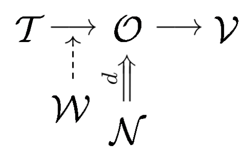


Figura 3.13: Marco argumentativo bipolar extendido con soporte deductivo $\mathsf{d\text{-}EBAF}_{3.10}$, correspondiente al Ejemplo 3.10.

Definición 3.7 (Marco Meta-argumentativo asociado a $\mathsf{d\text{-}EBAF}$) *Sea $\mathit{d\text{-}EBAF} = \langle \mathbb{A}, \longrightarrow, \stackrel{d}{\Longrightarrow}, \longrightarrow^2 \rangle$ un marco argumentativo bipolar extendido con soporte deductivo, donde $\mathbb{A} \subseteq \mathbb{U}$ es un conjunto de argumentos, $\longrightarrow \subseteq \mathbb{A} \times \mathbb{A}$ es una relación de derrota, $\stackrel{d}{\Longrightarrow} \subseteq \mathbb{A} \times \mathbb{A}$ es una relación de soporte deductivo, y $\longrightarrow^2 \subseteq (\mathbb{A} \cup \longrightarrow) \times (\longrightarrow \cup \stackrel{d}{\Longrightarrow})$ es una relación de derrota de segundo orden. El* marco meta-argumentativo asociado a $\mathsf{d\text{-}EBAF}$ *es $\mathit{EAF}^+ = \langle \mathbb{MA}, \longmapsto \rangle$, donde $\mathbb{MA} = \{acc(\mathcal{A}) \mid \mathcal{A} \in \mathbb{A}\} \cup \{X_{\mathcal{A},\mathcal{B}}, Y_{\mathcal{A},\mathcal{B}}, Z_{\mathcal{A},\mathcal{B}} \mid \mathcal{A}, \mathcal{B} \in \mathbb{A}\} \cup \{X_{\mathcal{A},\mathcal{B}\to\mathcal{C}}, Y_{\mathcal{A},\,\mathcal{B}\to\mathcal{C}}, X_{\mathcal{C},\,Z_{\mathcal{A},\mathcal{B}}}, Y_{\mathcal{C},\,Z_{\mathcal{A},\mathcal{B}}} \mid \mathcal{A}, \mathcal{B}, \mathcal{C} \in \mathbb{A}\}$ y $\longmapsto \subseteq \mathbb{MA} \times \mathbb{MA}$ es una relación de meta-derrota tal que:*

- *Si $\mathcal{A} \longrightarrow \mathcal{B}$, entonces $acc(\mathcal{A}) \longmapsto X_{\mathcal{A},\mathcal{B}}$, $X_{\mathcal{A},\mathcal{B}} \longmapsto Y_{\mathcal{A},\mathcal{B}}$ y $Y_{\mathcal{A},\mathcal{B}} \longmapsto acc(\mathcal{B})$.*
- *Si $\mathcal{A} \stackrel{d}{\Longrightarrow} \mathcal{B}$, entonces $acc(\mathcal{B}) \longmapsto Z_{\mathcal{A},\mathcal{B}}$ y $Z_{\mathcal{A},\mathcal{B}} \longmapsto acc(\mathcal{A})$.*
- *Si $\mathcal{A} \longrightarrow^2 (\mathcal{B} \longrightarrow \mathcal{C})$, entonces $acc(\mathcal{A}) \longmapsto X_{\mathcal{A},\mathcal{B}\to\mathcal{C}}$, $X_{\mathcal{A},\mathcal{B}\to\mathcal{C}} \longmapsto Y_{\mathcal{A},\mathcal{B}\to\mathcal{C}}$ y $Y_{\mathcal{A},\mathcal{B}\to\mathcal{C}} \longmapsto Y_{\mathcal{B},\mathcal{C}}$.*
- *Si $(\mathcal{A} \longrightarrow \mathcal{B}) \longrightarrow^2 (\mathcal{C} \longrightarrow \mathcal{D})$, entonces $Y_{\mathcal{A},\mathcal{B}} \longmapsto Y_{\mathcal{C},\mathcal{D}}$.*
- *Si $\mathcal{C} \longrightarrow^2 (\mathcal{A} \stackrel{d}{\Longrightarrow} \mathcal{B})$, entonces $acc(\mathcal{C}) \longmapsto X_{\mathcal{C},\,Z_{\mathcal{A},\mathcal{B}}}$, $X_{\mathcal{C},\,Z_{\mathcal{A},\mathcal{B}}} \longmapsto Y_{\mathcal{C},\,Z_{\mathcal{A},\mathcal{B}}}$ y $Y_{\mathcal{C},\,Z_{\mathcal{A},\mathcal{B}}} \longmapsto Z_{\mathcal{A},\mathcal{B}}$.*

Ejemplo 3.10 *Consideremos el $\mathit{d\text{-}BAF}_{3.9}$ del Ejemplo 3.9, y supongamos que le añadimos la derrota de segundo orden $\mathcal{W} \longrightarrow^2 (\mathcal{T} \longrightarrow \mathcal{O})$. El marco argumentativo bipolar extendido con soporte deductivo resultante es $\mathit{d\text{-}EBAF}_{3.10}$, ilustrado en la Figura 3.13, donde la derrota de segundo orden es representada utilizando una flecha punteada. El marco meta-argumentativo asociado a $\mathit{d\text{-}EBAF}_{3.10}$ es $\mathit{EAF}^+{}_{3.10}$, el cual se ilustra en la Figura 3.14. La única extensión preferida de $\mathit{EAF}^+{}_{3.10}$ es $\{acc(\mathcal{T}), acc(\mathcal{W}), Y_{\mathcal{W},\,\mathcal{T}\to\mathcal{O}}, acc(\mathcal{O}), acc(\mathcal{N}), Y_{\mathcal{O},\mathcal{V}}\}$. Por lo tanto, la única extensión preferida de $\mathit{d\text{-}EBAF}_{3.10}$ es $\{\mathcal{T}, \mathcal{W}, \mathcal{O}, \mathcal{N}\}$, ya que*

$$\begin{tikzcd}
acc(\mathcal{T}) \arrow[r, mapsto] & X_{\mathcal{T},\mathcal{O}} \arrow[r, mapsto] & Y_{\mathcal{T},\mathcal{O}} \arrow[r, mapsto] & acc(\mathcal{O}) \arrow[r, mapsto] \arrow[d, mapsto] & X_{\mathcal{O},\mathcal{V}} \arrow[r, mapsto] & Y_{\mathcal{O},\mathcal{V}} \arrow[r, mapsto] & acc(\mathcal{V}) \\
acc(\mathcal{W}) \arrow[r, mapsto] & X_{\mathcal{W},\,\mathcal{T}\rightarrow\mathcal{O}} \arrow[r, mapsto] & Y_{\mathcal{W},\,\mathcal{T}\rightarrow\mathcal{O}} \arrow[u, mapsto] & Z_{\mathcal{N},\mathcal{O}} \arrow[d, mapsto] & & & \\
& & & acc(\mathcal{N}) & & &
\end{tikzcd}$$

Figura 3.14: Marco meta-argumentativo $\mathsf{EAF}^{+}{}_{3.10}$ asociado al $\mathsf{d\text{-}EBAF}_{3.10}$ del Ejemplo 3.10.

la derrota de $\mathcal{T}$ sobre $\mathcal{O}$ es dejada sin efecto por $\mathcal{W}$. De esta manera, como $\mathcal{O}$ ya no está derrotado y puede ser aceptado, se efectiviza la derrota de $\mathcal{O}$ sobre $\mathcal{V}$. Finalmente, el argumento $\mathcal{N}$, quien soporta deductivamente a $\mathcal{O}$, también está aceptado.

Siguiendo la Definición 3.7, una derrota de segundo orden cuyo destino es la relación de soporte tiene por efecto anular las restricciones impuestas por el soporte deductivo. Cuando se tiene que $\mathcal{C} \longrightarrow^2 (\mathcal{A} \overset{d}{\Longrightarrow} \mathcal{B})$, la relación de meta-derrota es tal que $acc(\mathcal{C}) \longmapsto X_{\mathcal{C},\, Z_{\mathcal{A},\,\mathcal{B}}} \longmapsto Y_{\mathcal{C},\, Z_{\mathcal{A},\,\mathcal{B}}} \longmapsto Z_{\mathcal{A},\mathcal{B}}$. Así, si $acc(\mathcal{C})$ está aceptado, entonces $Z_{\mathcal{A},\mathcal{B}}$ queda derrotado. Como resultado, dado que el soporte deductivo de $\mathcal{A}$ para $\mathcal{B}$ es anulado por $\mathcal{C}$, si $\mathcal{B}$ no está aceptado puede ocurrir que a $\mathcal{A}$ esté aceptado y, a la inversa, si $\mathcal{A}$ está aceptado puede darse el caso que $\mathcal{B}$ no esté aceptado. Para ilustrar esta situación consideremos el siguiente ejemplo.

Ejemplo 3.11 *Supongamos que extendemos el marco argumentativo bipolar con soporte deductivo $\mathsf{d\text{-}BAF}_{3.9}$ del Ejemplo 3.9, incorporando la derrota de segundo orden $\mathcal{W} \longrightarrow^2 (\mathcal{N} \overset{d}{\Longrightarrow} \mathcal{O})$. En este caso, obtenemos el marco argumentativo bipolar extendido con soporte deductivo $\mathsf{d\text{-}EBAF}_{3.11}$ ilustrado en la Figura 3.15, donde la flecha punteada representa la derrota de segundo orden. Además, $\mathsf{EAF}^{+}{}_{3.11}$ es el marco meta-argumentativo asociado a $\mathsf{d\text{-}EBAF}_{3.11}$, ilustrado en la Figura 3.16. En este caso, el argumento $\mathcal{O}$ es derrotado por el argumento $\mathcal{T}$ y, por lo tanto, no está aceptado. Sin embargo, el argumento $\mathcal{N}$ está aceptado porque el soporte de $\mathcal{N}$ para $\mathcal{O}$ resulta anulado por la derrota del argumento $\mathcal{W}$. Tenemos entonces que la única extensión preferida de $\mathsf{EAF}^{+}{}_{3.11}$ es $\{acc(\mathcal{T}), Y_{\mathcal{T},\mathcal{O}}, acc(\mathcal{W}), Y_{\mathcal{W},\, Z_{\mathcal{N},\mathcal{O}}}, acc(\mathcal{N}), X_{\mathcal{O},\mathcal{V}}, acc(\mathcal{V})\}$, y la correspondiente extensión preferida de $\mathsf{d\text{-}EBAF}_{3.11}$ es $\{\mathcal{T}, \mathcal{W}, \mathcal{N}, \mathcal{V}\}$.*

Finalmente, en [BGvdTV10] los autores destacan que su propuesta para modelar soporte rebatible admite la representación de los dos tipos de derrotadores identificados por

$$\begin{tikzcd}
\mathcal{T} \arrow[r] & \mathcal{O} \arrow[r] & \mathcal{V} \\
& & \mathcal{W} \\
& \mathcal{N} \arrow[uu, Rightarrow, "d", ""{name=D, right}] &
\arrow[from=2-3, to=D, dashed]
\end{tikzcd}$$

Figura 3.15: Marco argumentativo bipolar extendido con soporte deductivo $\mathsf{d\text{-}EBAF}_{3.11}$, correspondiente al Ejemplo 3.11.

$$\begin{tikzcd}
acc(\mathcal{T}) \arrow[r, mapsto] & X_{\mathcal{T},\mathcal{O}} \arrow[r, mapsto] & Y_{\mathcal{T},\mathcal{O}} \arrow[r, mapsto] & acc(\mathcal{O}) \arrow[r, mapsto] \arrow[d, mapsto] & X_{\mathcal{O},\mathcal{V}} \arrow[r, mapsto] & Y_{\mathcal{O},\mathcal{V}} \arrow[r, mapsto] & acc(\mathcal{V}) \\
acc(\mathcal{W}) \arrow[r, mapsto] & X_{\mathcal{W},\, Z_{\mathcal{N},\mathcal{O}}} \arrow[r, mapsto] & Y_{\mathcal{W},\, Z_{\mathcal{N},\mathcal{O}}} \arrow[r, mapsto] & Z_{\mathcal{N},\mathcal{O}} \arrow[d, mapsto] & & & \\
& & & acc(\mathcal{N}) & & &
\end{tikzcd}$$

Figura 3.16: Marco meta-argumentativo $\mathsf{EAF}^{+}{}_{3.11}$ asociado al $\mathsf{d\text{-}EBAF}_{3.11}$ del Ejemplo 3.11.

Pollock [Pol87]. Por una parte, la derrota de tipo *rebutting* ocurre cuando $\mathcal{A} \stackrel{d}{\Longrightarrow} \mathcal{B}$ y $\mathcal{C} \longrightarrow \mathcal{B}$, y es capturada a través de la derrota mediada de $\mathcal{C}$ sobre $\mathcal{A}$). Por otra parte, la derrota de tipo *undercutting* toma lugar cuando $\mathcal{C} \longrightarrow^2 (\mathcal{A} \stackrel{d}{\Longrightarrow} \mathcal{B})$.

3.3. Marcos Argumentativos con Soporte por Necesidad (AFN)

En [NR10] se introdujeron por primera vez los *Marcos Argumentativos con Soporte por Necesidad (AFN)* (por su nombre en inglés, *Argumentation Frameworks with Necessities*). Esta propuesta extiende a los marcos argumentativos de [Dun95] incorporando una relación especializada de soporte entre argumentos. La relación de *soporte por necesidad* establece que si un argumento $\mathcal{A}$ soporta a un argumento $\mathcal{B}$, entonces $\mathcal{A}$ es necesario para obtener $\mathcal{B}$. De esta manera, si $\mathcal{B}$ está aceptado entonces $\mathcal{A}$ también está aceptado, y si $\mathcal{A}$ no está aceptado entonces $\mathcal{B}$ tampoco está aceptado. Los autores de [NR10] argumentan que, a diferencia de una relación de soporte general, la relación de necesidad tiene la ventaja de asegurar que su interacción con la relación de derrota genera nuevas derrotas que comparten la naturaleza de las derrotas directas. Además, afirman que esta especialización de la relación de soporte permite la generalización de las semánticas de aceptabilidad de forma natural, asegurando el cumplimiento de la noción de admisibilidad.

Definición 3.8 (Marco Argumentativo con Soporte por Necesidad) *Un* Marco Argumentativo con Soporte por Necesidad (AFN) *es una tupla* $\langle \mathbb{A}, \mathbb{R}, \mathbb{N} \rangle$*, donde* $\mathbb{A}$ *es un conjunto de argumentos,* $\mathbb{R} \subseteq \mathbb{A} \times \mathbb{A}$ *es una relación de derrota*[4]*, y* $\mathbb{N} \subseteq \mathbb{A} \times \mathbb{A}$ *es una relación de necesidad irreflexiva y transitiva.*

La definición de AFN presentada aquí corresponde a aquella introducida en [BNR12], donde la relación de soporte por necesidad $\mathbb{N}$ es irreflexiva y transitiva. Por otra parte, la relación de derrota $\mathbb{R}$ en el AFN es la misma que aquella definida para los marcos argumentativos de Dung. A partir de las derrotas originales y de la relación de soporte por necesidad es posible obtener dos nuevos tipos de derrota: si $\mathcal{A}$ derrota a $\mathcal{C}$ y $\mathcal{C}$ es necesario para $\mathcal{B}$, entonces $\mathcal{A}$ derrota a $\mathcal{B}$; y si $\mathcal{C}$ derrota a $\mathcal{B}$ y $\mathcal{C}$ es necesario para $\mathcal{A}$, entonces $\mathcal{A}$ derrota a $\mathcal{B}$. Al igual que en las secciones 3.1 y 3.2, nos referiremos a estas derrotas como *derrotas complejas*. Estas nuevas derrotas se encuentran formalizadas por la siguiente definición.

Definición 3.9 (Derrota Extendida) *Sea* $\langle \mathbb{A}, \mathbb{R}, \mathbb{N} \rangle$ *un AFN y sean* $\mathcal{A}, \mathcal{B} \in \mathbb{A}$*. Existe una* derrota extendida *de* $\mathcal{A}$ *sobre* $\mathcal{B}$*, denotada* $\mathcal{A}\mathbb{R}^{+}\mathcal{B}$*, si y sólo si* $\exists \mathcal{C} \in \mathbb{A}$ *tal que o bien* $\mathcal{A}\mathbb{R}\mathcal{C}$ *y* $\mathcal{C}\mathbb{N}\mathcal{B}$*, o* $\mathcal{C}\mathbb{R}\mathcal{B}$ *y* $\mathcal{C}\mathbb{N}\mathcal{A}$*. La derrota directa* $\mathcal{A}\mathbb{R}\mathcal{B}$ *es considerada como un caso particular de derrota extendida.*

Originalmente, en [NR10] se introdujo únicamente el primer tipo de derrota extendida, la cual coincide con la derrota secundaria propuesta en [CLS05] (ver Sección 3.1). Luego, en [NR11] se avanzó en la formalización de los AFN introduciendo, entre otras cosas, el segundo tipo de derrota extendida.

Es posible representar gráficamente un AFN mediante un multi-grafo dirigido, donde los nodos son argumentos y hay dos tipos de arcos correspondientes a las relaciones de derrota y soporte. En [NR10] los autores utilizan $\longrightarrow$ y $\dashrightarrow$ para, respectivamente, denotar derrota y soporte entre argumentos. Sin embargo, para mantener uniformidad en

[4] Al igual que lo ocurrido en los BAF, los autores se refieren a ella como relación de 'ataque'. Siguiendo la convención adoptada en esta tesis, para evitar confusiones nos referiremos a ella como relación de 'derrota'.

(a) (b)

Figura 3.17: (a) Marco argumentativo con soporte por necesidad $\mathsf{AFN}_{3.12}$, correspondiente al Ejemplo 3.12 y (b) las derrotas extendidas que origina.

la notación, en esta tesis utilizaremos $\mathcal{A} \longrightarrow \mathcal{B}$ para denotar $\mathcal{A}\mathbb{R}\mathcal{B}$ (derrota), y $\mathcal{A} \overset{n}{\Longrightarrow} \mathcal{B}$ para denotar $\mathcal{A}\mathbb{N}\mathcal{B}$ (soporte por necesidad). En este caso, la etiqueta 'n' sobre la flecha doble indica que la relación de soporte es interpretada como necesidad. Para simplificar la representación, el grafo solamente incluirá las necesidades "directas". Luego, aquellas obtenidas por transitividad sobre la relación de soporte podrán visualizarse al seguir los caminos de soporte en el grafo. A continuación se incluye un ejemplo que ilustra las derrotas extendidas del AFN y su representación gráfica.

Ejemplo 3.12 *Consideremos el marco argumentativo con soporte por necesidad $\mathsf{AFN}_{3.12}$ ilustrado en la Figura 3.17(a). Como puede observarse a partir de los caminos de soporte en el grafo que describe a $\mathsf{AFN}_{3.12}$, además de las necesidades directas, existe una necesidad de $\mathcal{A}$ sobre $\mathcal{C}$, y de $\mathcal{D}$ sobre $\mathcal{F}$. La Figura 3.17(b) ilustra las derrotas extendidas obtenidas a partir de $\mathsf{AFN}_{3.12}$, las cuales se hallan denotadas por flechas punteadas (a excepción de aquellas que son, además, derrotas directas). Dado que $\mathcal{D}$ derrota a $\mathcal{A}$ y $\mathcal{A}$ soporta a $\mathcal{B}$ y $\mathcal{C}$ (respectivamente, directa e indirectamente), existen derrotas extendidas del primer tipo de $\mathcal{D}$ sobre $\mathcal{B}$ y $\mathcal{C}$. Por otra parte, como $\mathcal{D}$ soporta a $\mathcal{E}$ y $\mathcal{F}$, también existen derrotas extendidas del segundo tipo de $\mathcal{E}$ y $\mathcal{F}$ sobre $\mathcal{A}$.*

El cómputo de aceptabilidad en los AFN es efectuado siguiendo la aproximación extensional de [Dun95]. En presencia de una relación de soporte por necesidad, un requerimiento esencial a satisfacer por los conjuntos de argumentos aceptables es evitar ciclos de necesidades, ya que reflejan una forma de bloqueo que puede ser vista como una instancia de la falacia conocida como "*begging the question*"[5]. En [NR10] los autores abordaron esta

[5]Este tipo de falacia ocurre cuando se intenta probar una proposición "p" utilizando a la misma proposición"p" como parte de la prueba.

cuestión asumiendo que la relación de necesidad de los AFN es acíclica. Luego, en [NR11] proveyeron una caracterización formal de la noción de *libertad de ciclos de necesidad* (en inglés, *necessity-cycle-freeness*). Finalmente, en [BNR12] brindaron una nueva caracterización de los AFN requiriendo que la relación de necesidad sea irreflexiva y transitiva, evitando entonces la posibilidad de generar ciclos de necesidades. Como se mencionó anteriormente, la Definición 3.8 toma en cuenta las restricciones para la relación de necesidad introducidas en [BNR12]. En consecuencia, se definen las nociones de *conjunto coherente* y *conjunto fuertemente coherente* (en inglés, *coherent set* y *strongly coherent set*) como sigue.

Definición 3.10 (Conjunto Coherente y Conjunto Fuertemente Coherente) *Sea* $\langle \mathbb{A}, \mathbb{R}, \mathbb{N} \rangle$ *un AFN y sea* $S \subseteq \mathbb{A}$*:*

- S *es* coherente *si y sólo si es cerrado bajo* $\mathbb{N}^{-1}$ *(i. e.,* $\forall \mathcal{A} \in S$*: si* $\exists \mathcal{B} \in \mathbb{A}$ *tal que* $\mathcal{B}\mathbb{N}\mathcal{A}$*, entonces* $\mathcal{B} \in S$*).*
- S *es* fuertemente coherente *si y sólo si es coherente y libre de conflictos con respecto a* $\mathbb{R}$ *(i. e.,* $\nexists \mathcal{A}, \mathcal{B} \in S$ *tal que* $(\mathcal{A}, \mathcal{B}) \in \mathbb{R}$*).*

El requerimiento de coherencia asegura que un conjunto de argumentos S contiene a todos los argumentos necesarios para cada uno de sus miembros. Por otra parte, dado que la coherencia concierne únicamente la relación de necesidad $\mathbb{N}$, si $\mathbb{N} = \emptyset$ entonces la noción de coherencia fuerte se reduce a la noción de libertad de conflictos propuesta en [Dun95]. Por ejemplo, dado el $\mathsf{AFN}_{3.12}$ del Ejmplo 3.12, algunos conjuntos coherentes son $\{\mathcal{A}\}$, $\{\mathcal{A}, \mathcal{B}, \mathcal{C}, \mathcal{D}\}$ y $\{\mathcal{D}, \mathcal{E}, \mathcal{F}\}$. En contraste, por ejemplo, el conjunto $\{\mathcal{B}, \mathcal{C}\}$ no es coherente ya que $\mathcal{A}\mathbb{N}\mathcal{B}$ y $\mathcal{A}$ no pertenece al conjunto. De los conjuntos coherentes arriba mencionados, únicamente $\{\mathcal{A}\}$ y $\{\mathcal{D}, \mathcal{E}, \mathcal{F}\}$ son fuertemente coherentes; el conjunto $\{\mathcal{A}, \mathcal{B}, \mathcal{C}, \mathcal{D}\}$ no es fuertemente coherente dado que $\mathcal{D}\mathbb{R}\mathcal{A}$ y, por lo tanto, no es libre de conflictos.

El cálculo de aceptabilidad para los AFN sigue los mismos principios que la aproximación Dung, reemplazando la noción de libertad de conflictos por la noción de coherencia fuerte. Luego, los conjuntos admisibles y las extensiones preferidas de un AFN son caracterizados de la siguiente manera.

Definición 3.11 (Conjunto Admisible y Extensión Preferida) *Sea* $\mathsf{AFN} = \langle \mathbb{A}, \mathbb{R}, \mathbb{N} \rangle$ *un marco argumentativo con soporte por necesidad y* $S \subseteq \mathcal{A}$*.*

- S *es un* conjunto admisible *si y sólo si es fuertemente coherente y si* $\exists \mathcal{B} \in \mathbb{A}, \exists \mathcal{A} \in S$ *tal que* $\mathcal{B}\mathbb{R}A$*, entonces para cada conjunto coherente* $S' \subseteq \mathbb{A} \backslash S$ *tal que* $\mathcal{B} \in S'$ *vale que* $\exists \mathcal{C} \in S, \exists \mathcal{D} \in S'$ *tal que* $\mathcal{C}\mathbb{R}\mathcal{D}$*.*
- S *es una* extensión preferida *de* AFN *si y sólo si es un conjunto admisible maximal con respecto a la inclusión de conjuntos.*

Dado el $\mathsf{AFN}_{3.12}$ del Ejemplo 3.12, algunos conjuntos admisibles son $\{\mathcal{D}\}$ y $\{\mathcal{D}, \mathcal{E}, \mathcal{F}\}$. El conjunto $\{\mathcal{D}\}$ es trivialmente admisible ya que es fuertemente coherente y el argumento $\mathcal{D}$ no es derrotado por ningún otro argumento de $\mathsf{AFN}_{3.12}$. Por otra parte, el conjunto $\{\mathcal{D}, \mathcal{E}, \mathcal{F}\}$ es fuertemente coherente y es defendido ante la derrota $\mathcal{C}$, ya que $\mathcal{D}$ derrota a $\mathcal{A}$ y $\mathcal{A}$ es un miembro de $\{\mathcal{A}, \mathcal{B}, \mathcal{C}\}$, el cual es el único conjunto coherente que contiene a $\mathcal{C}$. En contraste, el conjunto $\{\mathcal{A}, \mathcal{B}, \mathcal{C}\}$ no es admisible porque $\mathcal{D}$ derrota a $\mathcal{A}$ y el conjunto no provee defensa ante esta derrota. Por último, la única extensión preferida de $\mathsf{AFN}_{3.12}$ es $\{\mathcal{D}, \mathcal{E}, \mathcal{F}\}$.

En [NR11] los autores remarcan que la noción de admisibilidad que ellos proponen puede verse como una extensión de la noción de admisibilidad de Dung, donde la libertad de conflictos es reemplazada por la coherencia fuerte y la defensa tiene en cuenta las derrotas extendidas además de las derrotas directas. Luego, muestran que un conjunto de argumentos S es admisible si y sólo si es fuertemente coherente y para cada argumento $\mathcal{A}$ tal que $\mathcal{A}\mathbb{R}S$ (*i. e.*, $\exists \mathcal{B} \in S$ tal que $\mathcal{A}\mathbb{R}\mathcal{B}$), vale que $S\mathbb{R}^{+}\mathcal{B}$. Adicionalmente, introducen una serie de propiedades en relación a las extensiones de los AFN, tales como la preservación de propiedades de las extensiones preferidas existentes para los marcos argumentativos de Dung.

Asimismo, los autores remarcan que sus resultados también son válidos para la interpretación de soporte deductivo, para lo cual simplemente se debe reemplazar la relación de necesidad $\mathbb{N}$ por una relación de soporte deductivo $\mathbb{D}$, y utilizar clausura bajo $\mathbb{D}$ en lugar de $\mathbb{N}^{-1}$. En particular, destacan que las interpretaciones de soporte deductivo y soporte por necesidad son duales. Esta dualidad, también señalada en [CGGS13b, CLS13], establece que $\mathcal{A} \overset{n}{\Longrightarrow} \mathcal{B}$ es equivalente a $\mathcal{B} \overset{d}{\Longrightarrow} \mathcal{A}$. Por lo tanto, al invertir la dirección de la relación de soporte deductivo de [BGvdTV10], las derrotas mediadas y las derrotas soportadas se corresponden con las derrotas extendidas de [NR11]. Análogamente, al invertir la dirección de la relación de necesidad de [NR11], las derrotas extendidas se corresponden

$$\begin{tikzcd}
\mathcal{LG} \arrow[r, Rightarrow, "n"] & \mathcal{LCPL} & \mathcal{MUNG} \arrow[l, Rightarrow, "n"'] \\
\mathcal{LP} \arrow[u] \arrow[ur, dashed, bend right] & &
\end{tikzcd}$$

Figura 3.18: Marco Argumentativo con Soporte por Necesidad correspondiente al Ejemplo 3.13.

con las derrota mediada y la derrota soportada de [BGvdTV10]. A continuación se incluye una serie de ejemplos que ilustra esta dualidad con un mayor nivel de detalle.

Ejemplo 3.13 *Consideremos el escenario introducido en el Ejemplo 3.7, donde Liverpool ($\mathcal{L}$) o Manchester United ($\mathcal{MU}$) está por convertirse en el campeón de la Premier League ($\mathcal{PL}$). Liverpool será el campeón de la Premier League ($\mathcal{LCPL}$) si gana su último partido ($\mathcal{LG}$) o si Manchester United no gana el suyo ($\mathcal{MUNG}$). Recordemos que, como se mencionó en el Ejemplo 3.7, Liverpool y Manchester United no juegan entre sí su último partido. Además, se tiene que Liverpool pierde su último partido ($\mathcal{LP}$) y Manchester United no gana el suyo.*

Supongamos que queremos modelar esta situación adoptado una interpretación de soporte por necesidad, en cuyo caso la representación resultante es la ilustrada por la Figura 3.18. A partir de esta representación se tiene que "Liverpool pierde su último partido" derrota a "Liverpool gana su último partido" ($\mathcal{LP} \longrightarrow \mathcal{LG}$). Además, dada la interpretación de soporte por necesidad, existe una derrota extendida de $\mathcal{LP}$ sobre $\mathcal{LCPL}$, ilustrada en la Figura 3.18 mediante una flecha punteada. Como consecuencia, la inferencia del sistema indicaría que Liverpool no es el campeón de la Premier League. Claramente, este es un resultado indeseado porque Manchester United no ganó su último partido, transformando a Liverpool en campeón. Este resultado inesperado surge del uso equivocado de la interpretación de soporte. Recordemos que el soporte por necesidad establece que un argumento es aceptado si y sólo si todos *los argumentos que lo soportan están aceptados. Por lo tanto, al ocurrir que $\mathcal{LG}$ no está aceptado, se tiene que $\mathcal{LCPL}$ tampoco está aceptado.*

Supongamos ahora que en lugar de utilizar soporte por necesidad adoptamos una interpretación de soporte deductivo. La nueva representación para esta situación se halla ilustrada en la Figura 3.19. Dada la interpretación de soporte deductivo, la derrota de $\mathcal{LP}$ sobre $\mathcal{LCPL}$ (ocurrida en la representación de la Figura 3.18) ya no existe. Esto se debe a

$$\begin{tikzcd} \mathcal{LG} \arrow[r, Rightarrow, "d"] & \mathcal{LCPL} & \mathcal{MUNG} \arrow[l, Rightarrow, "d"'] \\ \mathcal{LP} \arrow[u] \end{tikzcd}$$

Figura 3.19: Marco Argumentativo Bipolar con Soporte Deductivo correspondiente al Ejemplo 3.13.

que las únicas derrotas adicionales que tienen sentido bajo una interpretación de soporte deductivo son la derrota soportada y la derrota mediada. Por lo tanto, a partir de la representación ilustrada en la Figura 3.19, las inferencias del sistema indican que, como es esperado, Liverpool es el campeón de la Premier League.

El ejemplo anterior muestra que diferentes interpretaciones de soporte pueden dar lugar a diferentes resultados en las inferencias de un sistema. Es por esto que no toda interpretación de soporte resulta adecuada para el correcto modelado de todos los escenarios posibles. En contraste, el siguiente ejemplo muestra que una interpretación que puede parecer apropiada para modelar una situación en particular, posteriormente puede dar lugar a resultados incorrectos o no deseados.

Ejemplo 3.14 *Consideremos un escenario similar al descripto por el Ejemplo 3.13, con la siguiente diferencia: Manchester United gana su último partido ($\mathcal{MUG}$). Dado este nuevo escenario, no hay motivos para adoptar una interpretación de soporte diferente al soporte deductivo empleado en el Ejemplo 3.13. Entonces, dada la interpretación de soporte deductivo entre argumentos, puede adoptarse la representación ilustrada en la Figura 3.20. En este caso, además de la derrota de $\mathcal{LP}$ sobre $\mathcal{LG}$, existe una derrota de $\mathcal{MUG}$ sobre $\mathcal{MUNG}$. Por otra parte, dada la interpretación de soporte deductivo, no hay derrotas complejas. Nótese que esta representación conduce al resultado incorrecto de que el argumento $\mathcal{LCPL}$ está aceptado, indicando que Liverpool es el campeón de la Premier League aún cuando perdió su último partido y Manchester United ganó el propio. En contraste, si adoptamos una interpretación de soporte por necesidad, el resultado indica que, como es esperado, Liverpool no es el campeón de la Premier League. La Figura 3.21 ilustra la representación asociada a esta situación utilizando soporte por necesidad. En este caso, existen derrotas extendidas de $\mathcal{LP}$ y $\mathcal{MUG}$ sobre $\mathcal{LCPL}$, conduciendo a la no aceptación del argumento $\mathcal{LCPL}$.*

$$\begin{tikzcd} \mathcal{LG} \arrow[r, Rightarrow, "d"] & \mathcal{LCPL} & \mathcal{MUNG} \arrow[l, Rightarrow, "d"'] \\ \mathcal{LP} \arrow[u] & & \mathcal{MUG} \arrow[u] \end{tikzcd}$$

Figura 3.20: Marco Argumentativo Bipolar con Soporte Deductivo correspondiente al Ejemplo 3.14.

$$\begin{tikzcd} \mathcal{LG} \arrow[r, Rightarrow, "n"] & \mathcal{LCPL} & \mathcal{MUNG} \arrow[l, Rightarrow, "n"'] \\ \mathcal{LP} \arrow[u] \arrow[ur, dashed, bend right] & & \mathcal{MUG} \arrow[u] \arrow[ul, dashed, bend left] \end{tikzcd}$$

Figura 3.21: Marco Argumentativo con Soporte por Necesidad correspondiente al Ejemplo 3.14.

Como se puede observar en los ejemplos anteriores, la elección de una interpretación apropiada para expresar soporte entre argumentos no es una tarea fácil. Si bien es necesario analizar la relación de soporte para cada escenario en particular, esto no es suficiente. Como se vio en el Ejemplo 3.13, la interpretación deductiva resultó adecuada para el modelado del soporte entre argumentos. Sin embargo, en el Ejemplo 3.14 se presentó un escenario similar en el cual, sin modificar los vínculos de soporte existentes entre argumentos (y por lo tanto manteniendo la interpretación de soporte adoptada), el añadido de nueva información condujo a la obtención de resultados no deseados. En consecuencia, la incorporación de una derrota directa resultó en la necesidad de cambiar la interpretación de soporte adoptada, lo cual claramente es una característica no deseada para un sistema argumentativo.

Cabe destacar que un aspecto relevante detrás de este problema es que, al tratarse de sistemas argumentativos abstractos, no existen pautas a seguir en la representación de conocimiento. Es por esto que, dada la representación adoptada para modelar un escenario, algunas interpretaciones de soporte resultan más apropiadas que otras. Para ilustrar esto, consideremos el siguiente ejemplo.

Ejemplo 3.15 *Dado el escenario introducido en el Ejemplo 3.13, consideremos la siguiente representación de conocimiento. En lugar de contar con el argumento "Liverpool*

$$\begin{tikzcd} \mathcal{LG} \arrow[r, Rightarrow, "n"] & \mathcal{LCPL}_1 \\ \mathcal{LP} \arrow[u] \arrow[ur, dashed, bend right] & \end{tikzcd} \qquad \begin{tikzcd} \mathcal{LCPL}_2 & \mathcal{MUNG} \arrow[l, Rightarrow, "n"'] \end{tikzcd}$$

Figura 3.22: Marco Argumentativo con Soporte por Necesidad correspondiente al Ejemplo 3.15.

es el campeón de la Premier League" ($\mathcal{LCPL}$), el cual es soportado por los argumentos "Liverpool gana su último partido" ($\mathcal{LG}$) y "Manchester United no gana su último partido" ($\mathcal{MUNG}$), supongamos que contamos con dos argumentos para representar que Liverpool puede ganar la Premier League de dos formas. El primer argumento expresa que "Liverpool es el campeón de la Premier League por ganar su último partido" ($\mathcal{LCPL}_1$), el cual es soportado por el argumento $\mathcal{LG}$. El otro argumento expresa que "Liverpool es el campeón de la Premier League provisto que Manchester United no gane su último partido" ($\mathcal{LCPL}_2$), el cual es a su vez soportado por el argumento $\mathcal{MUNG}$. En este caso, podemos adoptar una interpretación de soporte por necesidad, dando lugar a la representación ilustrada en la Figura 3.22.

Dada esta representación, existe una derrota extendida de $\mathcal{LP}$ sobre $\mathcal{LCPL}_1$, denotada en la Figura 3.22 mediante una flecha punteada. Luego, las inferencias del sistema en este caso determinan que, como es esperado, Liverpool es el campeón de la Premier League porque el argumento $\mathcal{LCPL}_2$ está aceptado. Por otra parte, nótese que la no aceptación del argumento $\mathcal{LCPL}_1$ refleja el hecho de que Liverpool ganó la Premier League no por haber ganado su último partido, sino porque Manchester United no ganó el propio.

3.4. Sistemas Argumentativos Evidenciales (**EAS**)

Los argumentos son piezas de razonamiento que permiten arribar a una conclusión a partir de un conjunto de premisas. Usualmente, las teorías argumentativas asumen que estas premisas (y, por lo tanto, los argumentos a los que pertenecen) siempre valen ya que los marcos argumentativos representan una imagen instantánea de los argumentos y las relaciones involucradas en el proceso de razonamiento. Sin embargo, aproximaciones alternativas como [ON08] consideran que los argumentos deben estar soportados por evidencia. El razonamiento evidencial involucra determinar qué argumentos son aplicables a partir de la evidencia disponible. De esta manera, la aproximación de *soporte evidencial* propuesta en [ON08] intenta capturar una noción particular: un argumento no puede estar aceptado a menos que esté *soportado por evidencia*.

En [ON08] los autores introdujeron los *Sistemas Argumentativos Evidenciales (EAS)* (por su nombre en inglés, *Evidential Argumentation Systems*) que extienden a los marcos argumentativos de Dung incorporando una relación de soporte especializada que captura la noción de soporte por evidencia. Esta relación de soporte permite distinguir entre argumentos *prima facie* y argumentos estándar. Los argumentos *prima facie* no requieren del soporte de otros argumentos para subsistir, mientras que los argumentos estándar deben estar vinculados mediante una cadena de soporte con al menos un argumento *prima facie*. Adicionalmente, el sistema propuesto en [ON08], inspirado en la propuesta de [NP06] donde múltiples argumentos pueden atacarse entre sí, permite la existencia de derrotas y soportes originados en un conjunto de argumentos.

Como se mencionó anteriormente, un argumento en un EAS estará aceptado únicamente si está soportado por una cadena de argumentos, cada uno de los cuales está finalmente soportado por evidencia. Al comienzo de esta cadena de soporte hay un argumento especial η que representa el soporte del entorno (*i. e.*, la existencia de evidencia que los sustenta). Para capturar esta intuición, en [ON08] los autores definen un sistema argumentativo evidencial de la siguiente manera.

Definición 3.12 (Sistema Argumentativo Evidencial) *Un* Sistema Argumentativo Evidencial (EAS) *es una tupla* $\langle \mathbb{A}, \mathbb{R}_d, \mathbb{R}_e \rangle$*, donde* $\mathbb{A}$ *es un conjunto de argumentos,* $\mathbb{R}_d \subseteq (2^{\mathbb{A}} \backslash \{\}) \times \mathbb{A}$ *es una relación de derrota, y* $\mathbb{R}_e \subseteq 2^{\mathbb{A}} \times \mathbb{A}$ *es una relación de soporte. En particular, se tiene que* $\nexists X \in 2^{\mathbb{A}}$*,* $\nexists \mathcal{Y} \in \mathbb{A}$ *tales que* $X \mathbb{R}_d \mathcal{Y}$ *y* $X \mathbb{R}_e \mathcal{Y}$*. Además, se asume la existencia de un argumento "especial"* $\eta \notin \mathbb{A}$*.*

Un elemento de la forma $\{\eta\}\mathbb{R}_e\mathcal{A}$ perteneciente a la relación de soporte representa el soporte del entorno para $\mathcal{A}$. Luego, $\mathcal{A}$ es llamado argumento *prima facie* ya que se encuentra soportado por evidencia. Dado que el entorno no requiere de soporte alguno, el argumento especial η no puede aparecer como el segundo elemento de un par en la relación de soporte $\mathbb{R}_e$. Además, cualquier argumento que se encuentre derrotado por el entorno estaría derrotado incondicionalmente, por lo que no tendría sentido considerarlo.

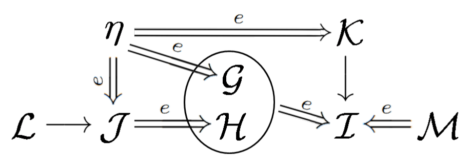


Figura 3.23: Sistema Argumentativo Evidencial $\mathsf{EAS}_{3.16}$ del Ejemplo 3.16.

Es por esto que el EAS no permite que el argumento especial η aparezca en la relación de derrota[6].

En [ON08] los autores emplean $S \dashrightarrow \mathcal{A}$ y $S \longrightarrow \mathcal{A}$ para denotar, respectivamente, que el conjunto S derrota y soporta al argumento $\mathcal{A}$. Siguiendo la convención adoptada para la notación gráfica introducida en las secciones previas, en esta tesis utilizaremos S $\longrightarrow \mathcal{A}$ para denotar $S\mathbb{R}_d\mathcal{A}$ (derrota) y S $\stackrel{e}{\Longrightarrow} \mathcal{A}$ para denotar $S\mathbb{R}_e\mathcal{A}$ (soporte evidencial). En este caso, la etiqueta 'e' sobre la flecha doble indica que la relación de soporte adopta una interpretación evidencial.

Ejemplo 3.16 *Sea* $\mathsf{EAS}_{3.16} = \langle \mathbb{A}_{3.16}, \mathbb{R}_{d_{3.16}}, \mathbb{R}_{e_{3.16}} \rangle$ *un sistema argumentativo evidencial, donde:*

$$\begin{aligned}
\mathbb{A}_{3.16} &= \{\mathcal{G},\ \mathcal{H},\ \mathcal{I},\ \mathcal{J},\ \mathcal{K},\ \mathcal{L},\ \mathcal{M}\} \\
\mathbb{R}_{d_{3.16}} &= \{(\{\mathcal{L}\},\mathcal{J}),(\{\mathcal{K}\},\mathcal{I})\} \\
\mathbb{R}_{e_{3.16}} &= \{(\{\eta\},\mathcal{G}),(\{\eta\},\mathcal{J}),(\{\eta\},\mathcal{K}),(\{\mathcal{J}\},\mathcal{H}),(\{\mathcal{G},\mathcal{H}\},\mathcal{I}),(\{\mathcal{M}\},\mathcal{I})\}
\end{aligned}$$

La representación gráfica de $\mathsf{EAS}_{3.16}$ *se halla ilustrada en la Figura 3.23.*

La relación de soporte de un EAS no captura por completo la intuición de que los argumentos deben estar soportados por evidencia, sino que simplemente expresa que un argumento se encuentra soportado por un conjunto de argumentos. Para capturar esta restricción, en [ON08] se define explícitamente la noción de *soporte evidencial* como sigue.

Definición 3.13 (Soporte Evidencial) *Sea* $\langle \mathbb{A}, \mathbb{R}_d, \mathbb{R}_e \rangle$ *un* EAS*. Un argumento* $\mathcal{A} \in \mathbb{A}$ *está* soportado evidencialmente (e-soportado) *por un conjunto* S *si y sólo si se satisface una de las siguientes condiciones:*

[6] En [ON08] los autores utilizan la terminología relación de 'ataque'. No obstante esto, siguiendo la convención adoptada en esta tesis, para evitar confusiones nos referiremos a $\mathbb{R}_d$ como relación de 'derrota'.

- $S\mathbb{R}_e\mathcal{A}$, *donde* $S = \{\eta\}$; *o*
- $\exists T \subset S$ *tal que* $T\mathbb{R}_e\mathcal{A}$ *y* $\forall \mathcal{B} \in T, \mathcal{B}$ *está e-soportado por* $S\backslash\{\mathcal{B}\}$.

Además, se dice que $\mathcal{A}$ *está e-soportado minimalmente por* S *si y sólo si* $\nexists S' \subset S$ *tal que* $\mathcal{A}$ *está e-soportado por* S'.

La noción de soporte evidencial para un argumento $\mathcal{A}$ requiere de evidencia (el argumento especial η) al inicio de la cadena de soporte que conduce, a través de varios argumentos, al argumento $\mathcal{A}$. Luego, basándose en esta noción de soporte evidencial, se definen las *derrotas soportadas por evidencia* con el objetivo de capturar aquellas derrotas que se encuentran respaldadas por evidencia.

Definición 3.14 (Derrota Soportada por Evidencia) *Sea* $\langle \mathbb{A}, \mathbb{R}_d, \mathbb{R}_e \rangle$ *un EAS y* $\mathcal{A} \in \mathbb{A}$. *Un conjunto* S *efectúa una* derrota soportada por evidencia (derrota e-soportada) *sobre* $\mathcal{A}$ *si y sólo si se satisfacen las siguientes condiciones:*

- $T\mathbb{R}_d\mathcal{A}$, *donde* $T \subseteq S$; *y*
- $\forall \mathcal{B} \in T$, $\mathcal{B}$ *está e-soportado por* S.

Además, una derrota e-soportada de S *sobre* $\mathcal{A}$ *se dice minimal si y sólo si* $\nexists S' \subset S$ *tal que* S' *efectúa una derrota e-soportada sobre* $\mathcal{A}$.

Por ejemplo, dado el sistema argumentativo evidencial $\mathsf{EAS}_{3.16}$ del Ejemplo 3.16, el argumento $\mathcal{I}$ está e-soportado por el conjunto $\{\eta, \mathcal{G}, \mathcal{H}, \mathcal{J}\}$, y existe una derrota e-soportada de $\{\eta, \mathcal{K}\}$ sobre $\mathcal{I}$. Nótese que $\mathcal{I}$ no está e-soportado por $\{\mathcal{M}\}$, ya que $\mathcal{M}$ no está a su vez e-soportado. De manera similar, si bien $\{\mathcal{L}\} \longrightarrow \mathcal{J}$, esta no es una derrota e-soportada ya que $\mathcal{L}$ no está e-soportado.

Siguiendo la aproximación de [Dun95], un argumento $\mathcal{A}$ de un EAS se dirá aceptable con respecto a un conjunto S si el conjunto defiende a $\mathcal{A}$ ante toda derrota que recibe. Sin embargo, los autores de [ON08] señalan que, dada la existencia de evidencia, el cómputo de aceptabilidad de argumentos en un EAS debe considerar únicamente las derrotas e-soportadas. Luego, para defender a un argumento ante una derrota e-soportada deberá derrotarse a dicha derrota. Para esto, se consideran las derrotas e-soportadas minimales, las cuales pueden ser derrotadas mediante la derrota de cualquier argumento perteneciente al conjunto minimal. Adicionalmente, los autores establecen que un requerimiento esencial para los argumentos aceptables es que deben estar soportados por evidencia.

Definición 3.15 (Aceptabilidad) *Sea* $\langle \mathbb{A}, \mathbb{R}_d, \mathbb{R}_e \rangle$ *un EAS y* $\mathcal{A} \in \mathbb{A}$*. El argumento* $\mathcal{A}$ *es* aceptable con respecto al conjunto S *si y sólo si se verifican las siguientes condiciones:*

- S *e-soporta a* $\mathcal{A}$*; y*
- *dada una derrota e-soportada minimal de* X *sobre el argumento* $\mathcal{A}$*,* $\exists T \subseteq S$ *tal que* $T\mathbb{R}_d\mathcal{X}$*, donde* $\mathcal{X} \in X$*, de manera que* $X\backslash\{\mathcal{X}\}$ *no efectúa una derrota e-soportada sobre* $\mathcal{A}$*.*

Luego, los autores de [ON08] continúan con la aproximación de Dung, definiendo un conjunto admisible S como un conjunto libre de conflictos (con respecto a $\mathbb{R}_d$) tal que cada argumento de S es aceptable con respecto a S. De esta manera, una *extensión preferida evidencial (extensión e-preferida)* de un EAS es caracterizada como un conjunto admisible maximal. Por último, los autores remarcan que también es posible computar las extensiones preferidas de un EAS siguiendo la exacta definición propuesta por Dung en [Dun95]. Para esto, simplemente debe restringirse la relación de soporte de modo que únicamente exista soporte entre el argumento especial η y todos los demás argumentos del sistema.

Por ejemplo, dado el sistema argumentativo evidencial $\mathsf{EAS}_{3.16}$ ilustrado en el Ejemplo 3.16, el argumento $\mathcal{H}$ es aceptable con respecto a $\{\eta, \mathcal{J}\}$ ya que está e-soportado por el conjunto y no existen derrotas e-soportadas contra él. Por el contrario, el argumento $\mathcal{I}$ no es aceptable con respecto a $\{\eta, \mathcal{G}, \mathcal{H}, \mathcal{J}\}$ porque el conjunto no defiende a $\mathcal{I}$ ante la derrota e-soportada de $\mathcal{K}$. No obstante esto, el conjunto $\{\eta, \mathcal{G}, \mathcal{H}, \mathcal{J}\}$ es admisible ya que es libre de conflictos, todos los argumentos en el conjunto están e-soportados por el conjunto, y no existen derrotas e-soportadas hacia ninguno de los argumentos del conjunto. En contraste, el conjunto $\{\eta, \mathcal{G}, \mathcal{H}, \mathcal{J}, \mathcal{M}\}$ no es admisible dado que, si bien es libre de conflictos y no existen derrotas e-soportadas contra los argumentos del conjunto, el argumento $\mathcal{M}$ no es aceptable con respecto al conjunto al no estar e-soportado por él. Finalmente, la única extensión e-preferida de $\mathsf{EAS}_{3.16}$ es $\{\eta, \mathcal{G}, \mathcal{H}, \mathcal{J}, \mathcal{K}\}$.

$$\begin{tikzcd}
 & \eta \arrow[ddl, Rightarrow, "e"'] \arrow[d, Rightarrow, "e"'] \arrow[dr, Rightarrow, "e"] & \\
 & \mathcal{MJ} \arrow[d, Rightarrow, "e"'] & \mathcal{BH} \arrow[d, Rightarrow, "e"'] \\
\mathcal{GL} \arrow[r] & \mathcal{BP} \arrow[r, Rightarrow, "e"] & \mathcal{B}
\end{tikzcd}$$

Figura 3.24: Sistema Argumentativo Evidencial $\mathsf{EAS}_{3.17}$, correspondiente al Ejemplo 3.17.

En lo que resta de esta sección discutiremos la relación existente entre el soporte evidencial y las interpretaciones de soporte presentadas en las secciones previas. Una diferencia clara entre la noción de soporte evidencial propuesta en [ON08] y las aproximaciones de [CLS05], [BGvdTV10] y [NR10] presentadas en las secciones 3.1, 3.2 y 3.3 respectivamente, es que la primera expresa el soporte que un conjunto de argumentos provee a otro argumento, mientras que las demás relacionan únicamente pares de argumentos. Teniendo en cuenta esto, es posible analizar la relación existente entre estas interpretaciones de soporte en el contexto de un sistema argumentativo evidencial en que los argumentos son derrotados y soportados por conjuntos unitarios de argumentos[7]. Para abordar esta comparación, consideremos el siguiente ejemplo inspirado en [NR10].

Ejemplo 3.17 *Consideremos un escenario donde un estudiante está analizando si obtendrá una beca para una universidad. Los estatutos de la universidad indican que un estudiante obtendrá una beca* ($\mathcal{B}$) *si obtuvo un título de bachiller con honores* ($\mathcal{BH}$) *o posee un bajo presupuesto* ($\mathcal{BP}$)*. Supongamos que el estudiante ha obtenido un bachillerato con honores y que tiene un bajo presupuesto dado que posee un trabajo de media jornada* ($\mathcal{MJ}$)*. Además, supongamos que el estudiante ganó un premio de un millón de pesos en la lotería* ($\mathcal{GL}$) *con anterioridad a la solicitud de la beca. Esta situación puede ser representada por el sistema argumentativo evidencial* $\mathsf{EAS}_{3.17}$ *ilustrado en la Figura 3.24. En este contexto, es esperable que el estudiante obtenga la beca dado que posee un grado de bachiller con honores. Esto se corresponde con las inferencias del sistema, ya que* $\{\eta, \mathcal{GL}, \mathcal{MJ}, \mathcal{BH}, \mathcal{B}\}$ *es la extensión e-preferida de* $\mathsf{EAS}_{3.17}$.

[7] Para lograr una mayor claridad, simplificaremos la notación de modo tal que un conjunto unitario de argumentos $\{\mathcal{A}\}$ será notado como $\mathcal{A}$.

$$\begin{tikzcd} & \mathcal{MJ} \arrow[d, Rightarrow, "n"'] & \mathcal{BH} \arrow[d, Rightarrow, "n"'] \\ \mathcal{GL} \arrow[r] \arrow[rr, dashed, bend right] & \mathcal{BP} \arrow[r, Rightarrow, "n"] & \mathcal{B} \end{tikzcd}$$

Figura 3.25: Marco Argumentativo con Soporte por Necesidad $\mathsf{AFN}_{3.17}$, correspondiente a la situación descripta en el Ejemplo 3.17.

Supongamos ahora que queremos representar la situación descripta en el Ejemplo 3.17 utilizando la interpretación de soporte por necesidad. Esta situación puede ser caracterizada por marco argumentativo con soporte por necesidad $\mathsf{AFN}_{3.17}$, ilustrado en la Figura 3.25. Observemos que esta nueva representación no incluye el argumento especial η. Esto se debe a que, como todos los argumentos en $\mathsf{EAS}_{3.17}$ están directa o indirectamente soportados por evidencia (*i. e.*, están e-soportados), es posible omitir la representación de η en el AFN correspondiente.

Dada la interpretación de soporte por necesidad, ocurre una derrota extendida de $\mathcal{GL}$ sobre $\mathcal{B}$ (denotada en la Figura 3.25 mediante una flecha punteada), resultando en que $\mathcal{B}$ no pertenece a la extensión preferida $\{\mathcal{GL}, \mathcal{MJ}, \mathcal{BH}\}$ de $\mathsf{AFN}_{3.17}$. Claramente, este es un resultado no deseado, el cual surge de la interpretación adoptada para la relación de soporte y su representación asociada. Recordemos que, bajo la interpretación de soporte por necesidad, si $\mathcal{B} \stackrel{n}{\Longrightarrow} \mathcal{A}$ y $\mathcal{C} \stackrel{n}{\Longrightarrow} \mathcal{A}$, entonces tanto $\mathcal{B}$ como $\mathcal{C}$ deben estar aceptados para que $\mathcal{A}$ esté aceptado. En particular, esto no ocurre en el escenario descripto por el Ejemplo 3.17, porque el estudiante puede obtener la beca ya sea por haber obtenido un bachillerato con honores o por contar con bajo presupuesto.

Analicemos qué sucede al representar la situación descripta por el Ejemplo 3.17 utilizando una interpretación deductiva de soporte. En tal caso, se obtiene la caracterización del marco argumentativo bipolar con soporte deductivo $\mathsf{d\text{-}BAF}_{3.17}$, ilustrado en la Figura 3.26. Al igual que en el caso anterior, el argumento η correspondiente al entorno no es incluido en la representación del $\mathsf{d\text{-}BAF}$. Luego, dada la interpretación deductiva de soporte, deben considerarse las derrotas mediadas y las derrotas soportadas. Por lo tanto, ocurre una derrota mediada de $\mathcal{GL}$ sobre $\mathcal{MJ}$, ilustrada en la Figura 3.26 con una flecha punteada. Sin embargo, esta derrota no afecta la pertenencia de $\mathcal{B}$ a la extensión preferida $\{\mathcal{GL}, \mathcal{BH}, \mathcal{B}\}$ de $\mathsf{d\text{-}BAF}_{3.17}$. Por último cabe destacar que, si bien la aproximación

$$
\begin{tikzcd}
 & \mathcal{MJ} \arrow[d, Rightarrow, "d"'] & \mathcal{BH} \arrow[d, Rightarrow, "d"'] \\
\mathcal{GL} \arrow[ur, dashed, bend left] \arrow[r] & \mathcal{BP} \arrow[r, Rightarrow, "d"] & \mathcal{B}
\end{tikzcd}
$$

Figura 3.26: Marco Argumentativo Bipolar con Soporte Deductivo $\mathsf{d\text{-}BAF}_{3.17}$, correspondiente al a situación descripta en el Ejemplo 3.17.

$$
\begin{tikzcd}
 & \eta \arrow[d, Rightarrow, "e"'] \arrow[ddl, Rightarrow, "e"'] & \\
 & \mathcal{MJ} \arrow[d, Rightarrow, "e"'] & \\
\mathcal{GL} \arrow[r] & \mathcal{BP} \arrow[r, Rightarrow, "e"] & \mathcal{B}
\end{tikzcd}
$$

Figura 3.27: Sistema Argumetativo Evidencial $\mathsf{EAS}_{3.18}$ del Ejemplo 3.18.

de soporte deductivo conduce al resultado esperado en esta situación, no resulta apropiada para representar correctamente algunos escenarios en los que se expresa soporte evidencial. Para ilustrar esto consideremos el siguiente ejemplo.

Ejemplo 3.18 *Supongamos un escenario similar al presentado en el Ejemplo 3.17, donde la única forma de obtener una beca es contar con un bajo presupuesto. Al igual que en el Ejemplo 3.17, el estudiante que solicita la beca posee un trabajo de media jornada y ganó la lotería recientemente. Esta nueva situación puede representarse mediante el sistema argumentativo evidencial* $\mathsf{EAS}_{3.18}$ *ilustrado en la Figura 3.27. En este caso, es esperable que el estudiante no obtenga la beca dado que por haber ganado la lotería ya no cuenta con un bajo presupuesto. Esto es capturado por la derrota e-soportada de* $\mathcal{GL}$ *sobre* $\mathcal{BP}$*, la cual evita que* $\mathcal{B}$ *esté e-soportado. De esta manera, el argumento* $\mathcal{B}$ *no pertenece a la extensión e-preferida* $\{\eta, \mathcal{GL}, \mathcal{MJ}\}$ *de* $\mathsf{EAS}_{3.18}$*, significando entonces que el estudiante no obtuvo la beca.*

Supongamos que se desea modelar la situación descripta en el Ejemplo 3.18 utilizando soporte deductivo. A tal efecto, consideremos el $\mathsf{d\text{-}BAF}_{3.18}$ ilustrado en la Figura 3.28. Dada esta representación, existe una derrota mediada de $\mathcal{GL}$ sobre $\mathcal{MJ}$. Además, $\mathcal{BP}$ no es un

$$\begin{tikzcd}
& \mathcal{MJ} \arrow[d, Rightarrow, "d"'] & \\
\mathcal{GL} \arrow[r] \arrow[ur, dashed, bend left] & \mathcal{BP} \arrow[r, Rightarrow, "d"] & \mathcal{B}
\end{tikzcd}$$

Figura 3.28: Marco Argumentativo Bipolar con Soporte Deductivo $\mathsf{d\text{-}BAF}_{3.18}$, correspondiente a la situación descripta en el Ejemplo 3.18.

$$\begin{tikzcd}
& \mathcal{MJ} \arrow[d, Rightarrow, "s"'] & \mathcal{BH} \arrow[d, Rightarrow, "s"'] \\
\mathcal{GL} \arrow[r] \arrow[rr, dashed, bend right] & \mathcal{BP} \arrow[r, Rightarrow, "s"] & \mathcal{B}
\end{tikzcd}$$

Figura 3.29: Marco Argumentativo Bipolar $\mathsf{BAF}_{3.19}$ del Ejemplo 3.19.

argumento aceptado de $\mathsf{d\text{-}BAF}_{3.18}$ ya que se encuentra derrotado por $\mathcal{GL}$. Sin embargo, el argumento $\mathcal{B}$ pertenece a la extensión preferida $\{\mathcal{GL}, \mathcal{B}\}$ de $\mathsf{d\text{-}BAF}_{3.18}$, lo cual es claramente un resultado no deseado. Recordemos que, dado $\mathcal{A} \overset{d}{\Longrightarrow} \mathcal{B}$, si $\mathcal{A}$ está aceptado entonces $\mathcal{B}$ también está aceptado, y si $\mathcal{B}$ no está aceptado entonces $\mathcal{A}$ tampoco está aceptado. En consecuencia, dado $\mathcal{A} \overset{d}{\Longrightarrow} \mathcal{B}$, es posible que el argumento $\mathcal{B}$ esté aceptado aún cuando el argumento $\mathcal{A}$ no lo está. Por lo tanto, no resultaría apropiado utilizar una interpretación deductiva de soporte en escenarios como el propuesto en el Ejemplo 3.18. Por otra parte, si se intenta modelar soporte evidencial utilizando la relación de soporte general del BAF (ver Sección 3.1), también puede ocurrir que se obtengan resultados no deseados. Para ilustrar esto, consideremos el siguiente ejemplo.

Ejemplo 3.19 *Supongamos que se quiere representar la situación descripta en el Ejemplo 3.17 utilizando un* BAF. *En tal caso, se obtiene el marco argumentativo bipolar* $\mathsf{BAF}_{3.19}$, *cuyo grafo de interacción bipolar se ilustra en la Figura 3.29. Al igual que en el caso del* AFN *y el* $\mathsf{d\text{-}BAF}$, *el* $\mathsf{BAF}_{3.19}$ *no incluye al argumento especial* η. *Dada esta representación, ocurre una derrota secundaria de* $\mathcal{GL}$ *sobre* $\mathcal{B}$, *conduciendo a la no aceptación del argumento* $\mathcal{B}$. *En consecuencia, la única extensión preferida de* $\mathsf{BAF}_{3.19}$ *es* $\{\mathcal{GL}, \mathcal{MJ}, \mathcal{BH}\}$. *Claramente, este es un resultado no deseado ya que el estudiante posee una forma alternativa para obtener la beca, dado que obtuvo un bachillerato con honores.*

Finalmente, hay otra característica que diferencia la aproximación evidencial de [ON08] de aquellas presentadas en las secciones 3.1, 3.2 y 3.3. Como se indica en la Definición 3.15, un argumento de un EAS será aceptable con respecto a un conjunto de argumentos si es defendido por el conjunto ante toda derrota e-soportada que recibe. Sin embargo, existe un requerimiento adicional que los argumentos aceptables deben satisfacer: deben estar soportados por evidencia (*i. e.*, deben estar directa o indirectamente soportados por el argumento especial η). Esto implica que, dado un EAS en el que no hay derrotas entre argumentos (*i. e.*, $\mathbb{R}_d = \emptyset$), puede ocurrir que algunos argumentos no estén aceptados. Específicamente, aquellos argumentos que no estén soportados por evidencia no formarán parte de las extensiones del sistema. De manera similar a los sistemas argumentativos evidenciales, en [RMF$^+$08, RMGS10] se propone una aproximación que distingue entre argumentos activos y argumentos potenciales o inactivos, dependiendo de si se encuentran soportados por evidencia o no. En contraste, en las propuestas de [CLS05], [BGvdTV10] y [NR10] (presentadas en las secciones 3.1, 3.2 y 3.3), en caso de no existir derrotas entre argumentos todos los argumentos del sistema estarán aceptados.

3.5. Otras Aproximaciones al uso de Soporte en Argumentación

En esta sección analizaremos otras propuestas existentes en la literatura que se relacionan con el tópico de estudio de este capítulo. Las aproximaciones presentadas en las secciones previas corresponden a formalismos de argumentación abstracta que consideran explícitamente una relación de soporte entre argumentos, adoptando diferentes interpretaciones para la noción de soporte. En contraste, en esta sección analizaremos otras propuestas que, si bien permiten expresar soporte entre argumentos, no consideran la existencia de soporte en forma explícita. En primer lugar, analizaremos la propuesta de [MGS06], donde se define una extensión de los marcos argumentativos de [Dun95] que incorpora, entre otras cosas, una relación de sub-argumento. Luego, consideraremos la propuesta de [BW10], donde se provee una generalización de los marcos argumentativos de Dung que incorpora condiciones de aceptación para los argumentos del sistema.

3.5.1. Sub-argumento como Soporte

En la literatura existe una gran variedad de aproximaciones que abordan la noción de sub-argumento, tales como [SL92, PS97, GS04, MGS06, Pra09]. En el contexto de un determinado argumento, puede ser necesario proveer sustento para las conclusiones intermedias, requiriendo la construcción de sub-argumentos para ellas. De esta manera, una premisa de un argumento se transforma en la conclusión de otro argumento. Por lo tanto, un sub-argumento es un argumento subordinado que provee sustento para una premisa, siendo un componente de un argumento más grande. En otras palabras, un sub-argumento constituye una línea de razonamiento que contribuye a la obtención de una conclusión, proveyendo entonces alguna clase de soporte para sus super-argumentos.

Si bien los trabajos arriba citados no consideran explícitamente a la relación de sub-argumento como una relación de soporte entre argumentos, es claro que, de alguna manera, contemplan que los sub-argumentos proveen soporte para sus super-argumentos. En particular, en [MGS06] se propone un marco argumentativo abstracto donde la relación de sub-argumento es explícita. Cabe destacar que el objetivo principal de dicho trabajo (y de subsiguientes trabajos como [MGS07, MGS08b, MGS08a]) concierne el estudio de la fuerza de los argumentos y de cómo esta afecta la aceptabilidad de los mismos.

En esta sección presentaremos la relación de sub-argumento de [MGS06]. Luego, discutiremos su relación con respecto a una interpretación particular de soporte entre argumentos: el soporte por necesidad presentado en la Sección 3.3. En [MGS06] los autores introducen un nuevo marco argumentativo que extiende el formalismo de [Dun95] mediante la incorporación de una relación de sub-argumento y una relación de preferencia entre argumentos. Por otra parte, en lugar de la relación de derrota de los marcos argumentativos de Dung, proponen una relación de conflicto simétrica que, al ser combinada con la relación de preferencia, permite obtener las derrotas entre los argumentos del sistema.

Definición 3.16 (Marco Argumentativo Abstracto con Sub-argumentos) *Un* Marco Argumentativo Abstracto con Sub-argumentos (AFS)[8] *es una tupla* $\langle \mathbb{A}, \sqsubseteq, \mathbb{C}, \mathbb{R} \rangle$,

[8]En [MGS06] los autores lo llaman *Marco Argumentativo Abstracto*. Sin embargo, para evitar confusiones con el Marco Argumentativo Abstracto de [Dun95], se decidió renombrarlo como *Marco Argumentativo Abstracto con Sub-argumentos (AFS)* (en inglés, *Abstract Argumentation Framework with Sub-arguments*).

donde $\mathbb{A}$ *es un conjunto finito de argumentos,* $\sqsubseteq\ \subseteq \mathbb{A} \times \mathbb{A}$ *es una relación de sub-argumento transitiva,* $\mathbb{C} \subseteq \mathbb{A} \times \mathbb{A}$ *es una relación de conflicto simétrica e irreflexiva, y* $\mathbb{R} \subseteq \mathbb{A} \times \mathbb{A}$ *es una relación de preferencia.*

La existencia de un par $(\mathcal{A}, \mathcal{B})$ en la relación de sub-argumento $\sqsubseteq$ expresa que $\mathcal{A}$ es un sub-argumento de $\mathcal{B}$ y, recíprocamente, $\mathcal{B}$ es un super-argumento de $\mathcal{A}$. En particular, todo argumento $\mathcal{A}$ es considerado un super-argumento (respectivamente, un super-argumento) de sí mismo. Por otra parte, la relación de conflicto $\mathbb{C}$ entre dos argumentos $\mathcal{A}$ y $\mathcal{B}$ denota el hecho de que estos argumentos no pueden ser simultáneamente aceptados dado que se contradicen entre sí. Por lo tanto, la relación de conflicto $\mathbb{C}$ de un AFS es simétrica. Además, la relación de conflicto debe exhibir un comportamiento racional con respecto a los sub-argumentos. Por este motivo, en [MGS06] se establece que la relación de conflicto debe satisfacer la propiedad de *herencia de conflicto*.

Definición 3.17 (Herencia de Conflicto) *Sea* $\langle \mathbb{A}, \sqsubseteq, \mathbb{C}, \mathbb{R} \rangle$ *un AFS y* $\mathcal{A}, \mathcal{B} \in \mathbb{A}$*. Si* $(\mathcal{A}, \mathcal{B}) \in \mathbb{C}$*, entonces* $(\mathcal{A}, \mathcal{B}_1) \in \mathbb{C}$*,* $(\mathcal{A}_1, \mathcal{B}) \in \mathbb{C}$ *y* $(\mathcal{A}_1, \mathcal{B}_1) \in \mathbb{C}$*, para todo par de argumentos* $\mathcal{A}_1$ *y* $\mathcal{B}_1$ *tales que* $\mathcal{A} \sqsubseteq \mathcal{A}_1$*,* $\mathcal{B} \sqsubseteq \mathcal{B}_1$*.*

Esta propiedad establece que si un argumento $\mathcal{A}$ está en conflicto con un argumento $\mathcal{B}$, entonces el conflicto sigue presente al considerar los super-argumentos de $\mathcal{A}$ y $\mathcal{B}$.

Ejemplo 3.20 *Consideremos un marco argumentativo abstracto con sub-argumentos* $\mathit{AFS}_{3.20} = \langle \mathbb{A}_{3.20}, \sqsubseteq_{3.20}, \mathbb{C}_{3.20}, \mathbb{R}_{3.20} \rangle$ *y sean* $\mathcal{N}, \mathcal{O} \in \mathbb{A}_{3.20}$ *tales que* $(\mathcal{N}, \mathcal{O}) \in \mathbb{C}_{3.20}$*,* $\mathcal{N} \sqsubseteq_{3.20} \mathcal{N}_1$*, y* $\mathcal{O} \sqsubseteq_{3.20} \mathcal{O}_1$*. Como* $\mathbb{C}_{3.20}$ *es simétrica y satisface la propiedad de herencia de conflicto, se tienen los conflictos ilustrados en la Figura 3.30.*

Con respecto a la relación de preferencia, en [MGS06] se introduce la siguiente notación: si $\mathcal{A}\mathbb{R}\mathcal{B}$ pero no ocurre que $\mathcal{B}\mathbb{R}\mathcal{A}$, entonces $\mathcal{A}$ es preferido ante $\mathcal{B}$, notado como $\mathcal{A} \succ \mathcal{B}$; si $\mathcal{A}\mathbb{R}\mathcal{B}$ y $\mathcal{B}\mathbb{R}\mathcal{A}$, entonces $\mathcal{A}$ y $\mathcal{B}$ son argumentos con igual preferencia relativa, notado como $\mathcal{A} \equiv \mathcal{B}$; y si no ocurre que $\mathcal{A}\mathbb{R}\mathcal{B}$ ni $\mathcal{B}\mathbb{R}\mathcal{A}$, entonces $\mathcal{A}$ y $\mathcal{B}$ son argumentos incomparables, notado como $\mathcal{A} \bowtie \mathcal{B}$. Como se mencionó anteriormente, las derrotas entre los argumentos de un AFS son obtenidas al aplicar preferencias sobre los argumentos conflictivos. Dados dos argumentos $\mathcal{A}$ y $\mathcal{B}$ tales que $(\mathcal{A}, \mathcal{B}) \in \mathbb{C}$, se considera la relación de preferencia $\mathbb{R}$. Si $\mathcal{A} \succ \mathcal{B}$ o $\mathcal{B} \succ \mathcal{A}$, entonces el argumento preferido es un *derrotador*

$(c_1)\ (\mathcal{N}, \mathcal{O}) \in \mathbb{C}_{3.20}$	$(c_3)\ (\mathcal{N}, \mathcal{O}_1) \in \mathbb{C}_{3.20}$	$(c_5)\ (\mathcal{O}, \mathcal{N}_1) \in \mathbb{C}_{3.20}$	$(c_7)\ (\mathcal{N}_1, \mathcal{O}_1) \in \mathbb{C}_{3.20}$
$(c_2)\ (\mathcal{O}, \mathcal{N}) \in \mathbb{C}_{3.20}$	$(c_4)\ (\mathcal{O}_1, \mathcal{N}) \in \mathbb{C}_{3.20}$	$(c_6)\ (\mathcal{N}_1, \mathcal{O}) \in \mathbb{C}_{3.20}$	$(c_8)\ (\mathcal{O}_1, \mathcal{N}_1) \in \mathbb{C}_{3.20}$

Figura 3.30: Conflictos entre los argumentos $\mathcal{N}$ y $\mathcal{O}$, y sus super-argumentos $\mathcal{N}_1$ y $\mathcal{O}_1$, correspondientes al Ejemplo 3.20.

propio del otro; y si $\mathcal{A} \bowtie \mathcal{B}$ o $\mathcal{A} \equiv \mathcal{B}$, entonces $\mathcal{A}$ y $\mathcal{B}$ son derrotadores por bloqueo. Nótese que si $\mathbb{R} = \emptyset$, entonces para todo par de argumentos $\mathcal{A}, \mathcal{B} \in \mathbb{A}$ se tiene que $\mathcal{A} \bowtie \mathcal{B}$.

A continuación discutiremos las similitudes existentes entre la relación de sub-argumento de [MGS06] y la relación de soporte por necesidad presentada en la Sección 3.3. En la literatura se han establecido diferentes restricciones que una relación de sub-argumento debería satisfacer. En particular, el *principio de composicionalidad* (en inglés, *compositionality principle*) [Vre97] captura la intuición de que un argumento no puede estar aceptado a menos que todos sus sub-argumentos estén aceptados. Esto implica que: *i*) si un argumento está aceptado, entonces todos sus sub-argumentos también están aceptados; y *ii*) si un argumento no está aceptado, entonces todos sus super-argumentos tampoco están aceptados.

Cabe destacar que las restricciones impuestas por el principio de composicionalidad se corresponden con aquellas impuestas por la interpretación de soporte por necesidad de los AFN. Recordemos que, bajo esta interpretación, dado $\mathcal{A} \overset{n}{\Longrightarrow} \mathcal{A}_1$, si $\mathcal{A}_1$ está aceptado entonces $\mathcal{A}$ también está aceptado (correspondiente a la primera restricción del principio de composicionalidad); y si $\mathcal{A}$ no está aceptado, entonces $\mathcal{A}_1$ tampoco está aceptado (segunda restricción del principio de composicionalidad). De manera similar, la propiedad de herencia de conflicto de los AFS captura estas restricciones al propagar los conflictos entre argumentos hacia sus super-argumentos.

Dado que los AFN no consideran la existencia de preferencias entre argumentos, primero analizaremos el caso de los AFS que poseen una relación de preferencia vacía (*i. e.*, $\mathbb{R} = \emptyset$). En tal caso, si $(\mathcal{A}, \mathcal{B}) \in \mathbb{C}$, entonces $\mathcal{A}$ derrota a $\mathcal{B}$ y viceversa, notado respectivamente como $\mathcal{A} \longrightarrow \mathcal{B}$ y $\mathcal{B} \longrightarrow \mathcal{A}$. Por ejemplo, consideremos el $\mathsf{AFS}_{3.20}$ del Ejemplo 3.20 y supongamos que $\mathbb{R}_{3.20} = \emptyset$. Como no existen preferencias entre argumentos, los conflictos $(c_1) - (c_8)$, ilustrados en el Figura 3.30, dan lugar a las derrotas $(d_1) - (d_8)$ ilustradas en la Figura 3.31. Consideremos ahora un marco argumentativo con soporte por necesidad que describa esta situación. El marco resultante es $\mathsf{AFN}_{3.20}$, ilustrado en la Figura 3.32,

$$
\begin{array}{llllllll}
(d_1) & \mathcal{N} \longrightarrow \mathcal{O} & (d_3) & \mathcal{N} \longrightarrow \mathcal{O}_1 & (d_5) & \mathcal{O} \longrightarrow \mathcal{N}_1 & (d_7) & \mathcal{N}_1 \longrightarrow \mathcal{O}_1 \\
(d_2) & \mathcal{O} \longrightarrow \mathcal{N} & (d_4) & \mathcal{O}_1 \longrightarrow \mathcal{N} & (d_6) & \mathcal{N}_1 \longrightarrow \mathcal{O} & (d_8) & \mathcal{O}_1 \longrightarrow \mathcal{N}_1
\end{array}
$$

Figura 3.31: Derrotas obtenidas a partir de la relación de preferencia vacía y los conflictos identificados para el $\mathsf{AFS}_{3.20}$ del Ejemplo 3.20.

$$
\begin{tikzcd}
\mathcal{N} \arrow[r, Rightarrow, "n"] \arrow[d, shift left] \arrow[dr] & \mathcal{N}_1 \arrow[d, shift left] \arrow[dl] \\
\mathcal{O} \arrow[r, Rightarrow, "n"] \arrow[u, shift left] \arrow[ur] & \mathcal{O}_1 \arrow[u, shift left] \arrow[ul]
\end{tikzcd}
$$

Figura 3.32: Marco Argumentativo con Soporte por Necesidad $\mathsf{AFN}_{3.20}$, correspondiente al $\mathsf{AFS}_{3.20}$ del Ejemplo 3.20.

donde la relación de soporte se corresponde con la relación de sub-argumento de $\mathsf{AFS}_{3.20}$, y la relación de derrota está determinada por las derrotas $(d_1) - (d_8)$ de la Figura 3.31.

En este caso, analizaremos si las derrotas extendidas de $\mathsf{AFN}_{3.20}$ conducen a comportamiento no deseado en las inferencias del sistema. En particular, dada la especificación de $\mathsf{AFN}_{3.20}$, se obtienen las siguientes derrotas extendidas del primer tipo: $\mathcal{N} \longrightarrow \mathcal{O}_1$, $\mathcal{O} \longrightarrow \mathcal{N}_1$, $\mathcal{N}_1 \longrightarrow \mathcal{O}_1$ y $\mathcal{O}_1 \longrightarrow \mathcal{N}_1$. Análogamente, se obtienen las siguientes derrotas extendidas del segundo tipo: $\mathcal{N}_1 \longrightarrow \mathcal{O}$, $\mathcal{O}_1 \longrightarrow \mathcal{N}$, $\mathcal{N}_1 \longrightarrow \mathcal{O}_1$ y $\mathcal{O}_1 \longrightarrow \mathcal{N}_1$. Nótese que todas estas derrotas extendidas son también derrotas directas y, por lo tanto, no modifican el estado de aceptabilidad de los argumentos con respecto a la relación de derrota original de $\mathsf{AFN}_{3.20}$.

El análisis precedente concierne un subconjunto de los AFS propuestos en [MGS06]; específicamente, aquellos marcos en los que no se consideran preferencias entre argumentos. A continuación consideraremos el caso de los AFS cuya relación de preferencia no es vacía. Por ejemplo, consideremos nuevamente el marco $\mathsf{AFS}_{3.20}$ del Ejemplo 3.20 y supongamos ahora que $\mathbb{R}_{3.20} \neq \emptyset$. Luego, analizaremos cómo los conflictos $(c_1) - (c_8)$ son resueltos mediante el uso de preferencias para finalmente obtener las derrotas asociadas. En particular, estas derrotas serán un subconjunto de las derrotas $(d_1) - (d_8)$ ilustradas en la Figura 3.31. En consecuencia, se considerarán estas derrotas en el contexto de un AFN, analizando si cada una de ellas da lugar a comportamiento no deseado o contra-intuitivo.

Las derrotas extendidas de un AFN se originan a partir de la combinación de una de-

$$\begin{tikzcd} \mathcal{N} \arrow[r, Rightarrow, "n"] \arrow[d] \arrow[dr, dashed] & \mathcal{N}_1 \arrow[dl, dashed] \\ \mathcal{O} \arrow[r, Rightarrow, "n"] & \mathcal{O}_1 \end{tikzcd}$$
(d_1)

$$\begin{tikzcd} \mathcal{N} \arrow[r, Rightarrow, "n"] & \mathcal{N}_1 \\ \mathcal{O} \arrow[r, Rightarrow, "n"] \arrow[u] \arrow[ur, dashed] & \mathcal{O}_1 \arrow[ul, dashed] \end{tikzcd}$$
(d_2)

$$\begin{tikzcd} \mathcal{N} \arrow[r, Rightarrow, "n"] \arrow[dr] & \mathcal{N}_1 \arrow[d, dashed] \\ \mathcal{O} \arrow[r, Rightarrow, "n"] & \mathcal{O}_1 \end{tikzcd}$$
(d_3)

$$\begin{tikzcd} \mathcal{N} \arrow[r, Rightarrow, "n"] & \mathcal{N}_1 \\ \mathcal{O} \arrow[r, Rightarrow, "n"] & \mathcal{O}_1 \arrow[ul] \arrow[u, dashed] \end{tikzcd}$$
(d_4)

$$\begin{tikzcd} \mathcal{N} \arrow[r, Rightarrow, "n"] & \mathcal{N}_1 \\ \mathcal{O} \arrow[r, Rightarrow, "n"] \arrow[ur] & \mathcal{O}_1 \arrow[u, dashed] \end{tikzcd}$$
(d_5)

$$\begin{tikzcd} \mathcal{N} \arrow[r, Rightarrow, "n"] & \mathcal{N}_1 \arrow[dl] \arrow[d, dashed] \\ \mathcal{O} \arrow[r, Rightarrow, "n"] & \mathcal{O}_1 \end{tikzcd}$$
(d_6)

$$\begin{tikzcd} \mathcal{N} \arrow[r, Rightarrow, "n"] & \mathcal{N}_1 \arrow[d] \\ \mathcal{O} \arrow[r, Rightarrow, "n"] & \mathcal{O}_1 \end{tikzcd}$$
(d_7)

$$\begin{tikzcd} \mathcal{N} \arrow[r, Rightarrow, "n"] & \mathcal{N}_1 \\ \mathcal{O} \arrow[r, Rightarrow, "n"] & \mathcal{O}_1 \arrow[u] \end{tikzcd}$$
(d_8)

Figura 3.33: Marcos Argumentativos con Soporte por Necesidad obtenidos a partir de la relación de soporte del $\mathsf{AFS}_{3.20}$ y las derrotas $(d_1) - (d_8)$ de la Figura 3.31.

rrota directa y la relación de soporte entre argumentos. Esto permite analizar las derrotas $(d_1)-(d_8)$ de la Figura 3.31 en forma aislada, ya que las derrotas extendidas que se generen a partir de ellas no darán lugar a posteriores derrotas extendidas. La Figura 3.33 ilustra una representación gráfica del AFN resultante en cada caso, donde la relación de soporte por necesidad se corresponde con la relación de sub-argumento de $\mathsf{AFS}_{3.20}$ y la relación de derrota considera, respectivamente, cada una de las derrotas $(d_1) - (d_8)$. En cada caso, las derrotas extendidas surgidas de la combinación de las relaciones de soporte y derrota del AFN son denotadas mediante flechas punteadas. A continuación se incluye un análisis detallado de las derrotas extendidas que se obtienen en cada caso de la Figura 3.33, a partir de la consideración individual de las derrotas $(d_1) - (d_8)$.

(d_1) $\mathcal{N} \longrightarrow \mathcal{O}$.

En este caso se obtienen las derrotas extendidas $\mathcal{N} \longrightarrow \mathcal{O}_1$ y $\mathcal{N}_1 \longrightarrow \mathcal{O}$. La derrota extendida $\mathcal{N} \longrightarrow \mathcal{O}_1$ captura el siguiente comportamiento: si $\mathcal{N}$ está aceptado, entonces $\mathcal{O}_1$ no está aceptado. Esta es una derrota extendida del primer tipo y está relacionada con la segunda restricción del principio de composicionalidad ya que $\mathcal{N} \longrightarrow \mathcal{O}$ y $\mathcal{O} \overset{n}{\Longrightarrow} \mathcal{O}_1$ (que corresponde a $\mathcal{O} \sqsubseteq \mathcal{O}_1$). De esta manera, si $\mathcal{N}$ está aceptado, entonces $\mathcal{O}$ no estará aceptado y por lo tanto $\mathcal{O}_1$ tampoco estará aceptado. Por otra parte, la derrota extendida $\mathcal{N}_1 \longrightarrow \mathcal{O}$ expresa que si $\mathcal{N}_1$ está aceptado, entonces $\mathcal{O}$ no está aceptado. En este caso se trata de una derrota extendida del segundo tipo, la cual está relacionada con la primera restricción del principio de composicionalidad: como $\mathcal{N} \stackrel{n}{\Longrightarrow} \mathcal{N}_1$ (*i. e.*, $\mathcal{N} \sqsubseteq \mathcal{N}_1$), si $\mathcal{N}_1$ está aceptado, entonces $\mathcal{N}$ también está aceptado. Luego, como $\mathcal{N} \longrightarrow \mathcal{O}$, tenemos que $\mathcal{O}$ no estará aceptado.

(d_2) $\mathcal{O} \longrightarrow \mathcal{N}$.
El análisis en este caso es análogo al realizado en el caso de (d_1).

(d_3) $\mathcal{N} \longrightarrow \mathcal{O}_1$.
En este caso se obtiene únicamente la derrota extendida $\mathcal{N}_1 \longrightarrow \mathcal{O}_1$. Esta corresponde al segundo tipo de derrota extendida y expresa que si $\mathcal{N}_1$ está aceptado, entonces $\mathcal{O}_1$ no está aceptado. Esta derrota extendida se encuentra asociada a la primera restricción del principio de composicionalidad: dado que $\mathcal{N} \stackrel{n}{\Longrightarrow} \mathcal{N}_1$ (*i. e.*, $\mathcal{N} \sqsubseteq \mathcal{N}_1$), si $\mathcal{N}_1$ está aceptado entonces $\mathcal{N}$ también está aceptado y, por lo tanto, al tener $\mathcal{N} \longrightarrow \mathcal{O}_1$, el argumento $\mathcal{O}_1$ no estará aceptado.

(d_4) $\mathcal{O}_1 \longrightarrow \mathcal{N}$.
La única derrota extendida obtenida en este caso es $\mathcal{O}_1 \longrightarrow \mathcal{N}$. Esta corresponde al primer tipo de derrota extendida y se encuentra relacionada con la segunda restricción del principio de composicionalidad. Dado que $\mathcal{O}_1 \longrightarrow \mathcal{N}$, si $\mathcal{O}_1$ está aceptado, entonces $\mathcal{N}$ no estará aceptado. Luego, como $\mathcal{N} \stackrel{n}{\Longrightarrow} \mathcal{N}_1$ (*i. e.*, $\mathcal{N} \sqsubseteq \mathcal{N}_1$), el argumento $\mathcal{N}_1$ tampoco estará aceptado.

(d_5) $\mathcal{O} \longrightarrow \mathcal{N}_1$.
El análisis en este caso es análogo al realizado en el caso de (d_3).

(d_6) $\mathcal{N}_1 \longrightarrow \mathcal{O}$.
El análisis en este caso es análogo al realizado en el caso de (d_4).

(d_7) $\mathcal{N}_1 \longrightarrow \mathcal{O}_1$.
En este caso no se obtienen derrotas extendidas.

(d_8) $\mathcal{O}_1 \longrightarrow \mathcal{N}_1$.
En este caso no se obtienen derrotas extendidas.

Cabe destacar que, si bien este análisis fue realizado para el caso particular del marco argumentativo abstracto con sub-argumentos $\mathsf{AFS}_{3.20}$ y su AFN asociado, los resultados obtenidos pueden generalizarse para cualquier AFS. Por una parte, como la relación de sub-argumento de un AFS es transitiva, la relación $\mathcal{N} \sqsubseteq \mathcal{N}_1$ (respectivamente, $\mathcal{O} \sqsubseteq \mathcal{O}_1$) puede haber sido obtenida por transitividad sobre otros pares de la relación $\sqsubseteq$. En particular, esto se vería reflejado en la relación de soporte por necesidad del AFN asociado mediante una cadena de soporte de $\mathcal{N}$ hacia $\mathcal{N}_1$ (respectivamente de $\mathcal{O}$ hacia $\mathcal{O}_1$). Por otra parte, como se mencionó anteriormente, las derrotas extendidas de un AFN no pueden ser utilizadas para obtener nuevas derrotas extendidas. Por lo tanto, el análisis por casos realizado para $\mathsf{AFS}_{3.20}$ y $\mathsf{AFN}_{3.20}$ resulta válido para cualquier AFS y su correspondiente AFN asociado.

Finalmente, el análisis efectuado muestra que existe una relación cercana entre la noción de sub-argumento de [MGS06] y la noción de soporte por necesidad de los AFN presentados en la Sección 3.3. En particular, la relación de sub-argumento sugiere que los sub-argumentos son *necesarios* para sus super-argumentos. Además, se observó que las restricciones impuestas por la relación de soporte por necesidad se corresponden con aquellas impuestas por el principio de composicionalidad. De esta manera, es posible establecer una correspondencia entre la relación de sub-argumento de [MGS06] y la relación de soporte por necesidad de los AFN.

3.5.2. Marcos Dialécticos Abstractos (ADF)

En esta sección analizaremos otra propuesta existente en la literatura que permite modelar la noción de soporte entre argumentos. Concretamente, consideraremos los *Marcos Dialécticos Abstractos (ADF)* (por su nombre en inglés, *Abstract Dialectical Frameworks*) propuestos en [BW10], los cuales proveen una generalización de los marcos argumentativos de [Dun95].

En [BW10] los autores introducen los ADF con el objetivo de añadir estándares de prueba a los marcos argumentativos de Dung. Un ADF es un grafo dirigido cuyos nodos representan argumentos, los cuales pueden estar aceptados o no, y los arcos entre los nodos representan dependencias. A cada argumento $\mathcal{X}$ en el grafo se le asocia una *condición de aceptación*, denotada $C(\mathcal{X})$. Básicamente, una condición de aceptación es una función proposicional cuyo valor de verdad está determinado por el valor correspondiente a las condiciones de aceptación de los argumentos $\mathcal{Y}$ tales que $(\mathcal{Y}, \mathcal{X})$ es un arco en el ADF (*i. e.*, el argumento $\mathcal{Y}$ es un "padre" de $\mathcal{X}$). De esta manera, la influencia que un nodo puede tener sobre otro se encuentra enteramente especificada a través de las condiciones de aceptación.

A diferencia de los AF propuestos por Dung, donde los arcos representan una relación particular entre argumentos (derrota), y los argumentos se encuentran aceptados a menos que estén derrotados y no sean defendidos, los ADF admiten la representación de diferentes dependencias entre argumentos. Por ejemplo, puede haber nodos que estén rechazados a menos que estén soportados por algunos nodos aceptados (dando lugar a vínculos de soporte). Además, pueden existir vínculos con diferente fuerza asociada, e incluso vínculos que representen el soporte o la derrota de un nodo dependiendo del contexto. De esta manera, la especificación de condiciones de aceptación individuales para los argumentos enriquece la capacidad expresiva de los ADF. En particular, en [BW10] los autores destacan que es posible representar un AF de Dung utilizando un ADF. Para lograr esto, debe realizarse lo siguiente: por cada par de argumentos $\mathcal{X}$, $\mathcal{Y}$ tales que $\mathcal{Y}$ derrota a $\mathcal{X}$ en el AF, se incluye un vínculo $(\mathcal{Y}, \mathcal{X})$ en el ADF; luego, para cada argumento $\mathcal{X}$ involucrado en estos vínculos, se define la condición de aceptación $C(\mathcal{X}) = \neg\mathcal{Y}$ que expresa que el argumento $\mathcal{X}$ estará aceptado si ninguno de sus padres está aceptado.

Como es mencionado por los autores en [BW10], un aspecto importante de su trabajo es la generalización de las semánticas estándar de Dung. En particular, afirman que la semántica grounded puede generalizarse para ADF arbitrarios. Por otra parte, enuncian que para las semánticas estable y preferida es necesario considerar nociones de soporte y derrota, las cuales se encuentran presentes en una sub-clase de los ADF llamada *Marcos Dialécticos Abstractos Bipolares (BADF)* (en inglés, *Bipolar Abstract Dialectical Frameworks*). Luego, los autores adaptan técnicas de [GL88] para evitar ciclos de soporte.

Con respecto a las aproximaciones presentadas en las secciones anteriores, en [BW10] los autores declaran que su propuesta es más general ya que, en lugar de añadir un segundo tipo de relación o vínculo entre argumentos, ellos permiten capturar una gran variedad de tipos de nodos y relaciones entre ellos. En particular, destacan que las condiciones de aceptación definidas en los ADF son más flexibles que las restricciones impuestas por las diferentes interpretaciones de soporte tales como soporte por necesidad, soporte deductivo o soporte evidencial. Por ejemplo, si $\mathcal{C}$ depende de $\mathcal{A}$ y $\mathcal{B}$, es posible modelar la siguiente restricción en un ADF: la aceptación de $\mathcal{B}$ y la no aceptación de $\mathcal{A}$ implican la aceptación de $\mathcal{C}$.

A pesar de la flexibilidad asociada al uso de condiciones de aceptación, dado que el estado de aceptación de un nodo en un ADF depende únicamente del estado de aceptación de sus padres, los ADF pueden no ser capaces de capturar las restricciones de aceptabi-

$$\begin{tikzcd} \mathcal{C} & \mathcal{LL} \arrow[l] & \mathcal{N} \arrow[l, Rightarrow, "s"'] \\ & \mathcal{V} \arrow[u] & \end{tikzcd}$$

Figura 3.34: Marco Argumentativo Bipolar $\mathsf{BAF}_{3.21}$ del Ejemplo 3.21.

lidad impuestas por algunas interpretaciones de soporte particulares. Para ilustrar esto, consideremos la interpretación de soporte deductivo. Si $\mathcal{A} \overset{d}{\Longrightarrow} \mathcal{B}$ y $\mathcal{C} \longrightarrow \mathcal{B}$, entonces existe la derrota mediada $\mathcal{C} \longrightarrow \mathcal{A}$. Esta derrota mediada se encuentra asociada a la restricción del soporte deductivo que expresa que si $\mathcal{B}$ no está aceptado, entonces $\mathcal{A}$ tampoco debe estar aceptado. Un ADF que modela esta situación poseería los nodos $\mathcal{A}$, $\mathcal{B}$ y $\mathcal{C}$, y los vínculos $(\mathcal{A}, \mathcal{B})$ (soporte) y $(\mathcal{C}, \mathcal{B})$ (ataque). Sin embargo, a partir de este ADF no sería posible definir una condición de aceptación capaz de asegurar que $\mathcal{A}$ no esté aceptado cuando $\mathcal{C}$ esté aceptado, dado que no existe vínculo entre $\mathcal{A}$ y $\mathcal{C}$ en el grafo. Una situación similar ocurre al considerar la relación de soporte de los BAF presentados en la Sección 3.1. Sin embargo, en este caso, en [BW10] los autores alegan que se debe a la diferencia entre la concepción de la noción de conflicto en los BAF y en los ADF. Para ilustrar esto, consideremos el siguiente ejemplo inspirado en [BW10].

Ejemplo 3.21 *Supongamos que una persona planea ir a correr por la tarde* ($\mathcal{C}$)*. El cielo está cubierto por nubes* ($\mathcal{N}$)*, lo cual es un indicio de posible lluvia* ($\mathcal{LL}$)*. Sin embargo, un reporte meteorológico confiable indica que la presencia de fuertes vientos* ($\mathcal{V}$) *alejará las nubes, negando la posibilidad de lluvias en la tarde. En este contexto, tenemos que* $\mathcal{N}$ *soporta a* $\mathcal{LL}$*,* $\mathcal{LL}$ *derrota a* $\mathcal{C}$*, y* $\mathcal{V}$ *derrota a* $\mathcal{LL}$*. Este escenario puede ser representado por el marco argumentativo bipolar* $\mathsf{BAF}_{3.21}$ *ilustrado en la Figura 3.34.*

Utilizando condiciones de aceptación apropiadas, y asumiendo que la derrota de $\mathcal{V}$ *sobre* $\mathcal{LL}$ *es más fuerte que el soporte provisto por* $\mathcal{N}$*, la extensión preferida del* ADF *correspondiente sería* $\{\mathcal{N}, \mathcal{V}, \mathcal{C}\}$*. Por otra parte, al considerar el* $\mathsf{BAF}_{3.21}$*,* $\{\mathcal{N}, \mathcal{V}, \mathcal{C}\}$ *no es un conjunto libre de conflictos*$^{+}$ *dado que existe una derrota soportada de* $\mathcal{N}$ *sobre* $\mathcal{C}$*. Por este motivo, la extensión preferida de* $\mathsf{BAF}_{3.21}$ *es* $\{\mathcal{N}, \mathcal{V}\}$*.*

Dados los resultados obtenidos en el Ejemplo 3.21, en [BW10] los autores expresan que el comportamiento de los BAF es contra-intuitivo. En particular, enuncian que el resultado esperado es que el argumento C pertenezca a la extensión del BAF ya que, ante la presencia de fuertes vientos, no habría nubes que hagan llover, apoyando así el plan de salir a correr.

La principal diferencia entre los ADF y los BAF es que, como se mencionó anteriormente, los ADF imponen restricciones de aceptabilidad mediante la definición de condiciones de aceptación para los nodos del grafo, las cuales únicamente pueden hacer referencia a los padres de los nodos. Por lo tanto, los ADF no consideran los conflictos entre argumentos que no se encuentran directamente relacionados en el grafo. En contraste, los BAF definen una serie de derrotas complejas que relacionan argumentos que no estaban directamente relacionados por las relaciones de derrota y soporte. Cabe destacar que esta diferencia también ocurre al considerar otras interpretaciones de soporte, tales como soporte deductivo (Sección 3.2) o soporte por necesidad (Sección 3.3). Esto se debe a que dichas aproximaciones también contemplan la definición de derrotas complejas, resultantes de la combinación de soportes y derrotas previamente existentes. Más aún, como se mostró anteriormente, el comportamiento modelado por las derrotas mediadas no puede ser capturado por los ADF. Análogamente, esta deficiencia también ocurre al considerar las derrotas extendidas de los AFN.

Finalmente, en [BW10] los autores expresan que para modelar la noción de *libertad de conflictos*$^{+}$ propuesta por los BAF (la cual considera la existencia de las derrotas soportadas y las derrotas secundarias) es necesario realizar las siguientes modificaciones en los vínculos del ADF: *i*) por cada derrota de $\mathcal{B}$ sobre $\mathcal{C}$ en el BAF tal que $\mathcal{B}$ está directa o indirectamente soportado por $\mathcal{A}$, se añade un vínculo entre $\mathcal{A}$ y $\mathcal{C}$ en el ADF (derrota soportada); y *ii*) dado un argumento $\mathcal{C}$ que está directa o indirectamente soportado por un argumento $\mathcal{B}$ tal que existe una derrota de $\mathcal{A}$ sobre $\mathcal{B}$ en el BAF, se añade un vínculo entre $\mathcal{A}$ y $\mathcal{C}$ en el ADF (derrota secundaria). Sin embargo, los autores afirman que, como sugiere la situación descripta en el Ejemplo 3.21, la consideración de estas derrotas adicionales puede conducir a resultados no deseados.

3.6. Conclusiones

En este capítulo se analizaron diferentes propuestas de la literatura que contemplan la existencia de una relación de soporte entre argumentos. En particular, el BAF presentado en la Sección 3.1 adopta una postura más general, simplemente considerando al soporte como una interacción positiva entre argumentos. En contraste, los d-BAF, AFN y EAS (respectivamente presentados en las secciones 3.2, 3.3 y 3.4) proponen la consideración de interpretaciones particulares para la noción de soporte: respectivamente, *soporte deductivo*, *soporte por necesidad* y *soporte evidencial*. Por otra parte, en la Sección 3.5 se consideraron aproximaciones que, si bien no contemplan la existencia de una relación de soporte de manera explícita, proveen mecanismos que permiten expresar el soporte entre argumentos. Se realizó un estudio comparativo entre las diferentes propuestas, destacando similitudes y diferencias entre las interpretaciones de soporte analizadas.

Se mostró que cada interpretación de soporte establece una serie de restricciones sobre la aceptabilidad de los argumentos que relaciona. En la mayoría de los casos, estas restricciones conducen a la consideración de nuevas derrotas entre argumentos, llamadas *derrotas complejas*, las cuales se infieren a partir de la combinación de derrotas ya existentes y la relación de soporte entre argumentos. En particular, se observó que estas derrotas complejas cobran sentido dependiendo de la interpretación adoptada para la relación de soporte, dado que refuerzan las restricciones de aceptabilidad impuestas por cada una de estas interpretaciones. Por lo tanto, dada una interpretación particular de soporte, es necesario considerar el conjunto de derrotas complejas correspondiente a la misma. De esta manera, no sería posible proveer un mecanismo formal que compute estas derrotas sin considerar la interpretación adoptada ya que, si estas derrotas fueran consideradas en forma arbitraria, las inferencias del sistema podrían conducir a resultados incorrectos. Las figuras 3.35 y 3.36 resumen diferentes aspectos correspondientes a las aproximaciones presentadas en las secciones 3.1, 3.2, 3.3 y 3.4, tales como la interpretación de soporte adoptada, las restricciones de aceptabilidad impuestas por la relación de soporte, y las derrotas complejas definidas por el formalismo.

A partir del análisis de las similitudes y diferencias entre estas interpretaciones, podemos concluir que algunas de ellas se encuentran estrechamente relacionadas. En particular, se mostró que las interpretaciones de soporte deductivo y soporte por necesidad presentadas en las secciones 3.2 y 3.3 son duales en el siguiente sentido: un argumento $\mathcal{A}$ soporta deductivamente a un argumento $\mathcal{B}$ si y sólo si $\mathcal{B}$ es necesario para $\mathcal{A}$. De manera similar, se mostró que existe una correspondencia directa entre la relación de sub-argumento presentada en la Sección 3.5.1 y la relación de soporte por necesidad. Por otra parte, se mostró que existe una característica esencial que diferencia al soporte evidencial presentado en la Sección 3.4 de las otras aproximaciones analizadas: un argumento aceptable de un sistema argumentativo evidencial debe estar soportado por evidencia. En consecuencia,

Formalismo	**Interpretación de Soporte**	**Restricciones de Aceptabilidad**	**Derrotas Complejas** ($\dashrightarrow$)	**Resultado ante Relación de Derrota Vacía**
Marco Argumentativo Bipolar (BAF)	Relación de Soporte General: $\overset{s}{\Longrightarrow}$	No especificadas.	$\mathcal{A} \overset{s}{\Longrightarrow} \mathcal{B} \longrightarrow \mathcal{C}$ $\mathcal{A} \longrightarrow \mathcal{B} \overset{s}{\Longrightarrow} \mathcal{C}$	Todos los argumentos del sistema están aceptados.
Marco Argumentativo Bipolar con Soporte Deductivo (d-BAF)	Relación de Soporte Deductivo: $\overset{d}{\Longrightarrow}$	Dado $\mathcal{A} \overset{d}{\Longrightarrow} \mathcal{B}$: - Si $\mathcal{A}$ está aceptado, entonces $\mathcal{B}$ está aceptado. - Si $\mathcal{B}$ no está aceptado, entonces $\mathcal{A}$ no está aceptado.	$\mathcal{A} \overset{d}{\Longrightarrow} \mathcal{B} \longrightarrow \mathcal{C}$ $\mathcal{A} \longrightarrow \mathcal{B} \overset{d}{\Longleftarrow} \mathcal{C}$	Todos los argumentos del sistema están aceptados.

Figura 3.35: Características de los BAF y d-BAF.

Formalismo	**Interpretación de Soporte**	**Restricciones de Aceptabilidad**	**Derrotas Complejas** ($\dashrightarrow$)	**Resultado ante Relación de Derrota Vacía**
Marco Argumentativo con Soporte por Necesidad (AFN)	Relación de Soporte por Necesidad: $\stackrel{n}{\Longrightarrow}$	Dado $\mathcal{A} \stackrel{n}{\Longrightarrow} \mathcal{B}$: - Si $\mathcal{B}$ está aceptado, entonces $\mathcal{A}$ está aceptado. - Si $\mathcal{A}$ no está aceptado, entonces $\mathcal{B}$ no está aceptado.	$\begin{tikzcd} \mathcal{A} \arrow[r] \arrow[rr, dashed, bend left] & \mathcal{B} \arrow[r, Rightarrow, "n"] & \mathcal{C} \end{tikzcd}$ $\begin{tikzcd} \mathcal{A} \arrow[rr, dashed, bend left] & \mathcal{B} \arrow[l, Rightarrow, "n"'] \arrow[r] & \mathcal{C} \end{tikzcd}$	Todos los argumentos del sistema están aceptados.
Sistema Argumentativo Evidencial (EAS)	Relación de Soporte Evidencial: $\stackrel{e}{\Longrightarrow}$	No especificadas.	No especificadas.	Se aceptan únicamente aquellos argumentos que estén soportados por evidencia (e-soportados).

Figura 3.36: Características de los AFN y EAS.

podría ocurrir que un argumento de un EAS que no es objeto de derrotas no pertenezca al conjunto de argumentos aceptados por no encontrarse soportado por evidencia.

Podemos concluir que el estudio de la noción de soporte constituye un prometedor tópico de investigación para la comunidad de Argumentación dentro del área de Inteligencia Artificial. Gran parte de los trabajos en la literatura de argumentación se concentran en el estudio de marcos argumentativos abstractos siguiendo el estilo de [Dun95], donde únicamente se consideran derrotas entre argumentos. No obstante esto, la incorporación de un relación de soporte ha demostrado ser de gran utilidad, aumentando así las capacidades de representación de los sistemas argumentativos.

Finalmente, podemos observar que la mayoría de los trabajos que abordan la existencia de soporte entre argumentos lo hacen en un contexto de argumentación abstracta. Como se mencionó en la Sección 3.3, la falta de guías y herramientas para la representación de conocimiento asociadas del uso de sistemas argumentativos abstractos puede conducir a problemas derivados de un incorrecto modelado de la información, lo cual puede ser resuelto al utilizar sistemas más concretos. Por otra parte, como se verá en el Capítulo 5, recientemente han surgido aproximaciones alternativas que estudian la noción de soporte en entornos más concretos. En particular, además de las interpretaciones de soporte analizadas en este capítulo, existen otras que restan ser estudiadas. Tal es el caso de la interpretación de soporte derivada del modelo para el diseño de argumentos propuesto por Toulmin en [Tou58], la cual será considerada por los formalismos desarrollados en los capítulos 4 y 5.

Capítulo 4

Marcos Argumentativos con Backing y Undercutting (BUAF)

En este capítulo se propone un formalismo llamado *Marco Argumentativo con Backing y Undercutting (BUAF)* (por su nombre en inglés, *Backing-Undercutting Argumentation Framework*) que extiende a los marcos argumentativos de [Dun95] mediante la incorporación de una relación especializada de soporte, una distinción entre diferentes tipos de ataque y una relación de preferencia entre argumentos. De esta manera, el BUAF provee un marco unificado que permite modelar ataque y soporte a inferencias en un contexto de argumentación abstracta. Dada su formalización como un marco argumentativo abstracto, el BUAF permitirá resolver los conflictos entre los argumentos mediante el uso de preferencias, para finalmente evaluar su aceptabilidad y así obtener los argumentos aceptados del sistema.

En el Capítulo 3 se analizaron diferentes propuestas de la literatura que contemplan la existencia de una relación de soporte entre argumentos. En particular, cada una de estas alternativas se encuentra asociada a una interpretación específica para la relación de soporte. En este capítulo se desarrollará otra propuesta que contempla una forma alternativa de soporte entre argumentos. Concretamente, se desarrollará un formalismo de argumentación abstracta que permite combinar dos nociones importantes en el área de argumentación. Por una parte, el sistema propuesto permitirá modelar el soporte existente entre las nociones de backing y warrant propuestas por Toulmin [Tou58]. Por otra parte, la propuesta desarrollada en este capítulo permitirá representar la noción de undercutting defeater de Pollock [Pol87].

4.1. Introducción

Los modelos argumentativos usualmente consideran a los argumentos como una pieza de razonamiento que provee una conexión entre premisas y una conclusión. No obstante, en [Tou58] Toulmin promulgó la idea de que los argumentos deberían ser analizados utilizando un formato más rico que el enfoque tradicional empleado en la lógica formal. Mientras que los análisis bajo la lógica utilizan la dicotomía de premisas y conclusión, Toulmin propuso un modelo para el diseño de argumentos que, además de datos y conclusión, distingue cuatro elementos: warrant, backing, calificador y derrotador (ver Sección 2.3).

Como se mencionó en la Sección 2.3, es posible identificar dos tipos de interacciones entre los elementos del esquema propuesto por Toulmin. En primer lugar, además de los datos que soportan la conclusión, el backing provee soporte para el warrant. En particular, el warrant constituye una regla de inferencia que provee la conexión entre las premisas y la conclusión. Luego, dado que el backing establece las razones bajo las cuales el warrant asociado es válido, podemos considerar que el backing provee soporte para el correspondiente warrant. Para ilustrar esto, consideremos el ejemplo introducido en previamente la Sección 2.3, ilustrado en la Figura 4.1, que versa acerca de la nacionalidad de un sujeto llamado Harry.

Dada la situación ilustrada en la Figura 4.1, el warrant constituye la regla general que establece que un hombre nacido en las Bermudas generalmente es Británico, permitiendo inferir, presumiblemente, que Harry es Británico por haber nacido en las Bermudas. Esto ocurre dada la existencia de estatutos y disposiciones legales que así lo establecen, correspondiendo al soporte brindado por el backing. Por otra parte, en [Tou58] Toulmin no proveyó mayores detalles acerca de la naturaleza de los derrotadores, considerándolos simplemente como condiciones de excepción para el argumento. De esta manera, sin pérdida de la generalidad, podemos asociar la noción de derrotador propuesta por Toulmin con la noción de derrotador para un argumento utilizada en la literatura de argumentación [PV02]. Luego, por ejemplo, el hecho de que los padres de Harry son extranjeros provee una condición de excepción para el argumento, dando lugar a una derrota sobre el mismo.

En cuanto a la naturaleza de los derrotadores, una contribución importante al área de argumentación fue realizada por Pollock en [Pol87]. En su trabajo, Pollock enunció que las razones rebatibles (las cuales pueden ser combinadas para obtener argumentos) poseen

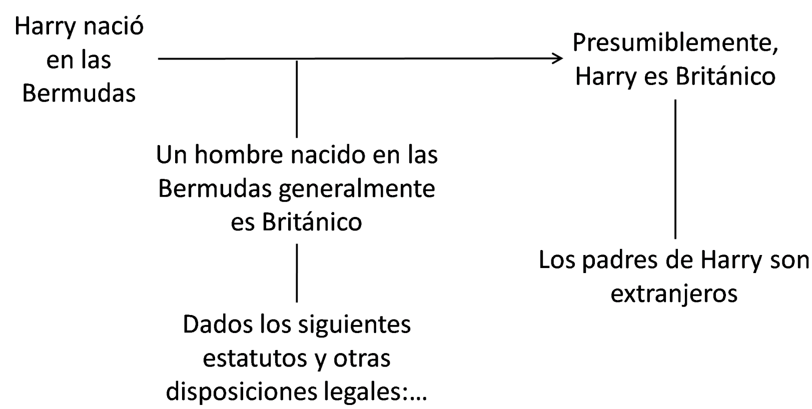


Figura 4.1: Ejemplo propuesto por Toulmin en [Tou58].

derrotadores, e identificó dos tipos de derrotadores: *rebutting defeaters* y *undercutting defeaters*. Mientras que los derrotadores del primer tipo derrotan la conclusión de un argumento proveyendo un argumento para la conclusión contraria, los del segundo tipo derrotan la conexión existente entre las premisas y la conclusión.

Para el caso del ejemplo ilustrado en la Figura 4.1, podemos considerar que el hecho de los padres de Harry son extranjeros constituye un undercutting defeater para el argumento. Esto se debe a que la información que establece que los padres de Harry son extranjeros no contradice la conclusión de que Harry es Británico, sino que constituye una razón en contra de la regla de inferencia expresada por el warrant. Es decir, el hecho de que los padres de Harry son extranjeros representa, de alguna manera, una excepción a la regla que establece que un hombre nacido en las Bermudas generalmente es Británico.

Tomando en cuenta la caracterización de derrotadores propuesta por Pollock y el modelo para el diseño de argumentos propuesto por Toulmin, podemos considerar que un undercutting defeater provee razones en contra del warrant asociado a un argumento. En consecuencia, podemos considerar que los backings brindan soporte a los warrants con el objetivo de defenderlos ante ataques de tipo undercutting, para así prevenir la ocurrencia de derrotas de este tipo.

A continuación se presentará un formalismo argumentativo que combina las ideas de Toulmin y Pollock en un contexto de argumentación abstracta. De esta manera, el formalismo propuesto proveerá un mecanismo para representar ataque y soporte para inferencias, capturando así la noción de undercutting defeater propuesta por Pollock y el soporte provisto por los backings de acuerdo al modelo de Toulmin.

El resto de este capítulo se encuentra organizado de la siguiente manera: en la Sección 4.2 se presentará la especificación del *Marco Argumentativo con Backing y Undercutting (BUAF)*, el cual extiende a los marcos argumentativos de Dung mediante la incorporación de ataque y soporte para inferencias, y una relación de preferencias para resolver los conflictos entre argumentos. En la Sección 4.3 se introducirán los diferentes tipos de derrota que se obtienen en un BUAF mediante la aplicación de preferencias sobre argumentos conflictivos, así como también aquellos resultantes de la consideración conjunta de argumentos de ataque y soporte para inferencias. La Sección 4.4 presentará diferentes nociones semánticas que proveen las bases para la definición formal de las semánticas de aceptabilidad de los BUAF. Además, se identificarán algunas propiedades deseables para los conjuntos de argumentos aceptables de los BUAF. Finalmente, en las secciones 4.5 y 4.6 se relacionará la propuesta desarrollada en este capítulo con otras aproximaciones existentes en la literatura y se elaborarán conclusiones pertinentes.

4.2. Especificación de los Marcos Argumentativos con Backing y Undercutting

Como se vio en la Sección 2.2.1, un AF se encuentra caracterizado por un conjunto de argumentos y una relación de derrota entre ellos. En esta sección se brindará una especificación para los *Marcos Argumentativos con Backing y Undercutting (BUAF)*, los cuales extienden a los AF de Dung mediante la incorporación de una relación especializada de soporte y una relación de preferencia, así como también distinguiendo entre diferentes tipos de ataque entre argumentos. Los resultados del estudio y desarrollo de este formalismo han sido publicados en [CGS11a], [CGS12b] y recientemente en [CGGS13b].

La relación de soporte de un BUAF permite representar el soporte que los backings de Toulmin le brindan a sus warrants asociados. Retomando el ejemplo de Toulmin acerca de la nacionalidad de Harry, el BUAF permitirá modelar el soporte que la existencia de estatutos y otras disposiciones legales proveen para la regla que establece que un hombre nacido en las Bermudas generalmente es Británico. Por otra parte, dado que los argumentos pueden ser combinados para obtener argumentos más complejos, esto da lugar a un anidamiento de argumentos de acuerdo al modelo propuesto por Toulmin. De esta manera, la conclusión de un argumento puede ser una premisa de otro argumento, dando lugar así a la existencia de sub-argumentos.

Como se introdujo en la Sección 2.3 los warrants constituyen una regla de inferencia que provee la conexión entre las premisas (datos) y conclusión de un argumento. Es por esto que, sin pérdida de la generalidad, podemos considerar que cada argumento posee un único warrant asociado, el cual provee el paso de inferencia de las premisas a la conclusión. De esta manera, los warrants correspondientes a pasos de inferencia intermedios en la construcción de un argumento estarán asociados a los sub-argumentos correspondientes, quedando indirectamente vinculados con el argumento final mediante una relación de sub-argumento. Esto ilustra la necesidad de considerar la existencia de sub-argumentos dentro de un BUAF. En consecuencia, la relación de soporte de un BUAF estará definida de manera tal que permita modelar, además del soporte entre backings y warrants, la relación de soporte existente entre sub-argumentos y super-argumentos. En particular, esta cualidad de la relación de soporte distingue a los BUAF de otras formalizaciones existentes en la literatura, las cuales adoptan una única interpretación para la relación de soporte.

Por otra parte, la relación de ataque de un BUAF distinguirá entre tres tipos de ataque: rebutting, undercutting y undermining. Los primeros dos tipos de ataque se encuentran relacionados con los derrotadores de tipo rebutting y undercutting propuestos por Pollock [Pol87]. Luego, el ataque de tipo undermining se encuentra relacionado con la noción de undermining defeater, ampliamente utilizada en al literatura de argumentación. Recordemos que, como se mencionó en la Sección 2.1, un ataque de tipo undermining constituye un ataque sobre una premisa de un argumento. Para ilustrar esta noción consideremos el ejemplo de Toulmin acerca de la nacionalidad de Harry. En tal caso, la existencia de un argumento que enuncia que Harry nació en París daría origen a un ataque de tipo undermining, dado que provee una razón en contra de la premisa que establece que Harry nació en las Bermudas. Finalmente, a partir de la relación de ataque entre argumentos, el BUAF utilizará una relación de preferencia para resolver los conflictos entre argumentos, dando lugar así a un conjunto de derrotas entre ellos.

Definición 4.1 (Marco Argumentativo con Backing y Undercutting) *Un* Marco Argumentativo con Backing y Undercutting (BUAF) *es una tupla* $\langle \mathbb{A}, \mathbb{R}, \mathbb{S}, \preceq \rangle$*, donde:*

- $\mathbb{A}$ *es un conjunto de argumentos;*
- $\mathbb{R} \subseteq \mathbb{A} \times \mathbb{A}$ *es una relación de ataque entre argumentos, distinguiendo entre ataques de tipo rebutting* $\mathbb{R}_b$*, undercutting* $\mathbb{U}_c$ *y undermining* $\mathbb{U}_m$ *(i. e.,* $\mathbb{R} = \mathbb{R}_b \cup \mathbb{U}_c \cup \mathbb{U}_m$*);*

- $\mathbb{S} \subseteq \mathbb{A} \times \mathbb{A}$ *es una relación de soporte entre argumentos, distinguiendo entre el soporte brindado por backings* $\mathbb{B}_k$ *y sub-argumentos* $\mathbb{S}_{\sqsubseteq}$ *(i. e.,* $\mathbb{S} = \mathbb{B}_k \cup \mathbb{S}_{\sqsubseteq}$*);*
- $\preceq \subseteq \mathbb{A} \times \mathbb{A}$ *es una relación de preferencia entre argumentos; y*
- *las relaciones de ataque y soporte son disjuntas (i. e.,* $\mathbb{R} \cap \mathbb{S} = \emptyset$*).*

Las relaciones de ataque y soporte de un BUAF deberán ser disjuntas para prevenir la ocurrencia de situaciones inconsistentes en el marco argumentativo. Es decir, si existe $\mathcal{A} \in \mathbb{A}$ tal que $(\mathcal{A}, \mathcal{B}) \in \mathbb{R}$ y $(\mathcal{A}, \mathcal{B}) \in \mathbb{S}$, entonces el argumento $\mathcal{A}$ estaría simultáneamente atacando y soportando a $\mathcal{B}$, lo cual es claramente un comportamiento no deseado. Como se mencionó anteriormente, distinguiremos tres tipos de ataque dentro de la relación $\mathbb{R}$ de un BUAF: los ataques de tipo rebutting, undercutting y undermining, respectivamente denotados como $\mathbb{R}_b$, $\mathbb{U}_c$, y $\mathbb{U}_m$. Consideraremos que un argumento no puede atacar a otro en más de una forma dentro de un BUAF. Por lo tanto, los subconjuntos de la relación de ataque $\mathbb{R}$ deberán ser disjuntos de a pares.

Por otra parte, la relación de soporte $\mathbb{S}$ de un BUAF se define de formal tal que distingue entre el soporte que los backings proveen a sus warrants asociados, representado mediante el subconjunto $\mathbb{B}_k$, y el soporte que los sub-argumentos brindan a sus super-argumentos, expresado mediante el subconjunto $\mathbb{S}_{\sqsubseteq}$. En particular, siguiendo la convención usualmente utilizada en la literatura, consideraremos que todo argumento es sub-argumento de sí mismo. Luego, aquellos argumentos que son sub-argumentos de otros argumentos serán llamados *sub-argumentos propios*. De esta manera, dado que nuestro interés radica en el uso de sub-argumentos para poder modelar el anidamiento de acuerdo al modelo de Toulmin, la relación $\mathbb{S}_{\sqsubseteq}$ de los BUAF corresponderá a la relación de *sub-argumento propio*.

Por último, la relación de preferencia $\preceq$ de un BUAF será utilizada para determinar el éxito de los ataques y así obtener las derrotas entre argumentos. A partir de esto, dados dos argumentos $\mathcal{A}$ y $\mathcal{B}$ vinculados mediante la relación de preferencia (*i. e.*, $(\mathcal{A}, \mathcal{B}) \in \ \preceq$), se tendrá que el argumento $\mathcal{B}$ es al menos tan preferido como el argumento $\mathcal{A}$, notado $\mathcal{A} \preceq \mathcal{B}$. Como es usual en la literatura, $\mathcal{A} \prec \mathcal{B}$ si $\mathcal{A} \preceq \mathcal{B}$ y $\mathcal{B} \not\preceq \mathcal{A}$, denotando que $\mathcal{B}$ es estrictamente más preferido que $\mathcal{A}$. De manera similar, diremos que $\mathcal{A} \equiv \mathcal{B}$ si $\mathcal{A} \preceq \mathcal{B}$ y $\mathcal{B} \preceq \mathcal{B}$, denotando que $\mathcal{A}$ y $\mathcal{B}$ son equivalentes. Finalmente, escribiremos $\mathcal{A} \bowtie \mathcal{B}$ cuando $\mathcal{A} \not\preceq \mathcal{B}$ y $\mathcal{B} \not\preceq \mathcal{A}$, denotando que $\mathcal{A}$ y $\mathcal{B}$ son incomparables de acuerdo a la relación de preferencia del BUAF.

Un BUAF puede representarse gráficamente a través de un multi-grafo dirigido, donde los nodos son argumentos y existen diferentes tipos de arcos denotando las relaciones de ataque y soporte. El ataque de un argumento $\mathcal{A}$ sobre un argumento $\mathcal{B}$ (*i. e.*, $(\mathcal{A}, \mathcal{B}) \in \mathbb{R}$) será notado como $\mathcal{A} \rightarrowtail \mathcal{B}$. Por otra parte, continuando con la notación introducida en el Capítulo 3, la relación de soporte del BUAF será representada mediante una flecha doble $\Longrightarrow$. En particular, dado que existe un especial interés en identificar las nociones de backing y undercutting, la representación gráfica distinguirá entre los diferentes tipos de ataque, así como también entre el soporte que los backings proveen a los warrants y el soporte que los sub-argumentos proveen a sus super-argumentos.

Dado el ataque de un argumento $\mathcal{A}$ sobre un argumento $\mathcal{B}$ (*i. e.*, $(\mathcal{A}, \mathcal{B}) \in \mathbb{R}$), utilizaremos la notación $\mathcal{A} \overset{R_b}{\rightarrowtail} \mathcal{B}$ cuando $(\mathcal{A}, \mathcal{B}) \in \mathbb{R}_b$, $\mathcal{A} \overset{U_c}{\rightarrowtail} \mathcal{B}$ cuando $(\mathcal{A}, \mathcal{B}) \in \mathbb{U}_c$, y $\mathcal{A} \overset{U_m}{\rightarrowtail} \mathcal{B}$ cuando $(\mathcal{A}, \mathcal{B}) \in \mathbb{U}_m$. En estos casos, las etiquetas 'R_b', 'U_c' y 'U_m' sobre la flecha con cola indican, respectivamente, que $\mathcal{A}$ es un atacante de tipo rebutting, undercutting o undermining para $\mathcal{B}$. Análogamente, dado el soporte de un argumento $\mathcal{A}$ para un argumento $\mathcal{B}$ (*i. e.*, $(\mathcal{A}, \mathcal{B}) \in \mathbb{S}$), utilizaremos la notación $\mathcal{A} \overset{b}{\Longrightarrow} \mathcal{B}$ cuando $(\mathcal{A}, \mathcal{B}) \in \mathbb{B}_k$, y $\mathcal{A} \overset{\sqsubseteq}{\Longrightarrow} \mathcal{B}$ cuando $(\mathcal{A}, \mathcal{B}) \in \mathbb{S}_{\sqsubseteq}$. Luego, las etiquetas '$b$' y '$\sqsubseteq$' sobre la flecha doble indican, respectivamente, que $\mathcal{A}$ es un argumento de backing para $\mathcal{B}$ y que $\mathcal{A}$ es un sub-argumento propio de $\mathcal{B}$. Para simplificar la representación gráfica, se incluirán únicamente las relaciones de sub-argumento "directas"; aquellas obtenidas por transitividad de la relación de sub-argumento podrán visualizarse al seguir los caminos de soporte indicados con $\overset{\sqsubseteq}{\Longrightarrow}$ en el grafo.

Consideremos nuevamente el ejemplo de Toulmin ilustrado en la Figura 4.1, el cual versa sobre de la nacionalidad de un sujeto llamado Harry. A partir de la situación descripta en este ejemplo es posible identificar los siguientes argumentos:

$\mathcal{H}$: *Harry nació en las Bermudas. Un hombre nacido en las Bermudas generalmente es Británico. Entonces, presumiblemente, Harry es Británico.*

$\mathcal{B}$: *Dados los siguientes estatutos y otras disposiciones legales:...*

$\mathcal{U}$: *Los padres de Harry son extranjeros.*

Para una mayor claridad, al describir los componentes de un BUAF en los ejemplos de este capítulo se especificarán directamente los subconjuntos no vacíos que integran las relaciones de ataque y soporte del BUAF. Es decir, en caso de existir, se identificarán

$$\mathcal{U} \xrightarrow{U_c} \mathcal{H} \xLeftarrow{b} \mathcal{B}$$

Figura 4.2: Representación gráfica correspondiente al $\mathsf{BUAF}_{4.1}$ del Ejemplo 4.1.

los elementos pertenecientes a $\mathbb{R}_b$, $\mathbb{U}_c$, $\mathbb{U}_m$, $\mathbb{B}_k$ y $\mathbb{S}_{\sqsubseteq}$ respectivamente. De esta manera, la especificación de un BUAF permitirá identificar los diferentes tipos de ataque y soporte que ocurren entre los argumentos del marco argumentativo. Análogamente, cuando resulte conveniente, algunas definiciones y proposiciones de este capítulo referenciarán directamente a los subconjuntos que componen las relaciones de ataque y soporte de un BUAF.

Ejemplo 4.1 *Una posible representación para el escenario descripto por el ejemplo de Toulmin está dada por el* $\mathsf{BUAF}_{4.1} = \langle \mathbb{A}_{4.1}, \mathbb{R}_{4.1}, \mathbb{S}_{4.1}, \preceq_{4.1} \rangle$*, donde:*

$$\mathbb{A}_{4.1} = \{\mathcal{H},\ \mathcal{B},\ \mathcal{U}\} \qquad \mathbb{B}_{k4.1} = \{(\mathcal{B}, \mathcal{H})\}$$
$$\mathbb{U}_{c4.1} = \{(\mathcal{U}, \mathcal{H})\} \qquad \preceq_{4.1} = \{(\mathcal{B}, \mathcal{U})\}$$

La Figura 4.2 ilustra la representación gráfica del $\mathsf{BUAF}_{4.1}$*. En particular, podemos observar que el* $\mathsf{BUAF}_{4.1}$ *no contiene ataques de tipo rebutting y undermining, ni tampoco posee argumentos vinculados por la relación de sub-argumento. Por otra parte, la existencia de estatutos y otras disposiciones legales provee soporte para el warrant del argumento* $\mathcal{H}$*, lo cual se encuentra expresado por el par* $(\mathcal{B}, \mathcal{H})$ *perteneciente a la relación de backing* $\mathbb{B}_{k4.1}$*. Además, el hecho de que los padres de Harry son extranjeros constituye un undercut*[1] *para la inferencia de* $\mathcal{H}$*, representado mediante el par* $(\mathcal{U}, \mathcal{H})$ *perteneciente a la relación de ataque* $\mathbb{U}_{c4.1}$ *de* $\mathsf{BUAF}_{4.1}$*.*

En las siguientes secciones se determinarán los diferentes tipos de derrota existentes entre argumentos de un BUAF, para luego identificar los conjuntos de argumentos que estarán finalmente aceptados de acuerdo a las diferentes semánticas de aceptabilidad consideradas.

[1]Dado un argumento $\mathcal{A}$ que produce un ataque de tipo undercutting sobre un argumento $\mathcal{B}$, diremos que $\mathcal{A}$ es un undercut para $\mathcal{B}$.

4.3. Obtención de Derrotas entre Argumentos

A partir de la especificación de un BUAF podemos identificar argumentos que se encuentran en conflicto al estar vinculados por la relación de ataque. En consecuencia, es necesario proveer un mecanismo que permita resolver estos conflictos, obteniendo así las derrotas entre los argumentos del sistema. Intuitivamente, dado que el BUAF cuenta con una relación de preferencia entre argumentos, estas preferencias serán utilizadas para determinar el éxito de los ataques. Siguiendo esta intuición, en esta sección se caracterizarán tres tipos de derrota que pueden ocurrir en el contexto de un BUAF. La primera de ellas corresponde a la resolución de los conflictos expresados en la relación de ataque del BUAF. Adicionalmente, se identificarán derrotas que surgen de la coexistencia de argumentos de backing y undercutting, para los cuales se mostrará que existe un conflicto. Luego, teniendo en cuenta que un argumento puede estar soportado por otros argumentos (mediante las relaciones de backing y sub-argumento), se definirá una noción de derrota que capture la situación en que cuando un argumento de soporte queda derrotado, se propague la derrota hacia los argumentos que este soporta.

Dada la relación de ataque de un BUAF, los ataques de tipo rebutting y undermining pueden resolverse directamente mediante la aplicación de preferencias sobre los argumentos en conflicto. De esta manera, tendremos que un argumento $\mathcal{A}$ derrota a un argumento $\mathcal{B}$ si lo ataca y la relación de preferencia establece que $\mathcal{A}$ no es menos preferido que $\mathcal{B}$. Por otra parte, a diferencia de otras aproximaciones como [Pra09], para determinar el éxito de los ataques de tipo undercutting se tomará en cuenta la existencia de los backings. Así, dado un argumento $\mathcal{A}$ que es un undercut para $\mathcal{B}$, si no existe backing alguno para $\mathcal{B}$ el ataque tendrá éxito; en caso contrario, será necesario comparar al argumento atacante con los backings existentes para el argumento atacado. Estas derrotas se identifican como *derrotas primarias*, y se encuentran caracterizadas por la siguiente definición.

Definición 4.2 (Derrota Primaria) *Sea* $\langle \mathbb{A}, \mathbb{R}, \mathbb{S}, \preceq \rangle$ *un BUAF y sean* $\mathcal{A}, \mathcal{C} \in \mathbb{A}$. *Diremos que* $\mathcal{A}$ derrota primariamente *a* $\mathcal{C}$*, notado* $\mathcal{A} \longrightarrow \mathcal{C}$*, si y sólo si se verifica una de las siguientes condiciones:*

- $(\mathcal{A}, \mathcal{C}) \in (\mathbb{R}_b \cup \mathbb{U}_m)$ *y* $\mathcal{A} \not\prec \mathcal{C}$*,*
- $(\mathcal{A}, \mathcal{C}) \in \mathbb{U}_c$ *y* $\nexists \mathcal{B} \in \mathbb{A}$ *tal que* $(\mathcal{B}, \mathcal{C}) \in \mathbb{B}_k$*, o*

$$\begin{tikzcd} \mathcal{A} \arrow[r, tail, "R_b/U_m"] \arrow[r, bend right] & \mathcal{C} \end{tikzcd} \qquad \begin{tikzcd} \mathcal{A} \arrow[r, tail, "U_c"] \arrow[r, bend right] & \mathcal{C} \end{tikzcd} \qquad \begin{tikzcd} \mathcal{A} \arrow[r, tail, "U_c"] \arrow[r, bend right] & \mathcal{C} & \mathcal{B} \arrow[l, Rightarrow, "b"'] \end{tikzcd}$$

$$\mathcal{A} \not\prec \mathcal{C} \qquad \nexists\ (\mathcal{B}, \mathcal{C}) \in \mathbb{B}_k \qquad \mathcal{A} \not\prec \mathcal{B}$$

(a) (b) (c)

Figura 4.3: Casos de derrota primaria para un BUAF, correspondientes a la Definición 4.2.

- $(\mathcal{A}, \mathcal{C}) \in \mathbb{U}_c$ *y* $\exists \mathcal{B} \in \mathbb{A}$ *tal que* $(\mathcal{B}, \mathcal{C}) \in \mathbb{B}_k$ *y* $\mathcal{A} \not\prec \mathcal{B}$.

Los diferentes casos de derrota primaria se hallan ilustrados en la Figura 4.3, donde las derrotas primarias se identifican mediante flechas sólidas sin cola. Cabe destacar que un ataque (ya sea de tipo rebutting, undercutting o undermining) no siempre tendrá éxito, con lo cual no necesariamente resultará en una derrota. El primer caso de la Definición 4.2, ilustrado en la Figura 4.3(a), considera conjuntamente los ataques de tipo rebutting y undermining. Esto se debe a que, dada la naturaleza abstracta de los argumentos, no es posible obtener mayor información que permita resolver estos ataques de diferente forma. Por lo tanto, para determinar su éxito simplemente se compara el argumento atacante con el argumento atacado de acuerdo a la relación de preferencia del BUAF.

Para los ataques de tipo undercutting, ilustrados en los casos (b) y (c) de la Figura 4.3, la Definición 4.2 considera la existencia de backings para el argumento atacado. Por una parte, ante la ausencia de backings, los ataques de tipo undercutting siempre tendrán éxito (Figura 4.3(b)). Por otra parte, en caso de existir algún backing, se efectúa una comparación entre el argumento atacante y los argumentos de backing para el argumento atacado (Figura 4.3(c)). Esto se debe al rol que se asigna a los argumentos de backing y undercutting en un BUAF. Como se mencionó anteriormente, un undercut en el contexto de un BUAF expresa un ataque sobre el warrant de un argumento. Luego, ante la presencia de undercuts, los backings son considerados para intentar prevenir la obtención de derrotas sobre los argumentos que soportan.

Ejemplo 4.2 *Consideremos un nuevo escenario que extiende a la situación descripta por el Ejemplo 4.1, donde se añade la información expresada por el siguiente argumento.*

$\mathcal{P}$: *El certificado de nacimiento de Harry indica que nació en París. Por lo tanto, Harry no nació en las Bermudas.*

$$\begin{tikzcd} \mathcal{U} \arrow[r, tail, "U_c"] & \mathcal{H} & \mathcal{B} \arrow[l, Rightarrow, "b"'] \\ & \mathcal{P} \arrow[u, tail, "U_m"] & \end{tikzcd}$$

Figura 4.4: Representación gráfica del $\mathsf{BUAF}_{4.2}$, correspondiente al Ejemplo 4.2.

En este caso, el argumento $\mathcal{P}$ origina un ataque de tipo undermining sobre el argumento $\mathcal{H}$, atacando, en particular, la premisa de que Harry nació en las Bermudas. Además, supongamos que se extiende la relación de preferencia del $\mathsf{BUAF}_{4.1}$ del Ejemplo 4.1 de manera tal que $\mathcal{H} \preceq \mathcal{P}$. El BUAF resultante de considerar estos nuevos elementos es $\mathsf{BUAF}_{4.2} = \langle \mathbb{A}_{4.2}, \mathbb{R}_{4.2}, \mathbb{S}_{4.2}, \preceq_{4.2} \rangle$, donde:

$$\begin{array}{ll} \mathbb{A}_{4.2} = \{\mathcal{H},\ \mathcal{B},\ \mathcal{U},\ \mathcal{P}\} & \mathbb{B}_{k4.2} = \{(\mathcal{B}, \mathcal{H})\} \\ \mathbb{U}_{c4.2} = \{(\mathcal{U}, \mathcal{H})\} & \preceq_{4.2} = \{(\mathcal{B},\mathcal{U}), (\mathcal{H}, \mathcal{P})\} \\ \mathbb{U}_{m4.2} = \{(\mathcal{P}, \mathcal{H})\} & \end{array}$$

cuya representación gráfica se ilustra en la Figura 4.4. A partir de $\mathsf{BUAF}_{4.2}$ se tiene que el argumento $\mathcal{P}$ derrota primariamente al argumento $\mathcal{H}$, dado que el ataque sobre la premisa de $\mathcal{H}$ es exitoso. Por otra parte, existe una derrota primaria de $\mathcal{U}$ sobre $\mathcal{H}$, correspondiente a un ataque exitoso de tipo undercutting. En particular, podemos observar que el éxito de este ataque se deriva de la preferencia de $\mathcal{U}$ por sobre el backing $\mathcal{B}$ de $\mathcal{H}$. Es decir, $\mathcal{U}$ derrota primariamente a $\mathcal{H}$ ya que $(\mathcal{U}, \mathcal{H}) \in \mathbb{U}_{c4.2}$, $(\mathcal{B}, \mathcal{H}) \in \mathbb{B}_{k4.2}$ y $\mathcal{B} \preceq_{4.2} \mathcal{U}$.

A continuación se incluye un nuevo ejemplo que ilustra otra situación donde ocurren derrotas primarias entre los argumentos de un BUAF.

Ejemplo 4.3 *Consideremos un escenario en el que se intenta determinar cuánto tiempo tomará viajar en automóvil desde una ciudad C_1 hacia otra ciudad C_2, donde C_1 y C_2 se encuentran a 300 km de distancia entre sí. Supongamos que contamos con los siguientes argumentos:*

$\mathcal{M}$: *Viajaremos de C_1 a C_2 <u>m</u>anejando a 120 km/h por autopista. Generalmente es posible manejar por autopista a una velocidad constante y sin necesidad de detenerse. Por lo tanto, llegaríamos a C_2 en dos horas y media.*

$\mathcal{A}$: *La <u>a</u>utopista AU_1 conecta las ciudades C_1 y C_2.*

$\mathcal{P}$: *El tramo de la autopista AU_1 comprendido entre las ciudades C_1 y C_2 no posee <u>p</u>eajes y admite una velocidad máxima de 120 km/h.*

$\mathcal{C}$: *El noticioso televisivo matutino informó que hubo un accidente en el sector de AU_1 comprendido entre las ciudades C_1 y C_2, por lo que <u>c</u>erraron temporalmente varios carriles de la autopista, frenando el tránsito en ese sector.*

$\mathcal{R}$: *El último noticioso radial informó que la autopista AU_1 fue <u>r</u>eabierta por completo hace una hora.*

$\mathcal{D}$: *En la radio también informaron que la autopista AU_1 quedó <u>d</u>añada luego del accidente. Por lo tanto, la velocidad máxima en un radio de 10 km desde el lugar del accidente es 50 km/h.*

En este caso, $\mathcal{A}$ es un sub-argumento de $\mathcal{M}$ ya que establece la existencia de la autopista AU_1 que comunica las ciudades C_1 y C_2, obteniendo así una premisa de $\mathcal{M}$. De manera similar, el argumento $\mathcal{P}$ es un backing para el argumento $\mathcal{M}$. Esto se debe a que $\mathcal{P}$ indica que el tramo de la autopista AU_1 comprendido entre las ciudades C_1 y C_2 no posee peajes y admite una velocidad máxima de 120 km/h, soportando el warrant de $\mathcal{M}$ que establece que generalmente las autopistas admiten manejar a velocidad constante y sin detenerse. En contraste, el argumento $\mathcal{C}$ constituye un undercut para $\mathcal{M}$ por expresar que el tráfico en el sector de AU_1 comprendido entre C_1 y C_2 se vio frenado a causa de un accidente. Por otra parte, el argumento $\mathcal{R}$ origina un ataque de tipo rebutting sobre $\mathcal{C}$ al contradecir la conclusión de que varios carriles de AU_1 se encuentran cerrados. En consecuencia, dada la naturaleza simétrica de este tipo de ataques, existe también un ataque rebutting de $\mathcal{C}$ sobre $\mathcal{R}$. Finalmente, existe un ataque undermining de $\mathcal{D}$ sobre $\mathcal{M}$, ya que $\mathcal{D}$ establece que la velocidad máxima en un tramo de AU_1 comprendido entre C_1 y C_2 está reducida a 50 km/h, contradiciendo así la premisa de $\mathcal{M}$ que indica que se viajará a una velocidad de 120 km/h.

Supongamos que contamos con una relación de preferencia que favorece a argumentos con información más reciente por sobre aquellos con información más antigua. De esta manera, es posible representar la situación aquí descripta mediante el $\mathsf{BUAF}_{4.3} = \langle \mathbb{A}_{4.3}, \mathbb{R}_{4.3}, \mathbb{S}_{4.3}, \preceq_{4.3} \rangle$, donde:

$$\begin{tikzcd}
& \mathcal{D} \arrow[d, tail, "U_m"'] & \\
\mathcal{A} \arrow[r, Rightarrow, "\sqsubseteq"] & \mathcal{M} & \mathcal{P} \arrow[l, Rightarrow, "b"'] \\
& \mathcal{C} \arrow[u, tail, "U_c"] \arrow[r, tail, shift left, "R_b"] & \mathcal{R} \arrow[l, tail, shift left, "R_b"]
\end{tikzcd}$$

Figura 4.5: Representación gráfica correspondiente al $\mathsf{BUAF}_{4.3}$ del Ejemplo 4.3.

$\mathbb{A}_{4.3} = \{\mathcal{M}, \mathcal{A}, \mathcal{P}, \mathcal{C}, \mathcal{R}, \mathcal{D}\}$

$\mathbb{R}_{b4.3} = \{(\mathcal{R}, \mathcal{C}), (\mathcal{C}, \mathcal{R})\}$

$\mathbb{U}_{c4.3} = \{(\mathcal{C}, \mathcal{M})\}$

$\mathbb{U}_{m4.3} = \{(\mathcal{D}, \mathcal{M})\}$

$\mathbb{S}_{\sqsubseteq 4.3} = \{(\mathcal{A}, \mathcal{M})\}$

$\mathbb{B}_{k4.3} = \{(\mathcal{P}, \mathcal{M})\}$

$\preceq_{4.3} = \{(\mathcal{C}, \mathcal{R}), (\mathcal{P}, \mathcal{C}), (\mathcal{M}, \mathcal{D})\}$

La representación gráfica de $\mathsf{BUAF}_{4.3}$ *se halla ilustrada en la Figura 4.5. En este caso, a partir de* $\mathsf{BUAF}_{4.3}$ *se obtienen derrotas primarias de* $\mathcal{D}$ *sobre* $\mathcal{M}$*, de* $\mathcal{C}$ *sobre* $\mathcal{M}$*, y de* $\mathcal{R}$ *sobre* $\mathcal{C}$*.*

La Definición 4.2 establece cómo se resuelven los conflictos expresados en la relación de ataque de un BUAF, dando lugar a las derrotas correspondientes. Sin embargo, de manera similar a lo ocurrido en las aproximaciones analizadas en el Capítulo 3, la coexistencia de las relaciones de ataque y soporte en el BUAF da origen a otros conflictos. En particular, nótese que la Definición 4.2 establece que, para la resolución de ataques de tipo undercutting, es necesario comparar el argumento atacante con un argumento de backing para el argumento atacado. Esto sugiere la existencia de un conflicto entre los argumentos de backing y undercutting: mientras que los últimos atacan la conexión entre premisas y conexión de un argumento, los primeros proveen soporte para dicha conexión. Por lo tanto, los argumentos de backing y undercutting para un mismo argumento no deberían ser aceptados simultáneamente. Más aún, dado que el conflicto arriba mencionado podría no estar explícitamente considerado en la relación de ataque de un BUAF, es necesario asegurar esta restricción de aceptabilidad. Esta situación da origen a la noción de *derrota implícita* que se define a continuación.

Definición 4.3 (Derrota Implícita) *Sea* $\langle \mathbb{A}, \mathbb{R}, \mathbb{S}, \preceq \rangle$ *un BUAF y sean* $\mathcal{A}, \mathcal{B}, \mathcal{C} \in \mathbb{A}$*. Si* $(\mathcal{B}, \mathcal{C}) \in \mathbb{B}_k$ *y* $(\mathcal{A}, \mathcal{C}) \in \mathbb{U}_c$*, entonces:*

$$\begin{tikzcd} \mathcal{A} \arrow[r, tail, "U_c"] \arrow[rr, bend right] & \mathcal{C} & \mathcal{B} \arrow[l, Rightarrow, "b"'] \arrow[ll, bend right] \end{tikzcd}$$

$\mathcal{A} \equiv \mathcal{B}$

(a)

$$\begin{tikzcd} \mathcal{A} \arrow[r, tail, "U_c"] \arrow[rr, bend right] & \mathcal{C} & \mathcal{B} \arrow[l, Rightarrow, "b"'] \arrow[ll, bend right] \end{tikzcd}$$

$\mathcal{A} \bowtie \mathcal{B}$

(b)

$$\begin{tikzcd} \mathcal{A} \arrow[r, tail, "U_c"] \arrow[rr, bend right] & \mathcal{C} & \mathcal{B} \arrow[l, Rightarrow, "b"'] \end{tikzcd}$$

$\mathcal{B} \prec \mathcal{A}$

(c)

$$\begin{tikzcd} \mathcal{A} \arrow[r, tail, "U_c"] & \mathcal{C} & \mathcal{B} \arrow[l, Rightarrow, "b"'] \arrow[ll, bend right] \end{tikzcd}$$

$\mathcal{A} \prec \mathcal{B}$

(d)

Figura 4.6: Casos de derrota implícita en función de la relación de preferencia de un BUAF.

- $\mathcal{A}$ derrota implícitamente *a $\mathcal{B}$, denotado $\mathcal{A} \longrightarrow \mathcal{B}$, si y sólo si $\mathcal{A} \not\prec \mathcal{B}$; y*
- $\mathcal{B}$ derrota implícitamente *a $\mathcal{A}$, denotado $\mathcal{B} \longrightarrow \mathcal{A}$, si y sólo si $\mathcal{B} \not\prec \mathcal{A}$.*

La Definición 4.3 provee un mecanismo para considerar explícitamente el conflicto existente entre el undercut $\mathcal{A}$ y el backing $\mathcal{B}$ de un argumento $\mathcal{C}$. Como se mencionó anteriormente, este conflicto se desprende de la naturaleza opuesta de estos argumentos, los cuales respectivamente atacan y soportan la conexión expresada por el warrant asociado al argumento $\mathcal{C}$. Luego, para resolver este conflicto es necesario analizar la preferencia entre $\mathcal{A}$ y $\mathcal{B}$, dando lugar a las siguientes alternativas: (a) $\mathcal{A} \equiv \mathcal{B}$, (b) $\mathcal{A} \bowtie \mathcal{B}$, (c) $\mathcal{B} \prec \mathcal{A}$ o (d) $\mathcal{A} \prec \mathcal{B}$. La Figura 4.6 ilustra las derrotas implícitas ocurridas en cada uno de estos casos, las cuales son denotadas mediante flechas sólidas sin cola.

A partir de la Figura 4.6 podemos observar que para las alternativas (a) y (b) ocurrirá que $\mathcal{A}$ y $\mathcal{B}$ se derrotan implícitamente entre sí, no pudiendo entonces resolverse el conflicto entre ellos. Esto sucede porque, en tales casos, la relación de preferencia establece que ambos argumentos son equivalentes o incomparables, asignándole la misma fuerza a cada uno de ellos. Por otra parte, las alternativas (c) y (d) plantean casos en que uno de los argumentos involucrados en el conflicto es estrictamente preferido al otro. En consecuencia, dada la alternativa (c) se tendrá que $\mathcal{A}$ derrota implícitamente a $\mathcal{B}$, mientras que para la alternativa (d) se tendrá que $\mathcal{B}$ derrota implícitamente a $\mathcal{A}$.

Ejemplo 4.4 *Dado el $\mathsf{BUAF}_{4.3}$ del Ejemplo 4.3 podemos identificar un conflicto entre el undercut $\mathcal{C}$ y el backing $\mathcal{P}$ del argumento $\mathcal{M}$. Luego, como la relación de preferencia de $\mathsf{BUAF}_{4.3}$ es tal que $\mathcal{P} \preceq_{4.3} \mathcal{C}$ y $\mathcal{C} \not\preceq_{4.3} \mathcal{P}$ (equivalentemente, $\mathcal{P} \prec_{4.3} \mathcal{C}$), existe una derrota implícita de $\mathcal{C}$ sobre $\mathcal{P}$ y no existe una derrota implícita de $\mathcal{P}$ sobre $\mathcal{C}$.*

Nótese que la derrota implícita de un undercut $\mathcal{A}$ por sobre un argumento de backing $\mathcal{B}$ se encuentra relacionada con la derrota primaria de $\mathcal{A}$ sobre el argumento $\mathcal{C}$ al cual

$\mathcal{B}$ soporta. Esto se debe a que, de acuerdo a la Definición 4.2, la derrota primaria de $\mathcal{A}$ sobre $\mathcal{C}$ se obtuvo a partir de una comparación entre el undercut $\mathcal{A}$ y el backing $\mathcal{B}$ de $\mathcal{C}$. Si consideramos los diferentes casos de derrota implícita ilustrados en la Figura 4.6, podemos observar que para las alternativas (a), (b) y (c) existirá también una derrota primaria de $\mathcal{A}$ sobre $\mathcal{C}$. En contraste, para la alternativa (d) ilustrada en la Figura 4.6, el ataque undercutting de $\mathcal{A}$ sobre $\mathcal{C}$ no tiene éxito y, por lo tanto, $\mathcal{A}$ no derrota implícitamente a $\mathcal{C}$ sino que $\mathcal{C}$ derrota implícitamente a $\mathcal{A}$. La relación existente entre estos tipos de derrota se halla formalizada por la siguiente proposición.

Proposición 4.1 *Sea $\langle \mathbb{A}, \mathbb{R}, \mathbb{S}, \preceq \rangle$ un BUAF y $\mathcal{A}, \mathcal{B}, \mathcal{C} \in \mathbb{A}$ tal que $(\mathcal{A}, \mathcal{C}) \in \mathbb{U}_c$ y $(\mathcal{B}, \mathcal{C}) \in \mathbb{B}_k$. Si $\mathcal{A}$ derrota implícitamente a $\mathcal{B}$, entonces $\mathcal{A}$ derrota primariamente a $\mathcal{C}$.*

Demostración: ver Apéndice A.

Cabe destacar que la relación establecida por la Proposición 4.1 se verifica en un solo sentido. Es decir, no siempre ocurre que si un undercut $\mathcal{A}$ derrota primariamente a un argumento $\mathcal{C}$, entonces existe un backing $\mathcal{B}$ de $\mathcal{C}$ tal que $\mathcal{A}$ derrota implícitamente a $\mathcal{B}$. Esto se debe a que el segundo caso de la Definición 4.2 establece que, ante la ausencia de argumentos de backing, los ataques de tipo undercutting siempre resultarán en derrotas primarias. Por otra parte, el Ejemplo 4.5 ilustra un escenario concreto en el que ocurre la relación identificada por la Proposición 4.1.

Ejemplo 4.5 *Sea $\mathsf{BUAF}_{4.3}$ el marco argumentativo con backing y undercutting del Ejemplo 4.3. Como se ilustra en el Ejemplo 4.3, el argumento $\mathcal{C}$ derrota primariamente al argumento $\mathcal{M}$ por ser preferido ante el backing $\mathcal{P}$ de $\mathcal{M}$. Luego, como se ilustra en el Ejemplo 4.4, esta preferencia da lugar a una derrota implícita de $\mathcal{C}$ sobre $\mathcal{P}$.*

Por otra parte, la Definición 4.2 establece que para que un ataque de tipo undercutting sea exitoso, el argumento atacante no debe ser menos preferido que *algún* backing del argumento atacado. Es decir, si bien el argumento atacado puede poseer otros backings que sean estrictamente preferidos por sobre el argumento atacante, el ataque será considerado exitoso dando lugar a una derrota primaria. Esto pareciera sugerir un comportamiento contra-intuitivo ya que, de haber considerado otro argumento de backing en la comparación, podría haber ocurrido que el undercut no tenga éxito. Sin embargo, como se ilustra en el siguiente ejemplo, la consideración de otros argumentos de backing para el argumento destinatario del undercut dan origen a otras derrotas implícitas, salvaguardando así el resultado esperado para esta situación.

Ejemplo 4.6 *Consideremos los siguientes argumentos para decidir si una determinada habitación de un edificio se encuentra iluminada por la noche.*

$\mathcal{I}$: *La habitación H del edificio E está iluminada porque posee una lámpara conectada al tomacorriente que se enciende automáticamente al caer la noche.*

$\mathcal{R}$: *El edificio E se encuentra conectado a la red de suministro de energía de la ciudad.*

$\mathcal{F}$: *Existe una falla eléctrica en el barrio del edificio E, interrumpiendo el suministro de energía en ese sector de la ciudad.*

$\mathcal{G}$: *El edificio E posee un grupo electrógeno de emergencia que opera a combustible y le suministra energía.*

Claramente existe un conflicto entre el argumento $\mathcal{F}$ y el argumento $\mathcal{I}$, dado que este último presupone la provisión de energía para el funcionamiento de la lámpara. Por otra parte, los argumentos $\mathcal{R}$ y $\mathcal{G}$ proveen soporte para el warrant del argumento $\mathcal{I}$ que establece el requerimiento de electricidad para que las lámparas funcionen. Por último, supongamos que el argumento $\mathcal{F}$ es preferido ante el argumento $\mathcal{R}$, mientras que el argumento $\mathcal{G}$ es preferido por sobre el argumento $\mathcal{F}$.

Los argumentos arriba expuestos y sus interacciones pueden ser caracterizados por el $\mathsf{BUAF}_{4.6} = \langle \mathbb{A}_{4.6}, \mathbb{R}_{4.6}, \mathbb{S}_{4.6}, \preceq_{4.6} \rangle$, donde:

$$\mathbb{A}_{4.6} = \{\mathcal{I}, \mathcal{R}, \mathcal{F}, \mathcal{G}\} \qquad \mathbb{B}_{k4.6} = \{(\mathcal{R}, \mathcal{I}), (\mathcal{G}, \mathcal{I})\}$$
$$\mathbb{U}_{c4.6} = \{(\mathcal{F}, \mathcal{I})\} \qquad \preceq_{4.6} = \{(\mathcal{R}, \mathcal{F}), (\mathcal{F}, \mathcal{G})\}$$

La Figura 4.7(a) ilustra la representación gráfica de $\mathsf{BUAF}_{4.6}$. En este caso, el argumento $\mathcal{F}$ es un undercut para $\mathcal{I}$. Por otra parte, los argumentos $\mathcal{R}$ y $\mathcal{G}$ son backings para $\mathcal{I}$. Teniendo en cuenta esto y la relación de preferencia $\preceq_{4.6}$ de $\mathsf{BUAF}_{4.6}$, el ataque de $\mathcal{F}$ sobre $\mathcal{I}$ resulta exitoso, dando origen a una derrota primaria. Además existe una derrota implícita de $\mathcal{F}$ sobre $\mathcal{R}$, por ser este último menos preferido que el argumento de undercut para $\mathcal{I}$. Sin embargo, la relación de preferencia establece que el backing $\mathcal{G}$ de $\mathcal{I}$ es preferido por sobre el

$$\begin{tikzcd} \mathcal{F} \arrow[r, tail, "U_c"] & \mathcal{I} & \mathcal{R} \arrow[l, Rightarrow, "b"'] \\ & \mathcal{G} \arrow[u, Rightarrow, "b"] & \end{tikzcd}$$

(a)

$$\begin{tikzcd} \mathcal{F} \arrow[r] \arrow[rr, bend left] & \mathcal{I} & \mathcal{R} \\ & \mathcal{G} \arrow[ul] & \end{tikzcd}$$

(b)

Figura 4.7: (a) Representación gráfica para el $\mathsf{BUAF}_{4.6}$ del Ejemplo 4.6 y (b) las derrotas que origina.

undercut $\mathcal{F}$, dando lugar a una derrota implícita de $\mathcal{G}$ sobre $\mathcal{F}$. De esta manera, el backing $\mathcal{G}$ provee una defensa ante las derrotas originadas por $\mathcal{F}$, reinstalando a los argumentos que este último derrotó. La Figura 4.7(b) *ilustra las derrotas aquí mencionadas, las cuales son denotadas mediante flechas sólidas.*

Como se mencionó anteriormente, la Definición 4.3 (derrota implícita) aborda los conflictos originados a partir de la coexistencia de argumentos de backing y undercutting en un BUAF. Sin embargo, dado el rol asignado a los argumentos de backing, es necesario considerar otros conflictos que pueden surgir. La caracterización brindada por Toulmin indica que los backings proveen las condiciones que justifican la conexión entre las premisas y conclusión de un argumento (*i. e.*, su warrant asociado). De esta manera, dado un backing $\mathcal{B}$ para un argumento $\mathcal{A}$, si $\mathcal{B}$ no está aceptado, entonces las condiciones bajo las cuales se sustenta el warrant de $\mathcal{A}$ no están satisfechas. Por lo tanto, el argumento $\mathcal{A}$ tampoco debería estar aceptado dado que su warrant ya no posee el soporte provisto por el backing $\mathcal{B}$.

Consideremos ahora la relación de sub-argumento del BUAF. En la literatura se han propuesto diferentes características que una relación de sub-argumento debería satisfacer. En particular, podemos destacar el principio de composicionalidad definido en [Vre97], el cual establece un conjunto de restricciones ampliamente aceptadas por la comunidad de argumentación. Como se mencionó en la Sección 3.5.1, este principio captura la restricción de que un argumento no puede ser aceptado a menos que todos sus sub-argumentos estén aceptados. Por lo tanto, para cumplir con este principio, el BUAF debe proveer un mecanismo para que las derrotas sobre un argumento sean propagadas hacia sus super-argumentos.

$$\begin{tikzcd} \mathcal{A} \arrow[r] \arrow[rr, bend right] & \mathcal{B} \arrow[r, Rightarrow, "b/\sqsubseteq"] & \mathcal{C} \end{tikzcd}$$

Figura 4.8: Caracterización gráfica de la Definición 4.4.

Con el objetivo de capturar las intuiciones arriba mencionadas en el contexto de los BUAF, a continuación se define la noción de *derrota indirecta*. De esta manera, las derrotas indirectas propagarán las derrotas sobre los backings y sub-argumentos hacia los argumentos que estos soportan.

Definición 4.4 (Derrota Indirecta) *Sea $\langle \mathbb{A}, \mathbb{R}, \mathbb{S}, \preceq \rangle$ un BUAF y sean $\mathcal{A}, \mathcal{C} \in \mathbb{A}$. $\mathcal{A}$* derrota indirectamente *a $\mathcal{C}$, denotado $\mathcal{A} \longrightarrow \mathcal{C}$, si y sólo si $\exists \mathcal{B} \in \mathbb{A}$ tal que $(\mathcal{B}, \mathcal{C}) \in \mathbb{S}$ y $\mathcal{A}$ derrota primariamente, derrota implícitamente o derrota indirectamente a $\mathcal{B}$.*

La Figura 4.8 provee una representación gráfica de la situación contemplada por la Definición 4.4, donde la flecha sólida de $\mathcal{A}$ hacia $\mathcal{C}$ corresponde a una derrota indirecta. Para ilustrar un ejemplo concreto de esta situación, consideremos el siguiente ejemplo.

Ejemplo 4.7 *Supongamos un escenario que extiende al descripto en el Ejemplo 4.6 mediante el agregado de la siguiente información:*

$\mathcal{V}$: El tanque de combustible del generador se encuentra $\underline{\text{v}}$acío.

Este nuevo escenario puede representarse mediante el BUAF$_{4.7}$ = $\langle \mathbb{A}_{4.7}, \mathbb{R}_{4.7}, \mathbb{S}_{4.7}, \preceq_{4.7} \rangle$:

$$\begin{array}{ll} \mathbb{A}_{4.7} = \{\mathcal{I}, \mathcal{R}, \mathcal{F}, \mathcal{G}, \mathcal{V}\} & \mathbb{B}_{k4.7} = \{(\mathcal{R}, \mathcal{I}), (\mathcal{G}, \mathcal{I})\} \\ \mathbb{U}_{c4.7} = \{(\mathcal{F}, \mathcal{I})\} & \preceq_{4.7} = \{(\mathcal{R}, \mathcal{F}), (\mathcal{F}, \mathcal{G}), (\mathcal{G}, \mathcal{V})\} \\ \mathbb{U}_{m4.7} = \{(\mathcal{V}, \mathcal{G})\} & \end{array}$$

cuya representación gráfica se ilustra en la Figura 4.9(a). En particular, los argumentos $\mathcal{I}, \mathcal{R}, \mathcal{F}$ y $\mathcal{G}$ corresponden a aquellos introducidos en el Ejemplo 4.6. Luego, el argumento $\mathcal{V}$ es un atacante para el argumento de backing $\mathcal{G}$, atacando en particular la premisa de que el grupo electrógeno posee combustible. Además, la relación de preferencia de BUAF$_{4.7}$ es tal que $\mathcal{G} \preceq_{4.7} \mathcal{V}$.

$$\begin{tikzcd} \mathcal{F} \arrow[r, tail, "U_c"] & \mathcal{I} & \mathcal{R} \arrow[l, Rightarrow, "b"'] \\ & \mathcal{G} \arrow[u, Rightarrow, "b"] & \mathcal{V} \arrow[l, tail, "U_m"'] \end{tikzcd}$$

(a)

$$\begin{tikzcd} \mathcal{F} \arrow[r] \arrow[rr, bend left] & \mathcal{I} & \mathcal{R} \\ & \mathcal{G} \arrow[ul] & \mathcal{V} \arrow[l] \arrow[ul] \end{tikzcd}$$

(b)

Figura 4.9: (a) Representación gráfica para el $\mathsf{BUAF}_{4.7}$ del Ejemplo 4.7 y (b) las derrotas que origina.

A partir de $\mathsf{BUAF}_{4.7}$ *se tiene que, además de las derrotas identificadas en el Ejemplo 4.6, el argumento* $\mathcal{V}$ *derrota primariamente al argumento* $\mathcal{G}$*. Luego, dado que* $\mathcal{G}$ *es un argumento de backing para* $\mathcal{I}$*, esto conduce a una derrota indirecta de* $\mathcal{V}$ *sobre* $\mathcal{I}$*. Todas estas derrotas se hallan ilustradas en la La Figura 4.9*(b) *mediante flechas sólidas.*

El Ejemplo 4.7 ilustra una situación en la que la derrota indirecta se origina a partir de una derrota primaria y el soporte provisto por un argumento de backing. Por otra parte, la Definición 4.4 admite la obtención de derrotas indirectas resultantes de combinar la relación de soporte del BUAF y otra derrota indirecta. En particular, la recursión en esta definición es necesaria para capturar las derrotas surgidas a partir del encadenamiento de argumentos de soporte. Por ejemplo, consideremos un BUAF con argumentos $\mathcal{A}, \mathcal{B}, \mathcal{C}$ y $\mathcal{D}$ tales que $\mathcal{D}$ soporta a $\mathcal{C}$, $\mathcal{C}$ soporta a $\mathcal{B}$, y existe una derrota primaria de $\mathcal{A}$ sobre $\mathcal{D}$. La Figura 4.10 ilustra esta situación, donde la derrota primaria es denotada mediante una flecha sólida sin etiqueta. Luego, como $\mathcal{A}$ derrota primariamente a $\mathcal{D}$, quien a su vez soporta a $\mathcal{C}$, existe una derrota indirecta de $\mathcal{A}$ sobre $\mathcal{C}$, ilustrada en la Figura 4.10 mediante una flecha sólida etiquetada con (i). Esta derrota indirecta sobre $\mathcal{C}$ refleja el hecho de haber perdido el soporte provisto por $\mathcal{D}$. De esta manera, $\mathcal{C}$ ya no posee las bases necesarias para proveer el soporte para el argumento $\mathcal{B}$. En consecuencia, a partir de la derrota indirecta de $\mathcal{A}$ sobre $\mathcal{C}$ y el soporte de $\mathcal{C}$ hacia $\mathcal{B}$, se obtiene una derrota indirecta de $\mathcal{A}$ sobre $\mathcal{B}$, ilustrada en la Figura 4.10 mediante una flecha sólida etiquetada con (ii).

A partir de las definiciones 4.2, 4.3 y 4.4, la siguiente definición caracteriza una noción general de derrota que unifica todas las situaciones en que se obtienen derrotas en el contexto de un BUAF.

$$\begin{tikzcd} \mathcal{A} \arrow[r] \arrow[rrr, bend left, "(ii)"] \arrow[rr, bend right, "(i)"'] & \mathcal{D} \arrow[r, Rightarrow, "b/\sqsubseteq"] & \mathcal{C} \arrow[r, Rightarrow, "b/\sqsubseteq"] & \mathcal{B} \end{tikzcd}$$

Figura 4.10: Ejemplo de derrota indirecta originada a partir del encadenamiento de argumentos de soporte.

Definición 4.5 (Derrota) *Sea $\langle \mathbb{A}, \mathbb{R}, \mathbb{S}, \preceq \rangle$ un BUAF y sean $\mathcal{A}, \mathcal{B} \in \mathbb{A}$. Diremos que $\mathcal{A}$* derrota *a $\mathcal{B}$, denotado $\mathcal{A} \longrightarrow \mathcal{B}$, si y sólo si $\mathcal{A}$ derrota primariamente, derrota implícitamente o derrota indirectamente a $\mathcal{B}$.*

Al comienzo de la Sección 4.2 se introdujo una representación gráfica para los BUAF mediante un multi-grafo dirigido que modela las relaciones de ataque y soporte del BUAF. En particular, la noción de derrota introducida en la Definición 4.5 surge de la resolución de los conflictos identificados a partir de las relaciones de ataque y soporte del BUAF. Es por esto que, dado que la noción de derrota resulta central para la obtención de los argumentos aceptados de un sistema argumentativo, a continuación introduciremos una nueva representación gráfica para los BUAF centrada en la noción de derrota. A partir de un BUAF es posible construir su *grafo de derrotas* asociado, en el que los nodos son los argumentos del BUAF y los arcos corresponden a la relación de derrota entre argumentos de acuerdo a la Definición 4.5. Dada la notación introducida en la Definición 4.5, la cual sigue la convención adoptada en esta tesis, los arcos del grafo de derrotas serán denotados mediante una flecha sólida $\longrightarrow$. Para ilustrar esta representación consideremos el siguiente ejemplo.

Ejemplo 4.8 *Consideremos el $\mathsf{BUAF}_{4.7}$ del Ejemplo 4.7. Su grafo de derrotas asociado es el ilustrado en la Figura 4.9(b). Luego, a partir del grafo de derrotas de $\mathsf{BUAF}_{4.7}$, se observa que existen derrotas (primarias) de $\mathcal{F}$ sobre $\mathcal{I}$ y de $\mathcal{V}$ sobre $\mathcal{G}$. Además, existe una derrota (implícita) de $\mathcal{F}$ sobre $\mathcal{R}$ y, por otra parte, una derrota (implícita) de $\mathcal{G}$ sobre $\mathcal{F}$. Finalmente, el argumento $\mathcal{V}$ derrota (indirectamente) al argumento $\mathcal{I}$.*

Como se mencionó anteriormente, el objetivo final consiste en determinar los argumentos aceptados de un BUAF. Para esto, como se verá en la Sección 4.4, será necesario considerar las derrotas ocurridas entre los argumentos del sistema. Cabe destacar que la relación de derrota caracterizada por la Definición 4.5 (y, en consecuencia, el grafo de derrotas asociado a un BUAF) no permite distinguir entre los diferentes tipos de derrota. Sin embargo, como se verá a continuación, será posible efectuar el cálculo de aceptabilidad de los argumentos de un BUAF abstrayéndose de la naturaleza de las derrotas entre ellos.

4.4. Semánticas de Aceptabilidad para los Marcos Argumentativos con Backing y Undercutting

En esta sección se abordará el cálculo de aceptabilidad para los argumentos de un BUAF, alcanzando de esta manera el objetivo primordial: determinar los argumentos aceptados del sistema. Como se vio en la Sección 4.3, los argumentos de un BUAF pueden estar en conflicto, dando así lugar a un conjunto de derrotas entre ellos. Por lo tanto, aquellos argumentos vinculados por una derrota no deberían ser aceptados simultáneamente. A partir de esto, los argumentos de un BUAF serán sujetos a un proceso de evaluación en el que se les asignará un estado de aceptación. De este modo, un argumento quedará aceptado si sobrevive de alguna manera a las derrotas que recibe. En particular, este proceso de evaluación estará dado por la definición de las semánticas de aceptabilidad del BUAF.

El cálculo de aceptabilidad de los argumentos de un BUAF será realizado siguiendo el enfoque extensional propuesto por Dung en [Dun95]. En primera medida, se definirá una serie de nociones semánticas que establecen los requerimientos mínimos que un conjunto de argumentos aceptables debe satisfacer. Luego, a partir de estas restricciones, se brindará una especificación formal de las semánticas de aceptabilidad para el BUAF. Finalmente, la aplicación de estas semánticas conducirá a la obtención de un conjunto de extensiones que identifican los conjuntos de argumentos aceptados del BUAF.

Un requisito esencial para todo conjunto de argumentos aceptados a partir de un BUAF es la consistencia. Esto se debe a que no resulta racional promover un argumento que simultáneamente ataca y soporta a otro argumento. De ocurrir esto, el argumento en cuestión conformaría una pieza de razonamiento inconsistente y, por lo tanto, no debería ser aceptado. Teniendo en cuenta el rol esencial de la consistencia en un proceso de razonamiento argumentativo, los BUAF son especificados de manera tal que aseguran su cumplimiento. En otras palabras, la Definición 4.1 impide la pérdida de consistencia al requerir que las relaciones de ataque y soporte de un BUAF sean disjuntas.

Por otra parte, dado que el BUAF incorpora una relación de soporte entre argumentos, es necesario analizar cómo la coexistencia de las relaciones de ataque y soporte puede afectar a la aceptabilidad de sus argumentos. Para ilustrar esto, consideremos el escenario presentado en el Ejemplo 4.1, correspondiente al ejemplo de Toulmin que versa acerca de la nacionalidad de Harry. La situación descripta por el ejemplo de Toulmin se halla caracterizada por el $\mathsf{BUAF}_{4.1}$ que se ilustra en la Figura 4.2 (ver pág. 118).

A partir de la representación provista por el $\mathsf{BUAF}_{4.1}$ se tiene que el argumento $\mathcal{U}$ es un undercut para $\mathcal{H}$, mientras que el argumento $\mathcal{B}$ es un backing para $\mathcal{H}$. Dado el ataque exitoso de $\mathcal{U}$ sobre $\mathcal{H}$, no resultaría razonable aceptar en simultáneo a estos argumentos. Por otra parte, dado que no existen conflictos entre $\mathcal{B}$ y $\mathcal{H}$, resultaría adecuado aceptar conjuntamente a dichos argumentos. Luego, la relación de ataque del $\mathsf{BUAF}_{4.1}$ no establece conflicto alguno entre los argumentos $\mathcal{U}$ y $\mathcal{B}$ con lo cual, en principio, no pareciera haber razones que impidan su aceptación conjunta. No obstante esto, si consideramos la naturaleza de la relación de backing y el ataque de tipo undercutting, el conjunto $\{\mathcal{U}, \mathcal{B}\}$ no constituiría una posición coherente con respecto al argumento $\mathcal{H}$. Esto se debe a que $\{\mathcal{U}, \mathcal{B}\}$ estaría, simultáneamente, proveyendo razones a favor y en contra del warrant asociado al argumento $\mathcal{H}$. En otras palabras, el conjunto $\{\mathcal{U}, \mathcal{B}\}$ no sería coherente con respecto al argumento $\mathcal{H}$, el cual es un elemento externo al conjunto. Por lo tanto, independientemente de la semántica adoptada para obtener el conjunto de argumentos aceptados de un BUAF, se requerirá que no exista en él ningún par de argumentos tal que sean backing y undercutting para un mismo argumento. Continuando con el ejemplo de Toulmin representado por el $\mathsf{BUAF}_{4.1}$ ilustrado en la Figura 4.2, el conjunto $\{\mathcal{U}, \mathcal{B}\}$ no estará contenido en ningún conjunto de argumentos aceptados a partir de $\mathsf{BUAF}_{4.1}$. Con el objetivo de formalizar esta intuición, a continuación se introduce la noción de *coherencia externa* para un conjunto de argumentos del BUAF.

Definición 4.6 (Coherencia Externa) *Sea* $\langle \mathbb{A}, \mathbb{R}, \mathbb{S}, \preceq \rangle$ *un BUAF y* $S \subseteq \mathbb{A}$. *El conjunto* S *es* externamente coherente *si y sólo si* $\forall \mathcal{C} \in \mathbb{A} : \nexists \mathcal{A}, \mathcal{B} \in S$ *tal que* $(\mathcal{A}, \mathcal{C}) \in \mathbb{U}_c$ *y* $(\mathcal{B}, \mathcal{C}) \in \mathbb{B}_k$.

Para ilustrar la noción de coherencia externa, consideremos el siguiente ejemplo.

$$
\begin{tikzcd}
\mathcal{J} \arrow[r, tail, "R_b"] & \mathcal{K} & & \mathcal{P} \arrow[dl, Rightarrow, "b"'] \\
\mathcal{L} \arrow[ur, Rightarrow, "\sqsubseteq"] & \mathcal{M} \arrow[l, tail, "R_b"'] & \mathcal{N} \arrow[l, tail, "U_m"'] & \mathcal{O} \arrow[l, tail, "U_c"']
\end{tikzcd}
$$

(a)

$$
\begin{tikzcd}
\mathcal{J} & \mathcal{K} & & \mathcal{P} \arrow[d, shift left] \\
\mathcal{L} & \mathcal{M} \arrow[l] \arrow[u] & \mathcal{N} \arrow[l] & \mathcal{O} \arrow[l] \arrow[u, shift left]
\end{tikzcd}
$$

(b)

Figura 4.11: (a) $\mathsf{BUAF}_{4.9}$ del Ejemplo 4.9 y (b) su grafo de derrotas asociado.

Ejemplo 4.9 *Sea $\mathsf{BUAF}_{4.9} = \langle \mathbb{A}_{4.9}, \mathbb{R}_{4.9}, \mathbb{S}_{4.9}, \preceq_{4.9} \rangle$ un marco argumentativo con backing y undercutting, donde:*

$$
\begin{array}{ll}
\mathbb{A}_{4.9} = \{\mathcal{J}, \mathcal{K}, \mathcal{L}, \mathcal{M}, \mathcal{N}, \mathcal{O}, \mathcal{P}\} & \mathbb{B}_{k4.9} = \{(\mathcal{P}, \mathcal{N})\} \\
\mathbb{R}_{b4.9} = \{(\mathcal{J}, \mathcal{K}), (\mathcal{M}, \mathcal{L})\} & \mathbb{S}_{\sqsubseteq 4.9} = \{(\mathcal{L}, \mathcal{K})\} \\
\mathbb{U}_{c4.9} = \{(\mathcal{O}, \mathcal{N})\} & \preceq_{4.9} = \{(\mathcal{J}, \mathcal{K}), (\mathcal{L}, \mathcal{M}), (\mathcal{M}, \mathcal{N})\} \\
\mathbb{U}_{m4.9} = \{(\mathcal{N}, \mathcal{M})\} &
\end{array}
$$

La Figura 4.11(a) ilustra la representación gráfica de $\mathsf{BUAF}_{4.9}$, mientras que la Figura 4.11(b) ilustra su correspondiente grafo de derrotas. En este caso podemos identificar, entre otros, los siguientes conjuntos externamente coherentes: $\{\mathcal{J}, \mathcal{K}, \mathcal{L}\}$ ya que el ataque de $\mathcal{J}$ sobre $\mathcal{K}$ es de tipo rebutting y $\mathcal{L}$ es un sub-argumento de $\mathcal{K}$; y $\{\mathcal{M}, \mathcal{N}\}$ ya que ninguno de estos dos argumentos soporta a un tercero. En contraste, por ejemplo, el conjunto $\{\mathcal{O}, \mathcal{P}\}$ no es externamente coherente ya que el argumento $\mathcal{P}$ es un backing para el argumento $\mathcal{N}$, mientras que el argumento $\mathcal{O}$ es un undercut para $\mathcal{N}$.

Nótese que la Definición 4.6 considera únicamente aquellas situaciones en las que el conjunto analizado contiene simultáneamente a un backing y a un undercut para un mismo argumento. La coherencia externa no toma en cuenta situaciones que involucren el soporte de sub-argumentos dado que, en tales casos, la información brindada por las relaciones de ataque y soporte no es suficiente para identificar conflictos adicionales. Es decir, dado un conjunto de argumentos $S = \{\mathcal{A}, \mathcal{B}\}$ donde el argumento $\mathcal{A}$ ataca a un argumento $\mathcal{C}$ (ya sea un ataque de tipo rebutting, undercutting o undermining) y el argumento $\mathcal{B}$ es un sub-argumento de $\mathcal{C}$, esta información no resulta suficiente para identificar un conflicto entre $\mathcal{A}$ y $\mathcal{B}$. Esto se ve reflejado en el Ejemplo 4.9, donde el conjunto $\{\mathcal{J}, \mathcal{K}, \mathcal{L}\}$ es externamente coherente y ocurre que $\mathcal{L}$ es un sub-argumento de $\mathcal{K}$ y $\mathcal{J}$ es un atacante de tipo rebutting para $\mathcal{K}$. De manera similar, si el conjunto $S = \{\mathcal{A}, \mathcal{B}\}$ es tal que $\mathcal{A}$ es un backing para un argumento $\mathcal{C}$ y $\mathcal{B}$ es un atacante de tipo rebutting o undermining para $\mathcal{C}$, la información brindada por las relaciones de ataque y soporte no resulta suficiente para identificar un conflicto entre $\mathcal{A}$ y $\mathcal{B}$.

Cabe destacar que la noción de coherencia externa para un conjunto de argumentos del BUAF verifica ciertas propiedades. Por una parte, dado un conjunto de argumentos C que no es externamente coherente, todo super-conjunto D de C tampoco será externamente coherente. Esto se debe a que D contendrá a todos los elementos de C, incluidos los argumentos responsables de que C no sea externamente coherente. Por otra parte, puede ocurrir que dados dos conjuntos C_1 y C_2 que son externamente coherentes, la unión $C_1 \cup C_2$ resulte en un conjunto que no es externamente coherente. Para ejemplificar esto consideremos el $\mathsf{BUAF}_{4.9}$ del Ejemplo 4.9. A partir de $\mathsf{BUAF}_{4.9}$ es posible obtener los conjuntos $\{\mathcal{M}, \mathcal{O}\}$ y $\{\mathcal{L}, \mathcal{P}\}$, los cuales son externamente coherentes. No obstante esto, el conjunto $\{\mathcal{M}, \mathcal{O}, \mathcal{L}, \mathcal{P}\}$ no es externamente coherente ya que el argumento $\mathcal{O}$ es un undercut para el argumento $\mathcal{N}$, mientras que el argumento $\mathcal{P}$ es un backing para $\mathcal{N}$.

Como se mencionó anteriormente, la noción de coherencia externa caracteriza un requerimiento que los conjuntos de argumentos aceptados a partir de un BUAF deberían satisfacer. Sin embargo, existe otra característica esencial que los conjuntos de argumentos aceptados deben satisfacer con el objetivo de constituir una posición coherente. Concretamente, ningún conjunto de argumentos debería aceptar simultáneamente a dos argumentos que se hallen relacionados por una derrota. Como se ilustra en el Ejemplo 4.9, esta restricción no es capturada por la noción de coherencia externa. Esto se debe a que, para el caso del Ejemplo 4.9, el conjunto $\{\mathcal{M}, \mathcal{N}\}$ es externamente coherente y existe una derrota del argumento $\mathcal{N}$ sobre el argumento $\mathcal{M}$. Por lo tanto, con el objetivo de capturar esta restricción sobre los conjuntos de argumentos del BUAF, a continuación se introduce la noción de *libertad de conflictos* en forma análoga a lo realizado por Dung en [Dun95].

Definición 4.7 (Conjunto Libre de Conflictos) *Sea* $\langle \mathbb{A}, \mathbb{R}, \mathbb{S}, \preceq \rangle$ *un BUAF y* $S \subseteq \mathbb{A}$. *S es un* conjunto libre de conflictos *si y sólo si* $\nexists \mathcal{A}, \mathcal{B} \in S$ *tal que* $\mathcal{A} \longrightarrow \mathcal{B}$.

La Definición 4.7 emplea la noción de derrota, permitiendo entonces que un conjunto libre de conflictos contenga argumentos que se atacan entre sí y no se derrotan. Esto no resulta un inconveniente ya que, como se mencionó en la Sección 4.3, un ataque (ya sea de tipo rebutting, undercutting o undermining) no siempre tendrá éxito. Por lo tanto, la noción de libertad de conflictos se define de manera tal que considera únicamente aquellos conflictos expresados mediante derrotas, los cuales corresponden a ataques exitosos (como en el caso de derrotas primarias para un BUAF), a la consideración explícita de conflictos derivados de la coexistencia de undercuts y backings (como en el caso de las derrotas implícitas), o a la propagación de derrotas a través de cadenas de soporte (como en el caso de las derrotas indirectas).

Por ejemplo, dado el $\mathsf{BUAF}_{4.9}$ del Ejemplo 4.9 y su grafo de derrotas ilustrado en la Figura 4.11(b), es posible obtener, entre otros, los siguientes conjuntos libres de conflictos: $\emptyset$, $\{\mathcal{N}\}$, $\{\mathcal{M}, \mathcal{O}\}$ y $\{\mathcal{J}, \mathcal{K}, \mathcal{L}\}$. En contraste, aquellos conjuntos para los cuales existe al menos una derrota que involucra a dos miembros del conjunto no serán libres de conflicto. Nuevamente, dado el $\mathsf{BUAF}_{4.9}$ del Ejemplo 4.9, el conjunto $\{\mathcal{M}, \mathcal{N}\}$ no es libre de conflictos ya que existe una derrota de $\mathcal{N}$ sobre $\mathcal{M}$; similarmente, el conjunto $\{\mathcal{K}, \mathcal{N}, \mathcal{O}\}$ no es libre de conflictos ya que, si bien $\mathcal{K}$ no se halla derrotado por (ni derrota a) $\mathcal{N}$ u $\mathcal{O}$, existe una derrota de $\mathcal{O}$ sobre $\mathcal{N}$.

Como se observa a partir del Ejemplo 4.9, un conjunto de argumentos puede ser externamente coherente incluso cuando no es libre de conflictos. Tal es el caso del conjunto $\{\mathcal{M}, \mathcal{N}\}$ que, como se mostró en el Ejemplo 4.9, es externamente coherente pero no es libre de conflictos. Sin embargo, existe una relación entre las nociones de libertad de conflictos y coherencia externa. En particular, como enuncia la siguiente proposición, la noción de libertad de conflictos es lo suficientemente fuerte como para asegurar las restricciones impuestas por la coherencia externa.

Proposición 4.2 *Sea $\langle \mathbb{A}, \mathbb{R}, \mathbb{S}, \preceq \rangle$ un BUAF y $S \subseteq \mathbb{A}$. Si S es un conjunto libre de conflictos, entonces S es externamente coherente.*

Demostración: ver Apéndice A.

La Proposición 4.2 establece que todo conjunto libre de conflictos es externamente coherente y, equivalentemente, que un conjunto no externamente coherente tampoco es libre de conflictos. Por otra parte, la recíproca de esta proposición no se cumple. Es decir, si un conjunto de argumentos es externamente coherente, esto *no* implica que el conjunto sea libre de conflictos. Una vez más, para ilustrar esto consideremos el $\mathsf{BUAF}_{4.9}$ del Ejemplo 4.9. Como se vio en el Ejemplo 4.9, el conjunto $\{\mathcal{M}, \mathcal{N}\}$ es externamente coherente. Sin embargo, $\{\mathcal{M}, \mathcal{N}\}$ no es libre de conflictos ya que existe una derrota de $\mathcal{N}$ sobre $\mathcal{M}$.

Al comienzo de esta sección se presentó la intuición de que un argumento quedará aceptado si de alguna manera sobrevive a las derrotas que recibe. Para capturar esta intuición en el contexto de los BUAF, a continuación se define la noción de *aceptabilidad* de manera análoga a lo realizado en [Dun95].

Definición 4.8 (Aceptabilidad) *Sea $\langle \mathbb{A}, \mathbb{R}, \mathbb{S}, \preceq \rangle$ un BUAF, $\mathcal{A} \in \mathbb{A}$ y $S \subseteq \mathbb{A}$. $\mathcal{A}$ es* aceptable con respecto a *S si y sólo si $\forall \mathcal{B} \in \mathbb{A}$ tal que $\mathcal{B} \longrightarrow \mathcal{A}$, $\exists \mathcal{C} \in S$ tal que $\mathcal{C} \longrightarrow \mathcal{B}$.*

Intuitivamente, un argumento $\mathcal{A}$ será aceptable con respecto a un conjunto de argumentos S siempre y cuando S lo defienda ante todas las derrotas que recibe. De esta manera, la noción de aceptabilidad se encuentra relacionada con la noción de reinstalación introducida en la Sección 2.1: dado que S defiende a $\mathcal{A}$ de las derrotas que recibe, S produce la reinstalación de $\mathcal{A}$. Por otra parte, la Definición 4.8 no establece restricciones sobre el conjunto de argumentos S con lo cual, en particular, el conjunto S podría no ser libre de conflictos. No obstante esto, como se verá más adelante en esta sección, las nociones de libertad de conflictos y aceptabilidad serán consideradas en forma conjunta para definir la noción de *admisibilidad*. El siguiente ejemplo ilustra la noción de aceptabilidad con respecto a un conjunto de argumentos del BUAF.

Ejemplo 4.10 *Dado el $BUAF_{4.9}$ del Ejemplo 4.9, el argumento $\mathcal{K}$ es aceptable con respecto a los conjuntos $\{\mathcal{N}\}$, $\{\mathcal{N}, \mathcal{P}\}$ y $\{\mathcal{N}, \mathcal{J}, \mathcal{O}\}$, entre otros. Además, el sub-argumento $\mathcal{L}$ de $\mathcal{K}$ también es aceptable con respecto a tales conjuntos. Por otra parte, el argumento $\mathcal{N}$ es aceptable con respecto al conjunto $\{\mathcal{P}\}$, entre otros. Asimismo, el backing $\mathcal{P}$ de $\mathcal{N}$ también es aceptable con respecto al conjunto $\{\mathcal{P}\}$ ya que se defiende a sí mismo de la derrota originada por el argumento $\mathcal{O}$.*

En particular, el conjunto $\{\mathcal{N}, \mathcal{J}, \mathcal{O}\}$ no es libre de conflictos ya que $\mathcal{O}$ derrota a $\mathcal{N}$. Asimismo, también ocurre que $\mathcal{K}$ y su sub-argumento $\mathcal{L}$ son aceptables con respecto al conjunto $\{\mathcal{M}, \mathcal{N}\}$ aún cuando $\mathcal{M}$ derrota a $\mathcal{K}$. Esto se debe a que el argumento $\mathcal{N}$ defiende a $\mathcal{K}$ ante la derrota de $\mathcal{M}$, lo cual se ve reflejado en el hecho de que el conjunto $\{\mathcal{M}, \mathcal{N}\}$ no es libre de conflictos.

La siguiente proposición establece que, como se puede observar en el Ejemplo 4.10, si un argumento es defendido por un conjunto de argumentos (*i. e.*, es aceptable con respecto a él), entonces todos los argumentos que lo soportan (ya sean backings o sub-argumentos) también estarán defendidos por ese conjunto.

Proposición 4.3 *Sea* $\langle \mathbb{A}, \mathbb{R}, \mathbb{S}, \preceq \rangle$ *un* BUAF*,* $\mathcal{A} \in \mathbb{A}$ *y* $S \subseteq \mathbb{A}$*. Si* $\mathcal{A}$ *es aceptable con respecto a* S*, entonces* $\forall \mathcal{B} \in \mathbb{A}$ *tal que* $(\mathcal{B}, \mathcal{A}) \in \mathbb{S}$ *se verifica que* $\mathcal{B}$ *es aceptable con respecto a* S*.*

Demostración: ver Apéndice A.

A partir de la Proposición 4.2 tenemos que la noción de libertad de conflictos subsume a la noción de coherencia externa. Además, por la Definición 4.1, todo BUAF posee la característica de ser consistente. Por lo tanto, podemos utilizar las nociones de libertad de conflictos y aceptabilidad para caracterizar los requisitos mínimos a satisfacer por los conjuntos de argumentos aceptados de un BUAF. De esta manera, siguiendo la aproximación de [Dun95], definiremos la noción de admisibilidad para los BUAF. Intuitivamente, un conjunto de argumentos será admisible si es libre de conflictos y defiende a todos sus elementos de las derrotas que reciben.

Definición 4.9 (Admisibilidad) *Sea* $\langle \mathbb{A}, \mathbb{R}, \mathbb{S}, \preceq \rangle$ *un* BUAF *y* $S \subseteq \mathbb{A}$*. El conjunto* S *es* admisible *si y sólo si es libre de conflictos y* $\forall \mathcal{A} \in S$ *se verifica que* $\mathcal{A}$ *es aceptable con respecto a* S*.*

Ejemplo 4.11 *De los conjuntos de argumentos mencionados en el Ejemplo 4.10, los únicos conjuntos admisibles son* $\{\mathcal{P}\}$ *y* $\{\mathcal{N}, \mathcal{P}\}$*. En particular, el conjunto* $\{\mathcal{N}\}$ *no es admisible ya que no defiende a* $\mathcal{N}$ *ante la derrota de* $\mathcal{O}$*. Por otra parte, los conjuntos* $\{\mathcal{N}, \mathcal{J}, \mathcal{O}\}$ *y* $\{\mathcal{M}, \mathcal{N}\}$ *no son admisibles ya que no son libres de conflictos.*

Las semánticas de aceptabilidad definen el conjunto de extensiones de un marco argumentativo, las cuales corresponden a conjuntos de argumentos que son aceptados en forma conjunta. Como se mencionó en la Sección 2.2.1.2, en esta tesis se adoptará un enfoque extensional, considerando las cuatro semánticas de aceptabilidad propuestas por Dung en [Dun95]. Por lo tanto, siguiendo la aproximación de Dung, a continuación se definen las extensiones *completa*, *preferida*, *estable* y *grounded* para un BUAF, las cuales se basan en las nociones de libertad de conflictos, aceptabilidad y admisibilidad introducidas previamente.

Definición 4.10 (Extensiones) *Sea* BUAF $= \langle \mathbb{A}, \mathbb{R}, \mathbb{S}, \preceq \rangle$ *un marco argumentativo con backing y undercutting y* $S \subseteq \mathbb{A}$ *un conjunto libre de conflictos:*

- *S es una* extensión completa *de BUAF si y sólo si $\forall \mathcal{A} \in \mathbb{A}$: si $\mathcal{A}$ es aceptable con respecto a S, entonces $\mathcal{A} \in S$.*

- *S es una* extensión preferida *de BUAF si y sólo si es un conjunto admisible maximal con respecto a la inclusión de conjuntos.*

- *S es una* extensión estable *de BUAF si y sólo si $\forall \mathcal{A} \in \mathbb{A} \backslash S$: $\exists \mathcal{B} \in S$ tal que $\mathcal{B} \longrightarrow \mathcal{A}$.*

- *S es la* extensión grounded *de BUAF si y sólo si es la extensión completa minimal con respecto a la inclusión de conjuntos.*

Ejemplo 4.12 *A partir del $\mathsf{BUAF}_{4.9}$ del Ejemplo 4.9 se obtienen los siguientes conjuntos de extensiones:*

- *las extensiones completas $\{\mathcal{J}\}$, $\{\mathcal{J},\mathcal{K}\}$, $\{\mathcal{J},\mathcal{L}\}$, $\{\mathcal{J},\mathcal{K},\mathcal{L}\}$, $\{\mathcal{J},\mathcal{K},\mathcal{L},\mathcal{N}\}$, $\{\mathcal{J},\mathcal{K},\mathcal{L},\mathcal{N},\mathcal{P}\}$, $\{\mathcal{J},\mathcal{M}\}$ y $\{\mathcal{J},\mathcal{M},\mathcal{O}\}$;*

- *las extensiones preferidas y estables $\{\mathcal{J},\mathcal{K},\mathcal{L},\mathcal{N},\mathcal{P}\}$ y $\{\mathcal{J},\mathcal{M},\mathcal{O}\}$; y*

- *la extensión grounded $\{\mathcal{J}\}$.*

Como se puede observar a partir del Ejemplo 4.12, la aplicación de semánticas de aceptabilidad alternativas conduce a la obtención de diferentes conjuntos de argumentos aceptados a partir del BUAF. Luego, siguiendo la aproximación presentada en la Sección 2.2.1.2, identificaremos los argumentos que se encuentran *escépticamente aceptados*, *crédulamente aceptados* o *rechazados* con respecto a una determinada semántica de aceptabilidad.

Definición 4.11 (Argumentos Aceptados y Argumentos Rechazados) *Sea $\mathsf{BUAF} = \langle \mathbb{A}, \mathbb{R}, \mathbb{S}, \preceq \rangle$ un marco argumentativo con backing y undercutting, $\mathsf{s} \in \{$completa, preferida, estable, grounded$\}$ una semántica de aceptabilidad y $\mathcal{A} \in \mathbb{A}$:*

- *El argumento $\mathcal{A}$ está* escépticamente aceptado *con respecto a s si y sólo si pertenece a toda extensión de BUAF bajo la semántica s.*

- *El argumento $\mathcal{A}$ está* crédulamente aceptado *con respecto a s si y sólo si pertenece a alguna extensión de BUAF (pero no a todas) bajo la semántica s.*

- *El argumento $\mathcal{A}$ está* rechazado *con respecto a* s *si y sólo si no pertenece a ninguna extensión de* BUAF *bajo la semántica* s.

Nótese que, dadas las características de la semántica grounded introducida en la Definición 4.10, todo argumento que esté aceptado con respecto a la semántica grounded estará, en particular, escépticamente aceptado. Esto se debe a que para todo BUAF existirá exactamente una extensión grounded. Por lo tanto, cuando se analice la aceptación de un argumento con respecto a la semántica grounded, no será necesario distinguir entre aceptación escéptica y aceptación crédula.

Ejemplo 4.13 *Consideremos el* $\mathsf{BUAF}_{4.9}$ *del Ejemplo 4.9, cuyas extensiones se ilustran en el Ejemplo 4.12. Dada la semántica completa, el argumento* $\mathcal{J}$ *se encuentra escépticamente aceptado ya que pertenece a toda extensión completa de* $\mathsf{BUAF}_{4.10}$*. Luego, los argumentos* $\mathcal{K}, \mathcal{L}, \mathcal{M}, \mathcal{N}, \mathcal{O}$ *y* $\mathcal{P}$ *están crédulamente aceptados con respecto a la semántica completa, mientras que no hay argumentos rechazados con respecto a esta semántica. Para las semánticas preferida y estable ocurre lo mismo que para la semántica completa, por lo que el argumento* $\mathcal{J}$ *está escépticamente aceptado, mientras que los argumentos* $\mathcal{K}, \mathcal{L}, \mathcal{M}, \mathcal{N}, \mathcal{O}$ *y* $\mathcal{P}$ *están crédulamente aceptados. Finalmente, el único argumento aceptado con respecto a la semántica grounded es* $\mathcal{J}$*.*

Las definiciones 4.8, 4.9 y 4.10 son análogas a aquellas presentadas en la Sección 2.2.1.2 para los AF de Dung. Por otra parte, recordemos que un AF está caracterizado por un conjunto de argumentos y una relación de derrota entre ellos. Por lo tanto, utilizando el conjunto de argumentos de un BUAF y la relación de derrota correspondiente a la Definición 4.5, es posible caracterizar un AF que acepta exactamente los mismos argumentos que el BUAF. Este resultado se halla formalizado mediante la siguiente proposición.

Proposición 4.4 *Sea* $\mathsf{BUAF} = \langle \mathbb{A}, \mathbb{R}, \mathbb{S}, \preceq \rangle$ *un marco argumentativo con backing y undercutting. Existe un marco argumentativo abstracto* $\mathsf{AF} = \langle \mathbb{A}, \longrightarrow \rangle$ *tal que para toda semántica* $\mathsf{s} \in \{completa,\ preferida,\ estable,\ grounded\}$ *las extensiones de* BUAF *y* AF *bajo* s *coinciden.*

Demostración: ver Apéndice A.

A partir de los resultados mostrados en esta sección, particularmente por la Proposición 4.4, los BUAF heredan todas las propiedades demostradas para los AF (ver Sección 2.2.1, o para mayores detalles referirse a [Dun95]). Por ejemplo, esto admite el uso, en el contexto de un BUAF, de semánticas de aceptabilidad alternativas propuestas en la literatura (ver [BCG11]). Asimismo, dado que los AF constituyen en la actualidad uno de los formalismos más utilizados para el desarrollo de nuevos avances en argumentación, esto permitirá a los BUAF aprovechar los resultados futuros que se elaboren sobre los AF.

4.5. Trabajo Relacionado

En este capítulo se presentó un formalismo de argumentación abstracta que toma como base los marcos argumentativos propuestos en [Dun95]. El formalismo desarrollado recibe el nombre de *Marco Argumentativo con Backing y Undercutting (BUAF)* y extiende a los AF de Dung mediante la incorporación de relaciones preferencia y soporte entre argumentos. Por otra parte, a diferencia del AF, el cual considera una relación de derrota entre argumentos, la especificación del BUAF considera una relación de ataque entre argumentos. Luego, a partir del uso de preferencias, los conflictos especificados por la relación de ataque del BUAF son resueltos, dando lugar a un conjunto de derrotas. Además, dada la coexistencia de argumentos de ataque y soporte para inferencias, el BUAF identifica una serie de conflictos que pueden no haber sido especificados mediante la relación de ataque. De esta manera, se caracteriza una serie de derrotas correspondientes a la resolución de tales conflictos.

En particular, el BUAF provee un marco unificado para representar las nociones de backing y undercutting propuestas por Toulmin [Tou58] y Pollock [Pol87]. Cabe destacar que, si bien las nociones de backing y undercutting han sido abordadas con anterioridad por diferentes aproximaciones existentes en la literatura, ninguna de las propuestas de argumentación abstracta desarrolladas hasta el momento había considerado estas nociones en forma conjunta. Por ejemplo, en [Pra09] se propone un sistema argumentativo llamado ASPIC+, en el cual los argumentos son provistos parcialmente de estructura interna. Luego, a partir de la estructura de los argumentos se identifican diferentes tipos de ataque entre argumentos, incluido el undercut. Sin embargo, la formalización brindada en [Pra09] no toma en cuenta la existencia de una relación de soporte entre argumentos. Asimismo, como se mencionó en la Sección 3.2.1, en [BGvdTV10] se propone una alternativa para modelar soporte rebatible, mediante la existencia de ataques sobre la relación de soporte. De esta manera, los autores de [BGvdTV10] capturan la noción de undercutting defeat propuesta por Pollock en el contexto de un marco argumentativo abstracto con soporte. No obstante esto, en [BGvdTV10] se adopta una interpretación deductiva para la relación de soporte entre argumentos, dejando así de lado la noción de backing propuesta por Toulmin. En contraste, como se verá en la Sección 5.7, en [Ver05] se propone una reconstrucción de las ideas de Toulmin utilizando DefLog [Ver03a], la cual permite además modelar la noción de undercutting defeater propuesta por Pollock.

Al igual que en [AC02b], el BUAF incorpora una relación de preferencia entre argumentos para determinar el éxito de los ataques. Sin embargo, a diferencia de la formalización propuesta en [AC02b], el BUAF no se centra en el estudio de preferencias, sino que estas son meramente utilizadas para la resolución de los conflictos entre argumentos. El uso de preferencias ha demostrado ser fundamental durante el proceso de razonamiento no monótono [BNT08]. Es por esto que existe una gran cantidad de trabajos que abordan el estudio del uso de preferencias en argumentación (ver *e. g.* [AC02b, AV11b, AV11a, Mod09]). En particular, en [AV11b] los autores muestran que las preferencias juegan dos roles en los marcos argumentativos: son utilizadas para *i*) computar las soluciones estándares (o extensiones) y para *ii*) refinar dichas soluciones. Por otra parte, en [Mod09] se introduce un marco argumentativo abstracto que permite modelar preferencias entre argumentos. El formalismo propuesto en [Mod09] no define explícitamente una relación de preferencia entre argumentos, sino que las preferencias son expresadas a nivel objeto mediante la incorporación de una relación de derrota de segundo nivel. Teniendo en cuenta esto último, podemos observar que el formalismo propuesto en [BCGG09, BCGG11] también permite el modelado de preferencias mediante una relación de derrota recursiva[2].

A continuación analizaremos la relación existente entre el BUAF y los formalismos argumentativos presentados en el Capítulo 3. Se considerarán los Marcos Argumentativos Bipolares (BAF), los Marcos Argumentativos Bipolares con Soporte Deductivo (d-BAF), los Marcos Argumentativos con Soporte por Necesidad (AFN) y los Sistemas Argumentativos Evidenciales (EAS), respectivamente introducidos en las secciones 3.1, 3.2, 3.3 y 3.4.

[2]En [BCGG09, BCG11] los autores utilizan la terminología "ataque". Sin embargo, al igual que en los marcos argumentativos de [Dun95] esta relación representa ataques exitosos. Por lo tanto, para evitar confusiones, utilizamos la terminología relación de "derrota".

En particular, el análisis se concentrará en estos formalismos por tratarse de aproximaciones que modelan una relación de soporte entre argumentos en forma explícita.

Para efectuar la comparación entre el BUAF y los formalismos arriba mencionados nos concentraremos en las interpretaciones de soporte adoptadas por cada uno de ellos. Recordemos que, de los formalismos arriba mencionados, los BAF, d-BAF y AFN definen un conjunto de derrotas que se infieren a partir de la consideración conjunta de las relaciones de derrota y soporte. Siguiendo la terminología utilizada en [CLS11, CLS13] estas son denominadas *derrotas complejas*. Cabe destacar que la importancia de las derrotas complejas definidas por cada una de estas aproximaciones radica en que refuerzan las restricciones de aceptabilidad impuestas por la correspondiente interpretación de soporte. Por lo tanto, para efectuar el análisis, se estudiará la relación existente entre las derrotas complejas definidas por elBUAF y los otros formalismos arriba mencionados. El análisis considerará en primera medida las derrotas complejas definidas por el BUAF, analizando su relación con las derrotas complejas definidas por otros formalismos. Luego, se pasará a considerar las derrotas complejas de otros formalismos que no son capturadas directamente por los BUAF.

A partir de las derrotas definidas en la Sección 4.3, consideraremos como derrotas complejas únicamente a las derrotas indirectas y las derrotas implícitas. En contraste, las derrotas primarias no serán consideradas como derrotas complejas dado que son obtenidas a partir de las relaciones de ataque y preferencia de los BUAF, es decir, sin considerar la relación de soporte[3]. Como se mostró en la Sección 4.3, las derrotas indirectas son obtenidas a partir de la combinación de otras derrotas y la relación de soporte de un BUAF. Por otra parte, las derrotas implícitas capturan los conflictos existentes entre argumentos de backing y undercutting que respectivamente soportan y atacan a un mismo argumento. En particular, estas últimas toman en cuenta, además, la relación de preferencia entre argumentos. No obstante esto, como se mencionó en la Sección 4.3, ante la presencia de un backing y un undercut para el mismo argumento, existirá al menos una derrota indirecta (ya sea del backing sobre el undercut o a la inversa). Por lo tanto, sin pérdida de generalidad, para efectuar el análisis inicialmente nos abstraeremos de la relación de preferencia del BUAF.

---

[3]Si bien el segundo y tercer caso de la Definición 4.2 toman en cuenta la existencia de un argumento de backing, este último es meramente considerado a los efectos de realizar la comparación entre argumentos para determinar el éxito de un ataque de tipo undercutting.

$$\begin{tikzcd} \mathcal{A} \arrow[r] \arrow[rr, bend right] & \mathcal{B} \arrow[r, Rightarrow, "b/\sqsubseteq"] & \mathcal{C} \end{tikzcd}$$

Figura 4.12: Derrota indirecta de un argumento $\mathcal{A}$ sobre un argumento $\mathcal{C}$ en un BUAF.

$$\begin{tikzcd} \mathcal{A} \arrow[r, tail, "U_c"] \arrow[rr, bend right] & \mathcal{C} & \mathcal{B} \arrow[l, Rightarrow, "b"'] \end{tikzcd}$$

Figura 4.13: Primer caso de derrota implícita en un BUAF.

Consideremos las derrotas indirectas caracterizadas en la Definición 4.4. Esta definición contempla, mediante la recursión, las derrotas indirectas surgidas de la consideración de una cadena de soportes. Por lo tanto, sin pérdida de la generalidad, podemos considerar el simple caso en que la derrota indirecta surge de la consideración de un único soporte (ya sea backing o sub-argumento). La Figura 4.12 ilustra la representación gráfica correspondiente a esta situación, donde se tiene una derrota indirecta de $\mathcal{A}$ sobre $\mathcal{C}$.

Podemos observar que la derrota indirecta ilustrada en la Figura 4.12 se corresponde con la derrota secundaria del BAF presentado en la Sección 3.1, así como también con el primer caso de derrota extendida del AFN introducido en la Sección 3.3. Esta correspondencia es de esperarse dado que la relación de soporte del BUAF combina las nociones de sub-argumento y backing. Por una parte, tomando en cuenta el principio de composicionalidad propuesto en [Vre97] (y como fue mencionado en la Sección 3.5.1), los sub-argumentos son considerados como necesarios para sus super-argumentos. Por otra parte, dado que una de las responsabilidades que un backing posee con respecto al argumento que soporta es establecer las circunstancias bajo las cuales la conexión entre sus premisas y conclusión (*i. e.*, su warrant asociado) es válida, podemos considerar que los argumentos de backing son necesarios para los argumentos que soportan.

Dado un BUAF en el que existen argumentos $\mathcal{A}, \mathcal{B}$, y $\mathcal{C}$ tales que $\mathcal{B}$ es un backing de $\mathcal{C}$ y $\mathcal{A}$ es un undercut para $\mathcal{C}$, la noción de derrota implícita caracterizada en la Definición 4.3 considera dos casos (no excluyentes): (*i*) $\mathcal{A}$ derrota implícitamente a $\mathcal{B}$ y (*ii*) $\mathcal{B}$ derrota implícitamente a $\mathcal{A}$. La Figura 4.13 ilustra el primer caso, en el que existe una derrota implícita de $\mathcal{A}$ sobre $\mathcal{B}$ denotada mediante una flecha sólida sin cola. En este caso, la derrota implícita se corresponde con la derrota mediada del d-BAF presentado en la Sección 3.2. Esta derrota tiene sentido en el contexto de un BUAF dado que, como la

$$\begin{tikzcd} \mathcal{A} \arrow[r, tail, "U_c"] & \mathcal{C} & \mathcal{B} \arrow[l, Rightarrow, "b"'] \arrow[ll, bend left] \end{tikzcd}$$

Figura 4.14: Segundo caso de derrota implícita en un BUAF.

Definición 4.2 indica, la determinación del éxito del undercut $\mathcal{A}$ sobre $\mathcal{C}$ involucra una comparación mediante el uso de preferencias entre los argumentos $\mathcal{A}$ y $\mathcal{B}$. Por lo tanto, es de esperarse que si $\mathcal{C}$ no está aceptado (por haber sido derrotado por $\mathcal{A}$) entonces su backing $\mathcal{B}$ tampoco lo esté, lo cual se corresponde con las restricciones impuestas por la interpretación deductiva de soporte.

El segundo caso de la Definición 4.3 es ilustrado por la Figura 4.14, donde la derrota implícita de $\mathcal{B}$ sobre $\mathcal{A}$ es denotada mediante una flecha sólida sin cola. A diferencia del primer caso de derrota implícita, ilustrado en la Figura 4.13, esta derrota no se corresponde con ninguna de las derrotas complejas identificadas por los formalismos presentados en el Capítulo 3. Esto se debe a que las derrotas implícitas son altamente dependientes de la interpretación de soporte del BUAF. Es decir, en un contexto en el que pueden existir undercuts y backings para un mismo argumento, las derrotas implícitas tienen por objetivo prevenir la aceptación conjunta de los argumentos de backing y undercutting.

A continuación analizaremos las derrotas complejas restantes propuestas por los formalismos de las secciones 3.1, 3.2 y 3.3. Concretamente, analizaremos el segundo tipo de derrota extendida del AFN y las derrotas soportadas de los BAF y d-BAF, dado que no se hallan directamente relacionadas con alguna de las derrotas complejas del BUAF.

Consideremos el segundo tipo de derrota extendida del AFN, la cual establece que si $\mathcal{B} \longrightarrow \mathcal{C}$ y $\mathcal{B} \overset{n}{\Longrightarrow} \mathcal{A}$, entonces $\mathcal{A} \longrightarrow \mathcal{C}$. Esta derrota no es explícitamente considerada por el BUAF. Sin embargo, como se vio en la Sección 4.4, el BUAF cuenta con una propiedad que establece que la aceptabilidad de un argumento implica la aceptabilidad de los argumentos que lo soportan (Proposición 4.3). Claramente, esta propiedad se mantiene al considerar cadenas de soporte. Esto se debe a que, como se mencionó anteriormente, la noción de derrota indirecta formalizada en la Definición 4.4 propaga las derrotas a través de cadenas de soporte. Por lo tanto, el comportamiento modelado por esta propiedad de los BUAF se corresponde con el comportamiento modelado por el segundo tipo de derrota extendida de los AFN, en el siguiente sentido: dados $\mathcal{B} \longrightarrow \mathcal{C}$ y $\mathcal{B} \overset{b}{\Longrightarrow} \mathcal{A}$ o $\mathcal{B} \overset{\sqsubseteq}{\Longrightarrow} \mathcal{A}$, si $\mathcal{A}$

está aceptado entonces $\mathcal{B}$ también estará aceptado y, por lo tanto, $\mathcal{C}$ no estará aceptado por hallarse derrotado por $\mathcal{B}$.

El análisis para el caso de las derrotas soportadas de los BAF y d-BAF es diferente. Una derrota soportada establece que si $\mathcal{B}$ soporta a $\mathcal{A}$ (respectivamente, $\mathcal{B} \overset{s}{\Longrightarrow} \mathcal{A}$ y $\mathcal{B} \overset{d}{\Longrightarrow} \mathcal{A}$) ya sea mediante un soporte directo o una cadena de soporte y $\mathcal{A} \longrightarrow \mathcal{C}$, entonces $\mathcal{B} \longrightarrow \mathcal{C}$. Las derrotas soportadas no son consideradas en el contexto de los BUAF. Esto es así porque, dadas las características de la relación de soporte del BUAF, si $\mathcal{B} \overset{b}{\Longrightarrow} \mathcal{A}$ o $\mathcal{B} \overset{\sqsubseteq}{\Longrightarrow} \mathcal{A}$ y $\mathcal{A} \longrightarrow \mathcal{C}$, el hecho de que el argumento $\mathcal{B}$ que origina el soporte esté aceptado no implica que $\mathcal{A}$ estará aceptado, habilitando la posibilidad de que $\mathcal{C}$ esté aceptado. Por ejemplo, si $\mathcal{B} \overset{b}{\Longrightarrow} \mathcal{A}$ y $\mathcal{B}$ está aceptado, puede ocurrir que $\mathcal{A}$ no esté aceptado dado que se halla derrotado por un argumento $\mathcal{D}$ que está aceptado y, en particular, efectuó un ataque de tipo rebutting sobre $\mathcal{A}$. De manera similar, si $\mathcal{B} \overset{\sqsubseteq}{\Longrightarrow} \mathcal{A}$ y $\mathcal{B}$ está aceptado, puede suceder que $\mathcal{A}$ no esté aceptado por ser objeto de una derrota producida por un argumento $\mathcal{D}$ que se halla aceptado y, en particular, ataca un sub-argumento de $\mathcal{A}$ diferente de $\mathcal{B}$. En cualquiera de estos dos casos, la derrota de $\mathcal{D}$ sobre $\mathcal{A}$ produce la reinstalación de $\mathcal{C}$, conduciéndolo a un estado de aceptación.

Con respecto a las derrotas mediadas del d-BAF, debemos tener en cuenta lo siguiente. Si bien el primer caso de derrota implícita del BUAF se corresponde con la derrota mediada del d-BAF, esta correspondencia no es total. Esto se debe a que la derrota implícita únicamente considera la situación de conflicto derivada de la coexistencia de argumentos de backing y undercutting. Es decir, cuando se consideren conflictos derivados de ataques de tipo rebutting o undermining y se utilice la relación de soporte como sub-argumento, no se obtendrán derrotas implícitas. Por lo tanto, por ejemplo, si tenemos un argumento $\mathcal{A}$ que es un derrotador de tipo rebutting para $\mathcal{C}$ y existe un backing $\mathcal{B}$ para $\mathcal{C}$, esta situación no conduce a una derrota implícita de $\mathcal{A}$ sobre $\mathcal{B}$. En particular, podemos notar que la diferencia entre la derrota mediada del d-BAF y la derrota indirecta del BUAF se desprende del hecho de que el d-BAF no distingue entre diferentes tipos de derrota y soporte. En contraste, el BUAF efectúa una distinción entre diferentes tipos de ataque (rebutting, undercutting y undermining) y la relación de soporte se encuentra definida de manera tal que modela el soporte provisto por los backings, así como también aquel provisto por los sub-argumentos.

El análisis efectuado muestra que el BUAF no provee los medios para representar las derrotas soportadas de los BAF y d-BAF. De manera similar, las derrotas mediadas del d-BAF son capturadas en el contexto de los BUAF únicamente cuando se consideran el soporte representado por los backings y los ataques de tipo undercutting. Sin embargo, se mostró que el BUAF captura (explícitamente o implícitamente) el comportamiento modelado por las derrotas extendidas del AFN. Por lo tanto, dado que, como se vio en la Sección 3.3, las interpretaciones de soporte por necesidad y soporte deductivo son duales, podemos considerar que el comportamiento modelado por las derrotas soportadas y las derrotas mediadas del d-BAF es capturado en el contexto del BUAF.

Podemos destacar otras diferencias existentes entre el BUAF y las aproximaciones presentadas en el Capítulo 3. Por una parte, la relación de soporte dl BUAF, al igual que en el d-BAF y el AFN, corresponde a una relación de soporte concreta. En contraste, la relación de soporte del BAF no adopta una interpretación específica, representando simplemente una interacción positiva entre argumentos. No obstante esto, existe una gran diferencia entre la relación de soporte del BUAF y aquellas correspondientes al AFN y al d-BAF: mientras que los últimos adoptan una única interpretación de soporte (respectivamente, soporte por necesidad y soporte deductivo), el BUAF emplea una relación de soporte que combina dos interpretaciones en un marco unificado: backing y sub-argumento.

Por otra parte, dado que la aproximación presentada en la Sección 3.4 no define derrotas complejas entre argumentos, no es posible efectuar una comparación directa entre el EAS y el BUAF. Sin embargo, las diferencias existentes entre el EAS y otras aproximaciones, señaladas en la Sección 3.4, son también aplicables en la comparación entre el EAS y el BUAF. Esto es, ante la ausencia de ataques o derrotas entre argumentos, todos los argumentos de un BUAF quedarán aceptados. En contraste, en un EAS únicamente estarán aceptados aquellos argumentos que se hallen soportados por evidencia. Además, mientras que las relaciones de ataque y soporte de un BUAF relacionan pares de argumentos, aquellas correspondientes a un EAS permiten expresar derrota y soporte originadas en un conjunto de argumentos.

Consideremos ahora las aproximaciones presentadas en la Sección 3.5. Como se mencionó en el Capítulo 3, los Marcos Argumentativos Abstractos con Sub-argumentos (AFS) y los Marcos Dialécticos Abstractos (ADF) se hallan relacionados con el estudio de soporte en argumentación. En particular, como se vio en la Sección 3.5.1, existe una correspondencia entre la relación de sub-argumento del AFS y la relación de soporte por necesidad del AFN, ya que los argumentos son *necesarios* para sus super-argumentos. Luego, podemos establecer una correspondencia entre la relación de soporte del BUAF y la relación de sub-argumento del AFS. Sin embargo, si bien el BUAF permite modelar la relación de sub-argumento del AFS, la relación de soporte del BUAF es más abarcativa, dado que también permite modelar el soporte provisto por los backings de Toulmin [Tou58].

Finalmente, análogamente a lo señalado en la Sección 3.5.2, a pesar de la flexibilidad asociada a las condiciones de aceptación de los ADF, estos no permiten capturar todas las restricciones de aceptabilidad impuestas por la relación de soporte de los BUAF. Por ejemplo, dado que las condiciones de aceptación para un nodo en un ADF se definen únicamente considerando sus nodos "padres", no es posible modelar las derrotas implícitas derivadas de la coexistencia de argumentos de backing y undercutting en los BUAF. Es decir, dados los argumentos $\mathcal{A}, \mathcal{B}$ y $\mathcal{C}$ tales que $\mathcal{A}$ es un undercut para $\mathcal{C}$ y $\mathcal{B}$ es un backing para $\mathcal{C}$, el BUAF captura el conflicto existente entre $\mathcal{A}$ y $\mathcal{B}$ mediante la noción de derrota implícita. En contraste, un ADF que modele esta situación poseerá únicamente vínculos entre los argumentos $\mathcal{A}$ y $\mathcal{C}$ (donde $\mathcal{A}$ es padre de $\mathcal{C}$), y entre los argumentos $\mathcal{B}$ y $\mathcal{C}$ (donde $\mathcal{B}$ es padre de $\mathcal{C}$). Luego, al definir las condiciones de aceptación de $\mathcal{A}$ o $\mathcal{B}$ no sería posible capturar el conflicto antes mencionado, dado que no existe vínculo alguno entre estos argumentos en el ADF.

La Figura 4.15 resume la relación existente entre las derrotas complejas definidas por el BUAF, y aquellas correspondientes a los formalismos presentados en el Capítulo 3. Esta relación es ilustrada mediante la identificación de las derrotas complejas que toman lugar, para cada formalismo, ante diferentes escenarios que combinan la coexistencia de derrota y soporte entre argumentos[4]. En particular, para cada caso, las derrotas complejas obtenidas por cada formalismo se hallan denotadas mediante flechas punteadas. Por último, cabe destacar que la Figura 4.15 incluye únicamente aquellos formalismos presentados en el Capítulo 3 que definen derrotas complejas entre los argumentos del sistema: BAF, d-BAF y AFN.

4.6. Conclusiones

En este capítulo se presentó un formalismo de argumentación abstracta llamado *Marco Argumentativo con Backing y Undercutting (BUAF)*. La propuesta aquí desarrollada fue inspirada en el trabajo de Toulmin [Tou58] y Pollock [Pol87]. Concretamente, el BUAF

[4]Los escenarios se hallan especificados en la primera fila de cada columna en la tabla.

	$\mathcal{A} \Longrightarrow \mathcal{B} \longrightarrow \mathcal{C}$	$\mathcal{A} \longrightarrow \mathcal{B} \Longrightarrow \mathcal{C}$	$\mathcal{A} \longrightarrow \mathcal{B} \Longleftarrow \mathcal{C}$	$\mathcal{A} \Longleftarrow \mathcal{B} \longrightarrow \mathcal{C}$
BUAF		$\mathcal{A} \longrightarrow \mathcal{B} \xRightarrow{b/\sqsubseteq} \mathcal{C}$ Derrota Indirecta	$\mathcal{A} \longrightarrow \mathcal{B} \xLeftarrow{b} \mathcal{C}$ Si $(\mathcal{A}, \mathcal{B} \in \mathbb{U}_c)$ y $\mathcal{A} \not\to \mathcal{C}$ Derrota Implícita (1° Caso) $\mathcal{A} \longrightarrow \mathcal{B} \xLeftarrow{b} \mathcal{C}$ Si $(\mathcal{A}, \mathcal{B} \in \mathbb{U}_c)$ y $\mathcal{C} \not\to \mathcal{A}$ Derrota Implícita (2° Caso)	Comportamiento capturado por la Proposición 4.3
BAF	$\mathcal{A} \xRightarrow{s} \mathcal{B} \longrightarrow \mathcal{C}$ Derrota Soportada	$\mathcal{A} \longrightarrow \mathcal{B} \xRightarrow{s} \mathcal{C}$ Derrota Secundaria		
d-BAF	$\mathcal{A} \xRightarrow{d} \mathcal{B} \longrightarrow \mathcal{C}$ Derrota Soportada		$\mathcal{A} \longrightarrow \mathcal{B} \xLeftarrow{d} \mathcal{C}$ Derrota Mediada	
AFN		$\mathcal{A} \longrightarrow \mathcal{B} \xRightarrow{n} \mathcal{C}$ Derrota Extendida (1° Caso)		$\mathcal{A} \xLeftarrow{n} \mathcal{B} \longrightarrow \mathcal{C}$ Derrota Extendida (2° Caso)

Figura 4.15: Relación entre derrotas complejas de los BUAF, BAF, d-BAF yAFN.

constituye una extensión al marco argumentativo de Dung (ver Sección 2.2.1) que permite modelar ataque y soporte para inferencias. Para lograr esto, el BUAF distingue entre diferentes tipos de ataque, así como también cuenta con una relación de soporte especializada que combina las nociones de backing y sub-argumento. De esta manera, el BUAF permite la representación de los backings de Toulmin y los undercuts de Pollock, dos nociones importantes en la comunidad de argumentación.

Si bien las nociones de backing y undercutting han sido abordadas con anterioridad por diferentes aproximaciones existentes en la literatura, ninguna de las propuestas de argumentación abstracta desarrolladas previamente las había considerado en forma conjunta. Por ejemplo, como se mencionó en la Sección 4.5, en [Pra09] y [BGvdTV10] se proveen formalizaciones de argumentación abstracta que permiten modelar la noción de undercutting defeater propuesta por Pollock. Por otra parte, en [Ver05] se propone una reconstrucción de las ideas de Toulmin que permite además modelar la noción de undercutting defeater propuesta por Pollock. En particular, dicho trabajo será analizado en la Sección 5.7 dado que la reconstrucción propuesta en [Ver05] es efectuada utilizando los conectivos lógicos de DefLog [Ver03a].

Al igual que el BUAF, muchas de las aproximaciones existentes en la literatura que proponen una relación de soporte entre argumentos adoptan una interpretación de soporte particular. Asimismo, todas estas aproximaciones definen un conjunto de derrotas complejas, las cuales refuerzan las restricciones de aceptabilidad impuestas por la correspondiente interpretación de soporte. Por lo tanto, dadas las diferencias existentes entre las interpretaciones de soporte adoptadas por cada uno de estos formalismos, las propuestas ya existentes en la literatura no capturan los conflictos implícitos que surgen en el BUAF a partir de la coexistencia de argumentos de backing y undercutting.

Como resultado del análisis efectuado en la Sección 4.5 podemos concluir que algunos aspectos de la relación de soporte del BUAF se corresponden con aspectos de otras interpretaciones de soporte propuestas en la literatura. En particular, se mostró que las restricciones impuestas por la relación de soporte por necesidad del AFN (análogamente, de la relación de soporte deductivo del d-BAF) son capturadas en el contexto del BUAF, dado que los argumentos de backing y los sub-argumentos son *necesarios* para los argumentos que estos soportan. Sin embargo, como se mencionó anteriormente, la relación de backing impone restricciones adicionales sobre la aceptabilidad de los argumentos, las cuales consideran de manera explícita la existencia de undercuts. Estas restricciones, como se vio en la Sección 4.3, se encuentran capturadas mediante la noción de derrota implícita entre argumentos de backing y undercutting. Por lo tanto, no es posible establecer una correspondencia total entre las relaciones de soporte por necesidad o deductivo y la relación de soporte del BUAF.

Por otra parte, se mostró que un BUAF puede ser mapeado a un AF cuya especificación se obtiene a partir del conjunto de argumentos del BUAF y la relación de derrota correspondiente a la Definición 4.5. De esta manera, es claro que los ejemplos y aplicaciones mostrados para el BUAF también son aplicables a los marcos argumentativos abstractos de Dung. Más aún, esta transformación permite a los BUAF heredar las propiedades demostradas para los AF, así como también aprovechar los resultados futuros que se elaboren sobre el formalismo de Dung.

Finalmente cabe destacar que, además de proveer un marco unificado para la representación de las nociones de backing y undercutting, el BUAF provee una herramienta específica e intuitiva para la representación de conocimiento y razonamiento en escenarios donde la información puede ser tanto atacada como soportada. En consecuencia, podemos considerar que el BUAF provee una herramienta de representación de más bajo nivel que el AF propuesto por Dung. Esto se debe a que, si bien ambos formalismos pertenecen a la categoría de marco argumentativo abstracto (porque se abstraen de la forma en que se construyen los argumentos), el BUAF abordan las relaciones entre argumentos con un mayor nivel de detalle que el AF.

A diferencia del AF, el BUAF considera una interacción positiva entre argumentos, la cual se halla caracterizada por la relación de soporte. Asimismo, en lugar de considerar una relación de derrota preestablecida, el BUAF especifica relaciones de ataque y preferencia que son utilizadas, en conjunto con la relación de soporte, para identificar las derrotas entre los argumentos del sistema. De esta manera, el BUAF se sitúa en un nivel intermedio entre el AF y sistemas argumentativos más concretos como los Sistemas Argumentativos Basados en Reglas (SABR). En particular, como se verá en el Capítulo 5, la especificación de un BUAF puede obtenerse a partir de los argumentos construidos utilizando el lenguaje de representación de un SABR y la identificación de las relaciones de ataque, soporte y preferencia entre ellos.

Capítulo 5

Programación en Lógica Rebatible Extendida (E-DeLP)

En este capítulo se desarrollará un sistema argumentativo basado en reglas llamado *Programación en Lógica Rebatible Extendida (E-DeLP)* (por su nombre en inglés, *Extended Defeasible Logic Programming*). Este formalismo extiende a la Programación en Lógica Rebatible [GS04], incorporando un conjunto de reglas de backing y undercutting. De esta manera, E-DeLP constituye un marco unificado para modelar ataque y soporte a inferencias en un sistema argumentativo basado en reglas, proveyendo así las bases para su implementación computacional.

De manera análoga a lo realizado en el Capítulo 4, el formalismo aquí propuesto permitirá modelar el soporte existente entre las nociones de backing y warrant propuestas por Toulmin [Tou58]. Asimismo, E-DeLP permitirá representar la noción de undercutting defeater propuesta por Pollock [Pol87].

5.1. Introducción

En las últimas décadas argumentación ha evolucionado como un atractivo paradigma para conceptualizar el razonamiento de sentido común [CML00, PV02, BCD07, BH08, RS09, AyC14], dando lugar a diferentes formalizaciones en la literatura. En particular, existen diferentes aproximaciones para modelar argumentación en entornos concretos, las cuales corresponden a Sistemas Argumentativos Basados en Reglas (SABR) (ver *e. g.*, [PS97, GS04, DKT06, AK07]).

Una crítica usualmente realizada sobre los SABR es que determinados patrones de razonamiento estudiados en áreas como el razonamiento legal y la filosofía, los cuales constituyen contribuciones importantes a la comunidad de argumentación, han sido simplificados o no considerados en la definición formal de los SABR desarrollados hasta el momento. Tal es el caso de las contribuciones realizadas por Toulmin [Tou58] y Pollock [Pol87]. En la Sección 2.3 se presentó el modelo para el diseño de argumentos propuesto por Toulmin, el cual identifica, además de *datos* y *conclusión*, elementos como *warrant*, *backing*, *derrotador* y *calificador*. Por otra parte, en [Pol87] Pollock propuso una clasificación de posibles derrotadores para un argumento, distinguiendo entre *rebutting defeaters* y *undercutting defeaters*. Mientras que los derrotadores del primer tipo surgen a partir de un ataque sobre la conclusión de un argumento, los últimos corresponden a un ataque sobre la conexión entre las premisas y conclusión del argumento, es decir, sobre la regla de inferencia asociada al mismo.

Continuando con los desarrollos presentados en el Capítulo 4, en este capítulo se abordarán las ideas de Toulmin y Pollock en el contexto de un SABR. En particular, en este capítulo se presentará un formalismo que extiende a la Programación en Lógica Rebatible (DeLP). Como se vio en la Sección 2.2.2, DeLP permite identificar argumentos cuyas conclusiones intermedias o finales se contradicen, capturando la noción de rebutting defeater propuesta por Pollock. Sin embargo, como se verá a continuación, en DeLP no es posible representar la noción de undercutting defeater de Pollock ni la noción de backing propuesta por Toulmin. Por lo tanto, en este capítulo se introducirá el formalismo de *Programación en Lógica Rebatible Extendida (E-DeLP)*, el cual permite expresar argumentos a favor y en contra de reglas de inferencia, capturando así las nociones de undercutting defeater y backing.

Backing y Undercutting en DeLP: analizando alternativas de representación

Cuando se comenzó con el desarrollo de un SABR que permita representar las nociones de backing y undercutting, se estudió la posibilidad de modelar estas nociones utilizando de las herramientas provistas por el formalismo de DeLP. Como resultado de este análisis, se concluyó que DeLP no posee los elementos de representación necesarios para modelar correctamente la noción de undercutting defeater propuesta por Pollock o la noción de backing propuesta por Toulmin. A continuación se presentará el análisis realizado ilustrando las diferentes alternativas de representación que fueron consideradas, así como también las limitaciones encontradas en cada caso.

Como punto de partida se consideró que los warrants de Toulmin podían ser representados mediante reglas rebatibles. Esto resulta razonable dado que en [Tou58] Toulmin identifica a los warrants como reglas generales que proveen el paso de inferencia entre las premisas y conclusión de un argumento. Luego, dado que los undercutting defeaters realizan ataques sobre reglas de inferencia, podemos considerarlos como razones en contra del uso de reglas rebatibles. De manera similar, dado que los backings de Toulmin proveen soporte para los warrants, estos pueden ser considerados como razones a favor del uso de reglas rebatibles.

Una alternativa para expresar soporte o ataque a una regla rebatible r : "$Cabeza \prec Cuerpo$" podría consistir en ubicar la regla rebatible r en la cabeza de otra regla cuyo cuerpo expresa, respectivamente, las razones a favor o en contra del uso de r. Sin embargo, como se vio en la Sección 2.2.2, el lenguaje de representación de DeLP admite únicamente el uso de un literal en la cabeza de una regla rebatible. Por lo tanto, para implementar esta alternativa sería necesario extender el formalismo de DeLP.

Otra posibilidad consistiría en asociar un literal especial a cada regla rebatible, correspondiente a una *etiqueta*. De esta manera, sería posible identificar a cada regla mediante un único literal, y expresar ataque o soporte a la regla mediante un ataque o soporte a su etiqueta. Sin embargo, como el lenguaje de DeLP no considera el uso de etiquetas, esta alternativa también implicaría la extensión del formalismo de DeLP. No obstante esto, a continuación consideraremos la posibilidad de simular etiquetas en DeLP, de manera tal que no se requiera la extensión de su lenguaje de representación.

La simulación de etiquetas asociadas a las reglas rebatibles de DeLP podría lograrse mediante la inclusión de un literal adicional en el cuerpo de las reglas. Por ejemplo, dada la regla rebatible r arriba mencionada, la regla resultante de añadir un literal especial en el cuerpo sería "$Cabeza \prec L_1, L_2, \ldots, L_n, r$", donde $Cuerpo = \{L_1, L_2, \ldots, L_n\}$ y el literal "r" intenta simular la etiqueta de la regla. Teniendo en cuenta esto, supongamos ahora que queremos representar que "a" provee razones rebatibles en contra del uso de esta nueva regla, así como también que "b" provee razones rebatibles en contra de"L_i" $(1 \leq i \leq n)$. Esto daría lugar a dos ataques originados, respectivamente, por las reglas rebatibles "$\sim r \prec a$" y "$\sim L_i \prec b$". Sin embargo, como se mostrará a continuación, esta representación conduce a una serie de problemas.

Por una parte, en DeLP no se efectúa distinción alguna entre los literales pertenecientes al cuerpo de una regla rebatible. Por lo tanto, no sería posible identificar el literal correspondiente a la "etiqueta simulada" de una regla rebatible. Más aún, teniendo en cuenta esto, no sería posible distinguir entre un ataque sobre un literal ordinario y un ataque sobre la regla (el cual se halla representado mediante un ataque sobre el literal que simula la etiqueta). De esta manera, por ejemplo, dada la regla rebatible "$Cabeza \prec L_1, L_2, \ldots, L_n, r$", DeLP no sería capaz de distinguir entre la naturaleza de los ataques originados por las reglas "$\sim r \prec a$" y "$\sim L_i \prec b$".

Por otra parte, si las etiquetas son asignadas de manera tal que permitan identificar unívocamente a las reglas rebatibles (*i. e.*, no existen dos reglas con la misma etiqueta), la inclusión de la etiqueta simulada en el cuerpo de las reglas podría obstruir el proceso de comparación de argumentos efectuado por DeLP. En particular, ya no sería posible utilizar *especificidad generalizada* (ver Sección 2.2.2.5) como criterio de comparación, dado que toda regla rebatible poseerá un literal (la etiqueta simulada) que no aparece en el cuerpo de ninguna otra regla rebatible.

Los problemas arriba mencionados muestran claramente que la simulación de etiquetas en DeLP no es posible. Por lo tanto, si queremos proveer un mecanismo para expresar ataque y soporte a reglas rebatibles mediante alguna de las alternativas propuestas, será necesario extender el lenguaje de representación de DeLP. En particular, en este capítulo se adoptará la primera alternativa presentada, mediante la incorporación de dos tipos de reglas al lenguaje de representación: reglas de backing y reglas de undercutting. De esta manera, el formalismo propuesto en este capítulo brindará los medios para modelar las ideas de Toulmin y Pollock en el contexto de la programación en lógica rebatible.

El resto de este capítulo se encuentra organizado de la siguiente manera: en la Sección 5.2 se introducirá el lenguaje de representación de E-DeLP, el cual incorpora las reglas de backing y undercutting, y se brindará una especificación para los programas E-DeLP. Partiendo de esta especificación, en la Sección 5.3 se presentará el mecanismo para la construcción de argumentos, identificando diferentes tipos de argumento. Seguidamente, en la Sección 5.4 se determinarán las situaciones que originan conflictos entre argumentos, distinguiendo entre tres tipos de ataque. Dados los argumentos construidos a partir de un programa E-DeLP, los ataques identificados y el criterio de comparación de argumentos, en la Sección 5.5 se proveerá la especificación de un Marco Argumentativo con Backing y Undercutting (BUAF) asociado al programa E-DeLP, sobre el cual se computarán las derrotas entre argumentos. Una vez determinadas las derrotas entre los argumentos de E-DeLP mediante el BUAF asociado, en la Sección 5.6 se procederá a efectuar el cálculo de aceptabilidad de los argumentos. Para tal fin, se presentarán dos alternativas: siguiendo la metodología empleada en el Capítulo 4, la Sección 5.6.1 considerará el uso de semánticas de aceptabilidad sobre el BUAF asociado al programa E-DeLP. Por otra parte, la Sección 5.6.2 propondrá una alternativa al cálculo de aceptabilidad utilizando árboles de dialéctica, tomando como punto de partida el conjunto de argumentos de E-DeLP y la relación de derrota obtenida a partir del BUAF asociado. Finalmente, una vez identificado el conjunto de argumentos aceptados, se definirá la noción de garantía para los literales de un programa E-DeLP, la cual permitirá determinar las inferencias del sistema.

5.2. Representación de Conocimiento en E-DeLP: lenguaje de representación extendido

En esta sección se introducirá la sintaxis de la *Programación en Lógica Rebatible Extendida (E-DeLP)*, tomando como base el lenguaje de representación de DeLP presentado en la Sección 2.2.2. Como se mencionó anteriormente, el lenguaje de E-DeLP permitirá modelar las nociones de undercutting defeater y backing propuestas por Pollock y Toulmin.

El lenguaje de representación de E-DeLP se define en términos de cinco conjuntos disjuntos: un conjunto de *hechos*, un conjunto de *reglas estrictas*, un conjunto de *reglas rebatibles*, un conjunto de *reglas de backing* y un conjunto de *reglas de undercutting*. Al igual que en DeLP, para definir la sintaxis de E-DeLP se emplearán los conceptos preliminares introducidos en la Sección 2.2.2.1. De esta manera, un literal "'L" será un átomo "A" o un átomo negado "$\sim A$", donde "$\sim$" representa la negación fuerte, y podrá ser utilizado como un hecho o como parte de una regla de un programa E-DeLP.

La representación de hechos, reglas estrictas y reglas rebatibles en E-DeLP es la misma que en DeLP. No obstante esto, para facilitar la lectura de las subsiguientes definiciones, a continuación se incluyen las definiciones correspondientes a estas nociones, introducidas previamente en la Sección 2.2.2.2.

Definición 5.1 (Hecho) *Un* hecho *es un literal fijo L, es decir, un átomo fijo o un átomo fijo negado.*

Definición 5.2 (Regla Estricta) *Una* regla estricta *es un par ordenado, denotado "Cabeza $\leftarrow$ Cuerpo", donde el primer elemento (Cabeza) es un literal fijo y el segundo elemento (Cuerpo) es un conjunto finito y no vacío de literales fijos.*
Una regla estricta con cabeza L_0 y cuerpo $\{L_1, \ldots, L_n\}$ $(n > 0)$ se escribirá también como $L_0 \leftarrow L_1, \ldots, L_n$.

Definición 5.3 (Regla Rebatible) *Una* regla rebatible *es un par ordenado, denotado "Cabeza $\prec$ Cuerpo", donde el primer elemento (Cabeza) es un literal fijo y el segundo elemento (Cuerpo) es un conjunto finito de literales fijos.*
Una regla rebatible con cabeza L_0 y cuerpo $\{L_1, \ldots, L_n\}$ $(n \geq 0)$ se escribirá también como $L_0 \prec L_1, \ldots, L_n$.

Los hechos, reglas estrictas y reglas rebatibles en E-DeLP poseen el mismo significado que en DeLP. Es decir, los hechos y reglas estrictas expresan información segura e indiscutible, mientras que las reglas rebatibles expresan información tentativa. Además, recordemos que una regla rebatible con cuerpo vacío se denota como "$Cabeza \prec$ " y recibe el nombre de *presuposición* [Nut88]. De esta manera, una presuposición "$P \prec$ " expresa que "*existen razones tentativas para creer en P*".

Los elementos incorporados por E-DeLP al lenguaje de representación son las *reglas de backing y undercutting* que, respectivamente, expresan soporte y ataque para reglas rebatibles. Así, la consideración de estos nuevos tipos de reglas permitirá argumentar acerca del uso de reglas rebatibles.

Definición 5.4 (Regla de Backing) *Una* regla de backing *es un par ordenado, denotado "$[Cabeza] \leftarrow\!\oplus\, [Cuerpo]$", donde el primer elemento (Cabeza) es una regla rebatible y el segundo elemento (Cuerpo) es un conjunto finito y no vacío de literales fijos.*
Una regla de backing con cabeza $R_{Cabeza} \prec R_{Cuerpo}$ y cuerpo $\{L_1, \ldots, L_n\}$ $(n > 0)$ se escribirá también como $[R_{Cabeza} \prec R_{Cuerpo}] \leftarrow\!\oplus\, [L_1, L_2, \ldots, L_n]$.

Definición 5.5 (Regla de Undercutting) *Una* regla de undercutting *es un par ordenado, denotado "$[Cabeza] \leftarrow\!\otimes\, [Cuerpo]$", donde el primer elemento (Cabeza) es una regla*

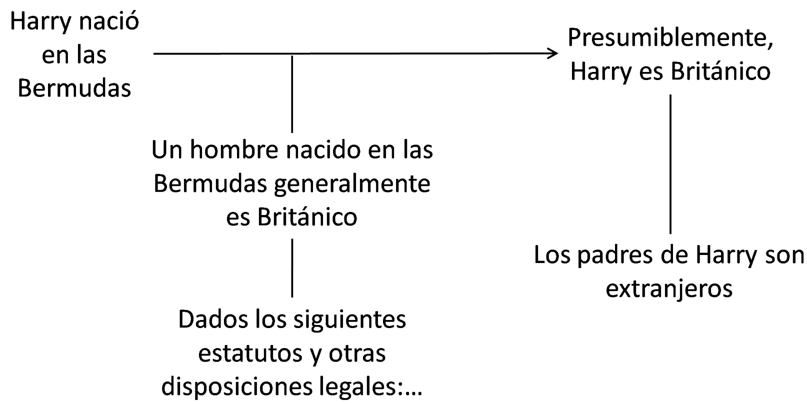


Figura 5.1: Ejemplo propuesto por Toulmin en [Tou58].

rebatible y el segundo elemento (Cuerpo) es un conjunto finito y no vacío de literales fijos. Una regla de undercutting con cabeza $R_{Cabeza} \prec R_{Cuerpo}$ *y cuerpo* $\{L_1, \ldots, L_n\}$ $(n > 0)$ *se escribirá también como* $[R_{Cabeza} \prec R_{Cuerpo}] \leftarrow\!\otimes [L_1, L_2, \ldots, L_n]$.

Sintácticamente, la única diferencia entre las reglas de backing y las reglas de undercutting es el uso de los símbolos '$\oplus$' y '$\otimes$'. Sin embargo, la semántica asociada a estos tipos de reglas es totalmente opuesta. Una regla de backing "$[Cabeza] \leftarrow\!\oplus [Cuerpo]$" expresa que "*razones para creer en Cuerpo proveen razones a favor del uso de la regla rebatible* $Cabeza$". En particular, consideraremos que una regla de backing establece las condiciones bajo las cuales el uso de la regla rebatible en cuestión es sensato. En otras palabras, una regla de backing establece condiciones bajo las cuales la regla rebatible que soporta sería aplicable. Por otra parte, una regla de undercutting "$[Cabeza] \leftarrow\!\otimes [Cuerpo]$" expresa que "*razones para creer en Cuerpo proveen razones en contra del uso de la regla rebatible* $Cabeza$". Para ilustrar estas nociones consideremos el siguiente ejemplo.

Ejemplo 5.1 *La Figura 5.1 ilustra un ejemplo propuesto por Toulmin, introducido previamente en la Sección 2.3, el cual versa acerca de la nacionalidad de un sujeto llamado Harry. Dada la formulación propuesta, el ejemplo instancia los elementos del modelo de Toulmin de la siguiente manera. Los datos se hallan representados por la información que indica que Harry nació en las Bermudas. Luego, la conclusión calificada del argumento establece que, presumiblemente, Harry es Británico. La conexión entre los datos y la conclusión del argumento se halla determinada por el warrant, correspondiente a la regla de* *inferencia que establece que un hombre nacido en las Bermudas generalmente es Británico. El warrant se halla soportado por el backing que enuncia la existencia de estatutos y otras disposiciones legales que lo avalan. Finalmente, el hecho de que los padres de Harry son extranjeros provee una condición de excepción para el argumento, dando origen a una derrota sobre el mismo. En particular, la derrota no se produce mediante un ataque sobre la conclusión, sino que el derrotador provee una condición de excepción para la regla de inferencia modelada por el warrant. Por lo tanto, el hecho de que los padres de Harry son extranjeros implica la existencia de un ataque de tipo undercutting sobre el argumento.*

Supongamos que deseamos modelar la situación descripta por el ejemplo de Toulmin utilizando el lenguaje de representación de E-DeLP. Los datos pueden representarse utilizando el literal "harry_nacio_bermudas", mientras que la conclusión se halla identificada mediante el literal "harry_britanico". Como se mencionó en la Sección 5.1, dado que el warrant constituye una regla de inferencia que provee la conexión entre los datos y la conclusión de un argumento, el mismo puede ser representado a través de una regla rebatible. De esta manera, el warrant en el ejemplo de Toulmin puede ser representado por la regla rebatible "harry_britanico —< harry_nacio_bermudas". En particular, nótese que el calificador presumiblemente *asociado a la conclusión no se incluye explícitamente, sino que se halla implícito en la naturaleza rebatible de la regla que modela el warrant. Por otra parte, el backing que establece la existencia de estatutos y otras disposiciones legales puede representarse mediante el literal "actas_parlamento_britanico". Luego, es posible modelar el soporte que el backing provee al warrant mediante la regla de backing "[harry_britanico —< harry_nacio_bermudas] ←⊕ [actas_parlamento_britanico]". Finalmente, el hecho de que los padres de Harry son extranjeros es representado mediante el literal "padres_harry_extranjeros". Esta información provee una condición de excepción para el argumento y, en particular, para la regla de inferencia expresada por el warrant. Por lo tanto, el hecho de que los padres de Harry son extranjeros provee una razón en contra del uso de la regla rebatible que modela el warrant, situación que se halla representada mediante la regla de undercutting "[harry_britanico —< harry_nacio_bermudas] ←⊗ [padres_harry_extranjeros]".*

Las definiciones 5.4 y 5.5 imponen una restricción sobre las reglas que pueden ser utilizadas en la cabeza de las reglas de backing y undercutting. En particular, la cabeza de este tipo de reglas sólo puede ser una regla rebatible y, por lo tanto, no es posible proveer soporte o ataque para reglas estrictas. Esta restricción resulta necesaria, dado que las reglas estrictas modelan información segura, proveyendo una conexión incondicional entre su antecedente (cuerpo) y consecuente (cabeza). Por otra parte nótese que, dado que las presuposiciones son un caso particular de reglas rebatibles, estas sí pueden ser utilizadas en la cabeza de reglas de backing o undercutting.

Las definiciones 5.1 a 5.5 describen los elementos utilizados por el lenguaje de representación de E-DeLP. En consecuencia, los *programas lógicos rebatibles extendidos* se hallan caracterizados por la siguiente definición.

Definición 5.6 (Programa Lógico Rebatible Extendido) *Un* programa lógico rebatible extendido *es un conjunto* $\mathcal{P}$ *de hechos, reglas estrictas, reglas rebatibles, reglas de backing y reglas de undercutting. Cuando resulte conveniente, un programa lógico rebatible extendido* $\mathcal{P}$ *será denotado como* $\mathcal{P} = (\Pi, \Delta, \Sigma)$*, distinguiendo el conjunto* Π *de hechos y reglas estrictas, el conjunto* Δ *de reglas rebatibles, y el conjunto* Σ *de reglas de backing y undercutting.*

Ejemplo 5.2 *Consideremos el escenario descripto por Toulmin, el cual fue presentado en el Ejemplo 5.1. El programa lógico rebatible extendido* $\mathcal{P}_{5.2} = (\Pi_{5.2}, \Delta_{5.2}, \Sigma_{5.2})$ *provee una formalización para dicho escenario en E-DeLP, donde:*

$\Pi_{5.2} = \{harry\_nacio\_bermudas,\ actas\_parlamento\_britanico,\ padres\_harry\_extranjeros\}$

$$\Delta_{5.2} = \{harry\_britanico \prec harry\_nacio\_bermudas\}$$

$$\Sigma_{5.2} = \left\{ \begin{array}{l} [harry\_britanico \prec harry\_nacio\_bermudas] \longleftarrow\!\oplus\ [actas\_parlamento\_britanico] \\ [harry\_britanico \prec harry\_nacio\_bermudas] \longleftarrow\!\otimes\ [padres\_harry\_extranjeros] \end{array} \right\}$$

El dato "harry_nacio_bermudas", el backing "'actas_parlamento_britanico" y el derrotador "padres_harry_extranjeros" son representados como hechos. El warrant se halla representado por la regla rebatible "$harry\_britanico \prec harry\_nacio\_bermudas$*", dado que razones para creer que Harry nació en las Bermudas proveen razones para creer que es Británico. El calificador "persumiblemente" se considera implícito en la naturaleza rebatible de la regla que modela el warrant. La regla de backing "*$[harry\_britanico \prec harry\_nacio\_bermudas] \longleftarrow\!\oplus\ [actas\_parlamento\_britanico]$*" expresa que el backing "actas_parlamento_britanico" provee el soporte correspondiente para el warrant. Finalmente, la regla de undercutting* *"[harry_britanico —< harry_nacio_bermudas] ←⊗ [padres_harry_extranjeros]" expresa que el derrotador "padres_harry_extranjeros" provee una condición de excepción para el warrant.*

Es importante remarcar que la existencia de reglas de backing y undercutting para una regla rebatible de un programa E-DeLP no es obligatoria. En particular, una regla rebatible para la cual no existen reglas de backing puede ser considerada como aplicable dada la ausencia de condiciones explícitas para su uso. Para ilustrar esta intuición, analicemos la naturaleza de los warrants de acuerdo a lo propuesto por Toulmin. En [Tou58] Toulmin enunció que pueden existir warrants de diferente tipo, los cuales confieren diferentes grados de fuerza a la conexión existente entre datos y conclusión de un argumento. Por una parte, hay warrants que conducen a la aceptación de una conclusión incondicionalmente, en cuyo caso la conexión provista por el warrant es indisputable. Dadas las características de este tipo de warrants, estos son representados en E-DeLP mediante reglas estrictas. Por otra parte, existen warrants que permiten arribar una conclusión tentativamente, los cuales son modelados a través de reglas rebatibles. En particular, un warrant de este tipo es el correspondiente al ejemplo de Toulmin, ilustrado en los ejemplos 5.1 y 5.2, dado que permite arribar a la conclusión *presumiblemente*.

De acuerdo a lo establecido por Toulmin, para aquellos warrants que pueden ser cuestionados no siempre se indicará algún backing. Por lo tanto, en tales casos, se tendrán reglas rebatibles (correspondientes a los warrants) para las cuales no existen reglas de backing. Por último, la presencia o ausencia de reglas de undercutting para una regla rebatible R depende de la existencia de condiciones de excepción para el warrant modelado por R. En conclusión, cada regla rebatible de un programa lógico rebatible extendido $\mathcal{P}$ poseerá cero o más reglas de backing y/o undercutting. El siguiente ejemplo, introducido por Pollock en [Pol87], ilustra un escenario en el que no se especifican backings, motivo por el cual su formalización en E-DeLP resulta en un programa lógico rebatible extendido sin reglas de backing.

Ejemplo 5.3 *Supongamos que vemos un objeto que luce de color rojo ("luce_rojo"), motivo por el cual existen razones para creer que es rojo ("es_rojo"). Sin embargo, si consideramos que la iluminación mediante una luz roja puede hacer que un objeto se vea rojo incluso cuando no lo es, y vemos que el objeto que luce de color rojo está siendo iluminado por una luz roja, entonces deberíamos dejar de concluir que es rojo (aunque* *podría, en efecto, serlo). De esta manera, el hecho de que el objeto está siendo iluminado por una luz roja ("iluminado_luz_roja") constituye un undercut para la inferencia dado que no representa una razón en contra de que el objeto sea rojo, sino que provee una razón en contra de la regla de inferencia que establece que el objeto es rojo porque luce de color rojo.*

Una posible representación para la situación arriba descripta se encuentra dada por el programa lógico rebatible extendido $\mathcal{P}_{5.3} = (\Pi_{5.3}, \Delta_{5.3}, \Sigma_{5.3})$*, donde:*

$$\Pi_{5.2} = \{luce\_rojo,\ iluminado\_luz\_roja\}$$
$$\Delta_{5.2} = \{es\_rojo \prec luce\_rojo\}$$
$$\Sigma_{5.2} = \{[es\_rojo \prec luce\_rojo] \leftarrow\!\otimes [iluminado\_luz\_roja]\}$$

En particular, el conjunto de hechos de $\mathcal{P}_{5.3}$ *modela la información conocida. Por otra parte, la regla rebatible "*$es\_rojo \prec luce\_rojo$*" expresa que si el objeto luce de color rojo, entonces existen razones para creer que es rojo. Finalmente, la regla de undercutting "*$[es\_rojo \prec luce\_rojo] \leftarrow\!\otimes [iluminado\_luz\_roja]$*" expresa la condición de excepción para la regla rebatible, la cual se basa en el hecho de que una luz roja puede hacer que un objeto se vea de color rojo incluso cuando no lo es.*

Las reglas de un programa E-DeLP son fijas, es decir, contienen literales fijos (sin variables). Sin embargo, al igual que en DeLP, se permitirá el uso de *esquemas de reglas* (estrictas, rebatibles, de backing y de undercutting) que contienen variables, las cuales serán denotadas con una letra mayúscula inicial. Siguiendo la convención usualmente adoptada en la literatura [Lif96], para cada esquema de reglas se definirá el conjunto de todas sus instancias fijas. En particular, dado un esquema de regla R, se define $Ground(R)$ como el conjunto de todas las instancias fijas de R. Es decir, el conjunto $Ground(R)$ agrupa todas las versiones de R reemplazando las variables por términos fijos, asumiendo que las variables con el mismo nombre corresponden al mismo elemento dentro de la regla. Luego, para un programa lógico rebatible extendido $\mathcal{P}$ se define el conjunto $Ground(\mathcal{P})$, el cual es la unión de los conjuntos $Ground(R)$ tales que R es un esquema de regla en $\mathcal{P}$. De esta manera, cuando se utilice el conocimiento expresado por el programa $\mathcal{P}$, se estará considerando, implícitamente, el conjunto $Ground(\mathcal{P})$.

El contar con esquemas de reglas en E-DeLP permite proveer una mejor representación para situaciones que podrían ser modeladas sin utilizar esquemas de reglas. Para ilustrar esto consideremos la definición de warrant propuesta por Toulmin, la cual expresa que los warrants son

> "*sentencias generales, hipotéticas, que actúan como puentes* [entre datos y conclusión] *y autorizan el paso de inferencia al cual se compromete el argumento*" [Tou58, pág. 98]

Por lo tanto, si consideramos el ejemplo propuesto por Toulmin que versa acerca de la nacionalidad de Harry, el warrant que expresa "*un hombre nacido en las Bermudas generalmente es Británico*" debería ser aplicable no sólo a Harry, sino a *cualquier* hombre nacido en las Bermudas. De esta manera, la naturaleza general asociada a los warrants de Toulmin puede ser capturada en E-DeLP mediante el uso de esquemas de reglas. Para ilustrar esto, consideremos el siguiente ejemplo.

Ejemplo 5.4 *Sea* $\mathcal{P}_{5.4} = (\Pi_{5.4}, \Delta_{5.4}, \Sigma_{5.4})$ *un programa lógico rebatible extendido que modela la situación descripta por el ejemplo de Toulmin (ver Figura 5.1) utilizando esquemas de reglas, donde:*

$$\Pi_{5.4} = \left\{ \begin{array}{l} nacio\_bermudas(harry) \\ actas\_parlamento\_britanico \\ padres\_extranjeros(harry) \end{array} \right\}$$

$$\Delta_{5.4} = \{es\_britanico(X) \prec nacio\_bermudas(X)\}$$

$$\Sigma_{5.4} = \left\{ \begin{array}{l} [es\_britanico(X) \prec nacio\_bermudas(X)] \leftarrow\!\oplus\, [actas\_parlamento\_britanico] \\ [es\_britanico(X) \prec nacio\_bermudas(X)] \leftarrow\!\otimes\, [padres\_extranjeros(X)] \end{array} \right\}$$

En este caso, el esquema de regla rebatible captura la naturaleza general del warrant, siendo entonces una regla de inferencia que se aplica para todo sujeto nacido en las Bermudas. De manera similar, las reglas esquemáticas de backing y undercutting reflejan razones a favor o en contra del esquema de regla rebatible. Por una parte, la regla de backing expresa una razón a favor del uso de cualquier instancia fija de la regla esquemática que modela el warrant. Por otra parte, la regla esquemática de backing establece que si una persona (por ejemplo, Harry) posee padres extranjeros, entonces existe una razón en contra del uso de la regla rebatible que modela el warrant para esa persona en particular. Por ejemplo, si consideramos el caso de Harry, la *regla rebatible "es_britanico(harry)* $\prec$ *nacio_Bermudas(harry)" corresponde a una instancia fija de la regla rebatible esquemática "es_britanico(X)* $\prec$ *nacio_bermudas(X)", donde la variable "X" se halla instanciada con el literal "harry". En consecuencia, esta regla rebatible expresa que si Harry nació en las Bermudas, entonces existen razones para creer que es Británico. De manera similar, la regla de undercutting "[es_britanico(harry)* $\prec$ *nacio_bermudas(harry)]* $\leftarrow\!\otimes$ *[padres_extranjeros(harry)]", correspondiente a una instancia fija de la regla de undercutting esquemática del conjunto* $\Sigma_{5.4}$*, expresa que si los padres de Harry son extranjeros, entonces existen razones en contra de la regla rebatible que establece que Harry es Británico.*

5.3. Construcción de Argumentos en E-DeLP

En esta sección se presentarán los elementos necesarios para la definición de la noción de *argumento*, correspondiente al segundo elemento de la estructura conceptual presentada en la Sección 2.1. En particular, a partir de un programa lógico rebatible extendido $\mathcal{P}$ será posible construir argumentos de diferente tipo: *argumentos conclusivos*, *argumentos de backing* y *argumentos de undercutting*. Por lo tanto, resulta necesario definir una noción de derivación que permita la construcción de los diferentes tipos de argumento arriba mencionados. Este mecanismo será llamado *derivación rebatible extendida* y se halla definido de la siguiente manera.

Definición 5.7 (Derivación Rebatible Extendida para un Literal) *Sea* $\mathcal{P} = (\Pi, \Delta, \Sigma)$ *un programa lógico rebatible extendido y* L *un literal. Una* derivación rebatible extendida *para* L *a partir de* $\mathcal{P}$ *es una secuencia finita de literales fijos* $L_1, L_2, \ldots, L_n = L$ *tal que cada literal* L_i *pertenece a la secuencia porque:*

(*a*) L_i *es un hecho en* Π*;*

(*b*) *existe en* Π *una regla estricta con cabeza* L_i *y cuerpo* $B_1, \ldots, B_k$*, donde todo literal* B_j *del cuerpo* $(1 \leq j \leq k)$ *es un elemento de la secuencia que precede a* L_i*; o*

(*c*) *existe en* Δ *una regla rebatible* R *con cabeza* L_i *y cuerpo* $B_1, \ldots, B_k$*, donde todo literal* B_j *del cuerpo* $(0 \leq j \leq k)$ *es un elemento de la secuencia que precede a* L_i *y se satisface una de las siguientes condiciones:*

i. no existe en Σ *una regla de backing con cabeza* R*, o*

ii. existe en Σ *una regla de backing con cabeza* R *y cuerpo* $S_1, \ldots, S_m$*, donde todo literal* S_p *del cuerpo* $(1 \leq p \leq m)$ *es un elemento de la secuencia que precede a* L_i*.*

La condición (c) de la Definición 5.7 expresa que si existen requerimientos para el uso de una regla rebatible R, estos deben ser satisfechos. Dada la naturaleza de los backings, una regla de backing para la regla rebatible R establece condiciones bajo las cuales el warrant modelado por la regla R es aplicable. Nótese que, dadas las características de los programas lógicos rebatibles extendidos, puede existir más de una regla de backing para R. Sin embargo, cada regla de backing para R expresa condiciones *alternativas* bajo las cuales el warrant modelado por R es aplicable. Por lo tanto, como expresa la condición $(c)ii.$ de la Definición 5.7, basta con tener una derivación rebatible para los literales que conforman el cuerpo de *una* regla de backing para R; es decir, es suficiente derivar los literales correspondientes a una de las condiciones alternativas bajo las cuales la regla R es aplicable. Para ilustrar esto consideremos el siguiente ejemplo.

Ejemplo 5.5 *Sea* $\mathcal{P}_{5.5} = (\Pi_{5.5}, \Delta_{5.5}, \Sigma_{5.5})$ *un programa lógico rebatible extendido, donde:*

$$\Pi_{5.5} = \left\{ \begin{array}{ll} noche & llave\_encendida(l) \\ electricidad & lampara\_rota(l) \\ lampara\_en\_habitac(l,h) & \end{array} \right\}$$

$$\Delta_{5.5} = \left\{ \begin{array}{l} luz\_encendida(X) \prec llave\_encendida(X) \\ habitac\_ilumninada(X) \prec dia \\ \sim habitac\_iluminada(X) \prec noche \\ habitac\_iluminada(X) \prec noche,\ lampara\_en\_habitac(Y,X), \\ \qquad\qquad\qquad\qquad\qquad luz\_encendida(Y) \end{array} \right\}$$

$$\Sigma_{5.5} = \left\{ \begin{array}{l} [luz\_encendida(X) \prec llave\_encendida(X)] \leftarrow\!\oplus [electricidad] \\ [luz\_encendida(X) \prec llave\_encendida(X)] \leftarrow\!\otimes [\sim electricidad] \\ [luz\_encendida(X) \prec llave\_encendida(X)] \leftarrow\!\oplus [\sim electricidad, \\ \qquad\qquad\qquad\qquad\qquad\qquad\qquad lampara\_emergencia(X)] \\ [luz\_encendida(X) \prec llave\_encendida(X)] \leftarrow\!\otimes [electricidad, \\ \qquad\qquad\qquad\qquad\qquad\qquad\qquad lampara\_rota(X)] \end{array} \right\}$$

El programa $\mathcal{P}_{5.5}$ *contiene información acerca del encendido de una lámpara y su correspondiente luz, con el objetivo de determinar si una habitación se hallará iluminada o no. En particular, la regla rebatible "*$luz\_encendida(X) \prec llave\_encendida(X)$*" modela la regla general que establece que si la llave de una lámpara* X *se halla encendida, entonces hay razones para creer que la luz de la lámpara* X *estará encendida. Luego, las reglas de backing y undercutting para esta regla rebatible expresan, respectivamente, razones a favor y en contra para su uso. Por una parte, la regla de backing "*$[luz\_encendida(X) \prec llave\_encendida(X)] \leftarrow\!\oplus\, [electricidad]$*" expresa que, para que la regla general referente al encendido de una lámpara* X *valga, debe haber electricidad. En consecuencia, la regla de undercutting "*$[luz\_encendida(X) \prec llave\_encendida(X)] \leftarrow\!\otimes\, [\sim electricidad]$*" expresa que la ausencia de electricidad es una razón bajo la cual no sería viable utilizar la regla rebatible arriba mencionada. Sin embargo, la regla de backing "*$[luz\_encendida(X) \prec llave\_encendida(X)] \leftarrow\!\oplus\, [\sim electricidad, lampara\_emergencia(X)]$*" expresa que, ante la ausencia de electricidad, la conexión modelada por la regla rebatible en cuestión es válida para lámparas de emergencia. Finalmente, la regla de undercutting "*$[luz\_encendida(X) \prec llave\_encendida(X)] \leftarrow\!\otimes\, [electricidad,\ lampara\_rota(X)]$*" expresa que por más que haya electricidad, si la lámpara* X *se encuentra rota, entonces la conexión expresada por la regla rebatible en cuestión no es válida.*

Dada la situación arriba descripta, las reglas de backing para la regla rebatible "$luz\_encendida(X) \prec llave\_encendida(X)$*" proveen condiciones alternativas bajo las cuales esta resulta aplicable. Es decir, para utilizar la regla rebatible deberá ocurrir que haya electricidad, o bien no haya electricidad y* X *sea una lámpara de emergencia. Luego, es posible obtener la siguiente derivación rebatible extendida para el literal "*$habitac\_iluminada(h)$*": noche,* $lampara\_en\_habitac(l, h)$*, electricidad,* $llave\_encendida(l)$*,* $luz\_encendida(l)$*,* $habitac\_iluminada(h)$*.*

Nótese que, en particular, la derivación para el literal "$habitac\_iluminada(h)$*" incluye al literal "electricidad", el cual corresponde al cuerpo de la primera regla de backing del programa* $\mathcal{P}_{5.5}$*. Por otra parte, también es posible obtener una derivación rebatible extendida para el literal "*$\sim habitac\_iluminada(h)$*", correspondiente a la secuencia: noche,* $\sim habitac\_iluminada(h)$*.*

La Definición 5.7 establece que para la obtención de una derivación pueden utilizarse hechos, reglas estrictas, reglas rebatibles y/o reglas de backing de un programa lógico rebatible extendido. En particular, pueden existir derivaciones rebatibles extendidas que sólo utilizan hechos y/o reglas estrictas, las cuales se encuentran caracterizadas por la siguiente definición.

Definición 5.8 (Derivación Estricta para un Literal) *Sea $\mathcal{P} = (\Pi, \Delta, \Sigma)$ un programa lógico rebatible extendido y L un literal para el cual existe una derivación rebatible extendida $L_1, L_2, \ldots L_n = L$. Diremos que L tiene una* derivación estricta *si todas las reglas de $\mathcal{P}$ utilizadas para obtener la secuencia $L_1, L_2, \ldots L_n = L$ son reglas estrictas.*

Ejemplo 5.6 *Dado el programa lógico rebatible extendido $\mathcal{P}_{5.5}$ del Ejemplo 5.5, podemos obtener derivaciones estrictas para todos los hechos del programa, correspondientes a las secuencias:*

- $noche$
- $electricidad$
- $lampara\_en\_habitac(l, h)$
- $llave\_encendida(l)$
- $lampara\_rota(l)$

El Ejemplo 5.5 muestra que, al igual que en DeLP, a partir de un programa lógico rebatible extendido es posible obtener derivaciones rebatibles para un literal y su complemento. En consecuencia, a continuación definiremos la noción de *conjunto contradictorio*, la cual permite identificar literales en conflicto.

Definición 5.9 (Conjunto Contradictorio) *Dado un programa lógico rebatible extendido $\mathcal{P}$, diremos que $C \subseteq \mathcal{P}$ es un* conjunto contradictorio *si y sólo si a partir de C es posible obtener derivaciones estrictas o rebatibles para un literal L y su complemento $\overline{L}$[1].*

Por ejemplo, si consideramos el programa lógico rebatible extendido $\mathcal{P}_{5.5}$ del Ejemplo 5.5, el conjunto $\Pi_{5.5}$ no es contradictorio. En contraste, el conjunto $\Pi_{5.5} \cup \Delta_{5.5} \cup \Sigma_{5.5}$

[1] Por la Definición 2.11, $\overline{L}$ es el complemento del literal L con respecto a la negación fuerte "$\sim$".

es contradictorio ya que permite obtener derivaciones rebatibles para los literales "*habitac_iluminada*(h)" y "$\sim$*habitac_iluminada*(h)".

Como se mencionó al comienzo de esta sección, la noción de derivación rebatible extendida (respectivamente, derivación estricta) provee un mecanismo para la construcción de *argumentos* en E-DeLP. Las definiciones 5.7 y 5.8 muestran que la noción de derivación contempla el uso de reglas estrictas, reglas rebatibles y reglas de backing. En contraste, las reglas de undercutting no son utilizadas para la obtención de derivaciones. No obstante esto, como se verá a continuación, las reglas de undercutting serán utilizadas para construir argumentos en contra del uso de reglas rebatibles. Concretamente, a partir de un programa lógico rebatible extendido $\mathcal{P}$ será posible construir argumentos de diferente tipo. Por una parte, se construirán *argumentos conclusivos* para literales. Por otra parte, será posible construir *argumentos de backing* y *argumentos de undercutting*, los cuales expresan, respectivamente, razones a favor y en contra del uso de reglas rebatibles.

Definición 5.10 (Argumento Conclusivo) *Sea* $\mathcal{P} = (\Pi, \Delta, \Sigma)$ *un programa lógico rebatible extendido y* L *un literal. Un* argumento conclusivo *para* L *es un par* $\langle \mathcal{A}, L \rangle$ *tal que verifica las siguientes condiciones:*

1. $\mathcal{A} \subseteq (\Delta \cup \Sigma)$*;*
2. *existe una derivación rebatible para* L *a partir de* $\Pi \cup \mathcal{A}$*;*
3. $\Pi \cup \mathcal{A}$ *es un conjunto no contradictorio; y*
4. $\mathcal{A}$ *es minimal:* $\nexists \mathcal{A}' \subset \mathcal{A}$ *tal que* $\mathcal{A}'$ *satisface las condiciones (2) y (3).*

Intuitivamente, un argumento conclusivo $\mathcal{A}$ para un literal L se caracteriza por un conjunto minimal y no contradictorio de reglas rebatibles y reglas de backing que permiten obtener una derivación rebatible para L. Nótese que, si bien la primera condición de la Definición 5.10 admitiría la inclusión de reglas de undercutting en el conjunto $\mathcal{A}$, estas quedan excluidas por la última condición que establece la minimalidad del argumento. Esto se debe a que, como se mencionó anteriormente, la noción de derivación rebatible extendida introducida en la Definición 5.7 no emplea reglas de undercutting. Para ilustrar la existencia de argumentos conclusivos consideremos el siguiente ejemplo.

Ejemplo 5.7 *A partir del programa lógico rebatible extendido* $\mathcal{P}_{5.5}$ *del Ejemplo 5.5 es posible construir los siguientes argumentos conclusivos:*

- $\langle \mathcal{A}_1, habitac\_iluminada(h)\rangle$, *con*

$$\mathcal{A}_1 = \left\{ \begin{array}{l} habitac\_iluminada(h) \prec noche, lampara\_en\_habitac(l,h), luz\_encendida(l) \\ luz\_encendida(l) \prec llave\_encendida(l) \\ [luz\_encendida(l) \prec llave\_encendida(l)] \leftarrow\!\oplus\, [electricidad] \end{array} \right\}$$

- $\langle \mathcal{A}_2, \sim\!habitac\_iluminada(h)\rangle$, *con*

$$\mathcal{A}_2 = \left\{ \sim\!habitac\_iluminada(h) \prec noche \right\}$$

- $\langle \mathcal{A}_3, luz\_encendida(l)\rangle$, *con*

$$\mathcal{A}_3 = \left\{ \begin{array}{l} luz\_encendida(l) \prec llave\_encendida(l) \\ [luz\_encendida(l) \prec llave\_encendida(l)] \leftarrow\!\oplus\, [electricidad] \end{array} \right\}$$

- $\langle \emptyset, noche\rangle$
- $\langle \emptyset, electricidad\rangle$
- $\langle \emptyset, lampara\_en\_habitac(l,h)\rangle$
- $\langle \emptyset, llave\_encendida(l)\rangle$
- $\langle \emptyset, lampara\_rota(l)\rangle$

Los argumentos de backing y de undercutting constituyen posiciones a favor y en contra del uso de reglas rebatibles. Por lo tanto, a diferencia de los argumentos conclusivos, en lugar de referenciar a un literal de un programa E-DeLP estos harán referencia explícita a una regla rebatible del programa. Es decir, en lugar de identificar un literal que constituye la conclusión del argumento, los argumentos de backing y undercutting identificarán la regla rebatible para la cual expresan razones a favor o en contra. Asimismo, para distinguir su naturaleza estos argumentos incluirán, respectivamente, la etiqueta "b"(backing) o "u" (undercutting).

Definición 5.11 (Argumento de Backing) *Sea* $\mathcal{P} = (\Pi, \Delta, \Sigma)$ *un programa lógico rebatible extendido y* $R \in \Delta$ *una regla rebatible. Un* argumento de backing *para* R *es un par* $\langle \mathcal{A}, R\rangle_b$ *tal que* $\mathcal{A} = \{[R] \leftarrow\!\oplus\, [L_1, \ldots, L_n]\} \cup \mathcal{A}'$ *y se verifican las siguientes condiciones:*

1. $\mathcal{A} \subseteq (\Delta \cup \Sigma)$;

2. para todo literal L_i $(1 \leq i \leq n)$ existe una derivación rebatible a partir de $\Pi \cup \mathcal{A}'$;

3. $\Pi \cup \mathcal{A}'$ es un conjunto no contradictorio; y

4. $\mathcal{A}$ es minimal: $\nexists \mathcal{B} \subset \mathcal{A}$ tal que $\mathcal{B}$ satisface las condiciones (2) y (3).

Definición 5.12 (Argumento de Undercutting) *Sea $\mathcal{P} = (\Pi, \Delta, \Sigma)$ un programa lógico rebatible extendido y $R \in \Delta$ una regla rebatible. Un* argumento de undercutting *para R es un par $\langle \mathcal{A}, R \rangle_u$ tal que $\mathcal{A} = \{[R] \leftarrow\!\otimes [L_1, \ldots, L_n]\} \cup \mathcal{A}'$ y se verifican las siguientes condiciones:*

1. $\mathcal{A} \subseteq (\Delta \cup \Sigma)$;

2. para todo literal L_i $(1 \leq i \leq n)$ existe una derivación rebatible a partir de $\Pi \cup \mathcal{A}'$;

3. $\Pi \cup \mathcal{A}'$ es un conjunto no contradictorio; y

4. $\mathcal{A}$ es minimal: $\nexists \mathcal{B} \subset \mathcal{A}$ tal que $\mathcal{B}$ satisface las condiciones (2) y (3).

Nótese que las definiciones 5.11 y 5.12 son análogas. Esto se debe a que la única diferencia entre estos argumentos es su naturaleza de soporte o ataque para una regla rebatible R, la cual se ve reflejada en la regla de backing o undercutting para R considerada por el argumento. En particular, las definiciones 5.11 y 5.12 establecen que la regla de backing o undercutting que da origen al argumento debe formar parte de la especificación del programa (determinado por la primera condición). Asimismo, en ambos casos, el proceso de construcción del argumento implica obtener una derivación rebatible para todo literal presente en el cuerpo de la regla de backing o undercutting (determinado por la segunda condición). Finalmente, al igual que los argumentos conclusivos, los argumentos de backing y undercutting deben ser no contradictorios y minimales (determinado por la tercera y cuarta condición). Para ilustrar estas nociones consideremos el siguiente ejemplo.

Ejemplo 5.8 *Dado el programa lógico rebatible extendido $\mathcal{P}_{5.5}$ del Ejemplo 5.5 es posible construir los siguientes argumentos de backing y undercutting:*

- $\langle \mathcal{A}_4, luz\_encendida(l) \prec llave\_encendida(l) \rangle_b$, *donde*
  $\mathcal{A}_4 = \left\{ \; [luz\_encendida(l) \prec llave\_encendida(l)] \leftarrow\!\oplus [electricidad] \; \right\}$

- $\langle \mathcal{A}_5, luz\_encendida(l) \prec llave\_encendida(l) \rangle_u$, *donde*

$$\mathcal{A}_5 = \Big\{ \ [luz\_encendida(l) \prec llave\_encendida(l)] \leftarrow\!\otimes [electricidad, lampara\_rota(l)] \ \Big\}$$

Cabe destacar que los tres tipos de argumento caracterizados por las definiciones 5.10, 5.11 y 5.12 son excluyentes. En los casos en que resulte conveniente nos abstraeremos del tipo de un argumento, refiriéndonos a él simplemente como argumento. Por ejemplo, dado el argumento $\langle \mathcal{A}, h \rangle$, podría ocurrir que este sea un argumento conclusivo (en cuyo caso h es un literal) o que sea un argumento de backing o undercutting (en cuyos casos h es una regla rebatible). Teniendo en cuenta esto, es posible caracterizar la noción de *sub-argumento* como se indica a continuación.

Definición 5.13 (Sub-argumento) *Sea $\mathcal{P}$ un programa lógico rebatible extendido y sean $\langle \mathcal{A}, h \rangle$, $\langle \mathcal{B}, q \rangle$ dos argumentos construidos a partir de $\mathcal{P}$. Diremos que $\langle \mathcal{B}, q \rangle$ es un* sub-argumento *de $\langle \mathcal{A}, h \rangle$ (respectivamente, $\langle \mathcal{A}, h \rangle$ es un* super-argumento *de $\langle \mathcal{B}, q \rangle$) si y sólo si $\mathcal{B} \subseteq \mathcal{A}$.*

En particular, si $\mathcal{B}$ es un subconjunto propio de $\mathcal{A}$ (i. e., $\mathcal{B} \subset \mathcal{A}$), diremos que $\langle \mathcal{B}, q \rangle$ es un sub-argumento propio *de $\langle \mathcal{A}, h \rangle$.*

La definición anterior establece que todo argumento es un sub-argumento de sí mismo. Por otra parte, la noción de sub-argumento propio corresponde a una relación entre argumentos diferentes. Esta noción, como se verá en la Sección 5.4, resultará de gran utilidad para la identificación de conflictos entre los argumentos construidos a partir de un programa E-DeLP. Para ilustrar la noción de sub-argumento (en particular, sub-argumento propio), consideremos el siguiente ejemplo.

Ejemplo 5.9 *Dado el programa lógico rebatible extendido $\mathcal{P}_{5.5}$ del Ejemplo 5.5 y los argumentos identificados en los ejemplos 5.7 y 5.8, tenemos que el argumento conclusivo $\langle \mathcal{A}_3, luz\_encendida(l) \rangle$ es un sub-argumento propio del argumento conclusivo $\langle \mathcal{A}_1, habitac\_iluminada(h) \rangle$. Asimismo, el argumento de backing $\langle \mathcal{A}_4, luz\_encendida(l) \prec llave\_encendida(l) \rangle_b$ es un sub-argumento propio de los argumentos conclusivos $\langle \mathcal{A}_3, luz\_encendida(l) \rangle$ y $\langle \mathcal{A}_1, habitac\_iluminada(h) \rangle$. Por otra parte, todo argumento cuyo conjunto de reglas sea vacío será un sub-argumento propio de cualquier otro argumento construido a partir de $\mathcal{P}_{5.5}$. Por* *ejemplo, los argumentos* $\langle \emptyset, llave\_encendida(l) \rangle$ *y* $\langle \emptyset, noche \rangle$*, entre otros, son sub-argumentos propios de* $\langle \mathcal{A}_1, habitac\_iluminada(h) \rangle$, $\langle \mathcal{A}_2, {\sim}habitac\_iluminada(h) \rangle$, $\langle \mathcal{A}_3, luz\_encendida(l) \rangle$, $\langle \mathcal{A}_4, luz\_encendida(l) \prec llave\_encendida(l) \rangle_b$ *y* $\langle \mathcal{A}_5, luz\_encendida(l) \prec llave\_encendida(l) \rangle_u$.

Dadas las características de los argumentos conclusivos, argumentos de backing y argumentos de undercutting, de acuerdo a lo establecido por las definiciones 5.10, 5.11 y 5.12, estos poseerán a lo sumo una regla de undercutting. Concretamente, dado que en la obtención de derivaciones rebatibles no se emplean reglas de undercutting, los argumentos conclusivos y de backing no poseerán reglas de undercutting. Por otra parte, un argumento de undercutting $\langle \mathcal{A}, R \rangle_u$ poseerá exactamente una regla de undercutting: la regla que expresa razones en contra de la regla rebatible R. A partir de esto, podemos identificar la siguiente propiedad sobre la relación de sub-argumento de E-DeLP.

Proposición 5.1 *Sea* $\mathcal{P}$ *un programa lógico rebatible extendido y* $\langle \mathcal{A}, h \rangle$ *un argumento construido a partir de* $\mathcal{P}$*. Todo sub-argumento propio de* $\langle \mathcal{A}, h \rangle$ *es un argumento conclusivo o un argumento de backing.*

Demostración: ver Apéndice A.

5.4. Identificación de Conflictos: ataques entre argumentos de E-DeLP

En esta sección presentaremos los diferentes conflictos que pueden ocurrir entre argumentos de E-DeLP. Estos conflictos determinarán la relación de ataque entre argumentos de E-DeLP, constituyendo así el tercer elemento de la estructura conceptual presentada en la Sección 2.1. Como se mencionó anteriormente, el uso de la negación fuerte "$\sim$" permite la representación de literales complementarios en E-DeLP. Más aún, es posible obtener derivaciones para un par de literales complementarios, e incluso construir argumentos para esos literales. Claramente, para tales argumentos existirá un conflicto, ya que sus conclusiones se hallarán en contradicción. Asimismo, también identificaremos la existencia de un conflicto entre argumentos tales que al considerar sus conclusiones en forma conjunta con el conocimiento estricto de un programa E-DeLP, el conjunto resultante es contradictorio. Por otra parte, dada la existencia de argumentos de undercutting, identificaremos un conflicto entre un argumento de undercutting para una regla R y aquel argumento que contenga a la regla rebatible R. Teniendo en cuenta estas intuiciones, en esta sección se caracterizará la relación de ataque para los argumentos de E-DeLP, distinguiendo entre ataques de tipo *rebutting*, *undercutting* y *undermining*.

La siguiente definición corresponde a la noción de *desacuerdo* entre literales, la cual generaliza el conflicto existente entre un par de literales complementarios. A partir de la noción de desacuerdo, existirá un conflicto entre cualquier par de literales tales que su consideración conjunta con el conocimiento estricto de un programa E-DeLP conduce a un conjunto contradictorio.

Definición 5.14 (Literales en Desacuerdo) *Sea $\mathcal{P} = (\Pi, \Delta, \Sigma)$ un programa lógico rebatible extendido. Diremos que dos literales h_1 y h_2 están en* desacuerdo *si y sólo si el conjunto $\Pi \cup \{h_1, h_2\}$ es contradictorio.*

Cabe destacar que la noción de *desacuerdo* en E-DeLP se corresponde con aquella presentada para DeLP en la Sección 2.2.2. Luego, tomando como base la noción de desacuerdo, el ataque de tipo *rebutting* entre un par de argumentos de E-DeLP se define de la siguiente manera.

Definición 5.15 (Ataque Rebutting) *Sea $\mathcal{P}$ un programa lógico rebatible extendido. Sean $\langle \mathcal{A}_1, h_1 \rangle$ y $\langle \mathcal{A}_2, h_2 \rangle$ dos argumentos construidos a partir de $\mathcal{P}$, donde $\langle \mathcal{A}_1, h_1 \rangle$ es un argumento conclusivo. Diremos que ocurre un* ataque rebutting *de $\langle \mathcal{A}_1, h_1 \rangle$ sobre $\langle \mathcal{A}_2, h_2 \rangle$ si existe un sub-argumento conclusivo $\langle \mathcal{A}, h \rangle$ de $\langle \mathcal{A}_2, h_2 \rangle$ tal que los literales h_1 y h están en desacuerdo.*

Dado un ataque rebutting de $\langle \mathcal{A}_1, h_1 \rangle$ sobre $\langle \mathcal{A}_2, h_2 \rangle$ diremos que $\langle \mathcal{A}_1, h_1 \rangle$ es un rebut *para $\langle \mathcal{A}_2, h_2 \rangle$. Además, sea $\langle \mathcal{A}, h \rangle$ un sub-argumento de $\langle \mathcal{A}_2, h_2 \rangle$ tal que no existe un sub-argumento propio $\langle \mathcal{B}, q \rangle$ de $\langle \mathcal{A}, h \rangle$ tal que los literales h_1 y q están en desacuerdo. En tal caso, diremos que $\langle \mathcal{A}, h \rangle$ es un* sub-argumento de desacuerdo *en este ataque.*

La Definición 5.15 considera el conflicto existente entre el argumento atacante $\langle \mathcal{A}_1, h_1 \rangle$ y un sub-argumento $\langle \mathcal{A}, h \rangle$ del argumento atacado $\langle \mathcal{A}_2, h_2 \rangle$, dando así lugar a dos alternativas: *i*) el sub-argumento de desacuerdo es $\langle \mathcal{A}_2, h_2 \rangle$, con lo cual $\mathcal{A} = \mathcal{A}_2$ y $h = h_2$; o *ii*) $\langle \mathcal{A}, h \rangle$ es un sub-argumento propio de $\langle \mathcal{A}_2, h_2 \rangle$. Para el primer caso, como los literales h_1

y h_2 están en desacuerdo, se tendrá que el ataque rebutting es simétrico. Es decir, tendremos que $\langle \mathcal{A}_1, h_1 \rangle$ es un rebut para de $\langle \mathcal{A}_2, h_2 \rangle$ y viceversa. Para el segundo caso, como el sub-argumento de desacuerdo $\langle \mathcal{A}, h \rangle$ es un sub-argumento propio de $\langle \mathcal{A}_2, h_2 \rangle$, tendremos que de $\langle \mathcal{A}_1, h_1 \rangle$ es un rebut para $\langle \mathcal{A}_2, h_2 \rangle$, pero $\langle \mathcal{A}_2, h_2 \rangle$ no es un rebut para $\langle \mathcal{A}_1, h_1 \rangle$. No obstante esto, como los literales h_1 y h están en desacuerdo, existirá un conflicto simétrico entre $\langle \mathcal{A}_1, h_1 \rangle$ y $\langle \mathcal{A}, h \rangle$. En consecuencia, tendremos que $\langle \mathcal{A}_1, h_1 \rangle$ es un rebut para $\langle \mathcal{A}, h \rangle$ y viceversa. Nótese que el comportamiento obtenido en los casos $i)$ y $ii)$ refleja el hecho de que, como se vio en la Sección 2.1, los ataques de tipo rebutting poseen una naturaleza simétrica. Para ilustrar este tipo de ataque, consideremos el siguiente ejemplo.

Ejemplo 5.10 *Consideremos el programa lógico rebatible extendido $\mathcal{P}_{5.5}$, correspondiente al Ejemplo 5.5, y los argumentos conclusivos $\langle \mathcal{A}_1, habitac\_iluminada(h) \rangle$ y $\langle \mathcal{A}_2, {\sim}habitac\_iluminada(h) \rangle$ introducidos en el Ejemplo 5.7. En este caso, los literales $habitac\_iluminada(h)$ y ${\sim}habitac\_iluminada(h)$ están en desacuerdo. Por lo tanto, tenemos que $\langle \mathcal{A}_1, habitac\_iluminada(h) \rangle$ es un rebut para $\langle \mathcal{A}_2, {\sim}habitac\_iluminada(h) \rangle$ y viceversa.*

El siguiente ejemplo ilustra un escenario en el que ocurren ataques de tipo rebutting donde el sub-argumento de desacuerdo es, en particular, un sub-argumento propio del argumento atacado.

Ejemplo 5.11 *Sea $\mathcal{P}_{5.11} = (\Pi_{5.11}, \Delta_{5.11}, \emptyset)$ un programa lógico rebatible extendido, donde:*

$$\Pi_{5.11} = \left\{ \; e,\ f \; \right\} \qquad \Delta_{5.11} = \left\{ \begin{array}{ll} a \prec b & d \prec e \\ b \prec c, d & {\sim}b \prec d, f \\ c \prec & {\sim}c \prec f \end{array} \right\}$$

Consideremos los siguientes argumentos construidos a partir de $\mathcal{P}_{5.11}$:

- $\langle \mathcal{B}_1, {\sim}b \rangle$, *con* $\mathcal{B}_1 = \{({\sim}b \prec d, f), (d \prec e)\}$[2].
- $\langle \mathcal{B}_2, a \rangle$, *con* $\mathcal{B}_2 = \{(a \prec b), (b \prec c, d), (c \prec\ ), (d \prec e)\}$.
- $\langle \mathcal{B}_3, b \rangle$, *con* $\mathcal{B}_3 = \{(b \prec c, d), (c \prec\ ), (d \prec e)\}$.

[2]Para una mayor claridad en la notación, en algunos casos se utilizarán paréntesis para distinguir los elementos dentro de un conjunto.

- $\langle \mathcal{B}_4, {\sim}c \rangle$, *con* $\mathcal{B}_4 = \{{\sim}c \prec f\}$.
- $\langle \mathcal{B}_5, c \rangle$, *con* $\mathcal{B}_5 = \{c \prec \}$.

En particular, tenemos que $\langle \mathcal{B}_1, {\sim}b \rangle$ *es tanto un rebut para* $\langle \mathcal{B}_2, a \rangle$ *como para* $\langle \mathcal{B}_3, b \rangle$ *y, en ambos casos, el sub-argumento de desacuerdo es* $\langle \mathcal{B}_3, b \rangle$*. En consecuencia,* $\langle \mathcal{B}_3, b \rangle$ *es un rebut para* $\langle \mathcal{B}_1, {\sim}b \rangle$*, y el sub-argumento de desacuerdo en este ataque es* $\langle \mathcal{B}_1, {\sim}b \rangle$*. Por otra parte, el argumento* $\langle \mathcal{B}_4, {\sim}c \rangle$ *es un rebut para los argumentos* $\langle \mathcal{B}_2, a \rangle$, $\langle \mathcal{B}_3, b \rangle$ *y* $\langle \mathcal{B}_5, c \rangle$*, en cuyos casos el sub-argumento de desacuerdo es* $\langle \mathcal{B}_5, c \rangle$*. Luego,* $\langle \mathcal{B}_5, c \rangle$ *es también un rebut para* $\langle \mathcal{B}_4, {\sim}c \rangle$*, donde el sub-argumento de desacuerdo es* $\langle \mathcal{B}_4, {\sim}c \rangle$*.*

Como se mencionó anteriormente, las reglas de undercutting de un programa E-DeLP son utilizadas para la obtención de argumentos de undercutting, los cuales constituyen una posición en contra del uso de reglas rebatibles. Luego, es posible identificar un conflicto entre un argumento de undercutting para la regla rebatible R y aquellos argumentos que hagan uso de dicha regla. De esta manera, un argumento de undercutting para una regla rebatible R originará un *ataque undercutting* sobre todo argumento que contenga a la regla rebatible R.

Definición 5.16 (Ataque Undercutting) *Sea* $\mathcal{P}$ *un programa lógico rebatible extendido. Sean* $\langle \mathcal{A}, r \rangle$ *y* $\langle \mathcal{B}, h \rangle$ *dos argumentos construidos a partir de* $\mathcal{P}$*, donde* $\langle \mathcal{A}, r \rangle$ *es un argumento de undercutting para la regla rebatible* r*. Diremos que ocurre un* ataque undercutting *de* $\langle \mathcal{A}, r \rangle$ *sobre* $\langle \mathcal{B}, h \rangle$ *si* $r \in \mathcal{B}$*.*

Dado un ataque undercutting de $\langle \mathcal{A}, r \rangle$ *sobre* $\langle \mathcal{B}, h \rangle$ *diremos que* $\langle \mathcal{A}, r \rangle$ *es un* undercut *para* $\langle \mathcal{B}, h \rangle$*. Además, si* $\langle \mathcal{C}, q \rangle$ *es el sub-argumento minimal (con respecto a* $\subseteq$*) de* $\langle \mathcal{B}, h \rangle$ *que contiene a la regla* r *(i. e., no existe un sub-argumento propio* $\langle \mathcal{D}, p \rangle$ *de* $\langle \mathcal{C}, q \rangle$ *tal que* $r \in \mathcal{D}$*), diremos que* $\langle \mathcal{C}, q \rangle$ *es el* sub-argumento de desacuerdo *en este ataque.*

En particular, la Definición 5.16 considera el conflicto existente entre el argumento atacante $\langle \mathcal{A}, r \rangle$ y todos los sub-argumentos del argumento atacado $\langle \mathcal{B}, h \rangle$ que contengan a la regla rebatible r. En consecuencia, $\langle \mathcal{A}, r \rangle$ será un undercut para todo sub-argumento de $\langle \mathcal{B}, h \rangle$ que contenga a la regla rebatible r. No obstante esto, como se verá en la Sección 5.5, resulta necesario identificar cuál es el sub-argumento más chico de $\langle \mathcal{B}, h \rangle$ que contiene a la regla r. Este sub-argumento, por ejemplo $\langle \mathcal{C}, q \rangle$, es identificado como el sub-argumento de desacuerdo dado que es el sub-argumento que origina el conflicto entre $\langle \mathcal{A}, r \rangle$ y $\langle \mathcal{B}, h \rangle$. Para ilustrar este tipo de ataque consideremos el siguiente ejemplo.

Ejemplo 5.12 *Sea $\mathcal{P}_{5.5}$ el programa lógico rebatible extendido del Ejemplo 5.5. Consideremos los argumentos construidos a partir de $\mathcal{P}_{5.5}$ que se ilustran en los ejemplos 5.7 y 5.8. El argumento $\langle \mathcal{A}_5, luz\_encendida(l) \prec llave\_encendida(l) \rangle_u$ es un undercut para el argumento $\langle \mathcal{A}_1, habitac\_iluminada(h) \rangle$. Asimismo, $\langle \mathcal{A}_5, luz\_encendida(l) \prec llave\_encendida(l) \rangle_u$ es un undercut para el sub-argumento $\langle \mathcal{A}_3, luz\_encendida(l) \rangle$ de $\langle \mathcal{A}_1, habitac\_iluminada(h) \rangle$. En particular, el sub-argumento de desacuerdo, en ambos casos, es $\langle \mathcal{A}_3, luz\_encendida(l) \rangle$.*

A continuación caracterizaremos los ataques de tipo undermining, los cuales constituyen un caso particular de ataque rebutting. Recordemos que todo argumento es construido sobre la base de premisas. Para el caso de los argumentos construidos a partir de un programa E-DeLP, las premisas se hallan conformadas por hechos y presuposiciones del programa. De acuerdo a la Definición 5.15, un ataque rebutting constituye un ataque sobre la conclusión de un argumento (ya sea la conclusión final o la conclusión de uno de sus sub-argumentos). Luego, dadas las características de los argumentos de E-DeLP, puede ocurrir que la conclusión de un argumento sea, en particular, una de sus premisas. Tal es el caso de los argumentos cuya conclusión es un hecho o una presuposición del programa E-DeLP. Por otra parte, la tercera condición de las definiciones 5.10, 5.11 y 5.12 establece que los argumentos deben ser consistentes con respecto al conocimiento estricto de un programa E-DeLP. Por lo tanto, esto implica que las únicas premisas atacables de un argumento son las presuposiciones. En consecuencia, cuando ocurra un ataque rebutting tal que la conclusión del argumento (sub-argumento) atacado es una presuposición, diremos que ocurre un ataque de tipo *undermining*.

Definición 5.17 (Ataque Undermining) *Sea $\mathcal{P} = (\Pi, \Delta, \Sigma)$ un programa lógico rebatible extendido. Sean $\langle \mathcal{A}_1, h_1 \rangle$ y $\langle \mathcal{A}_2, h_2 \rangle$ dos argumentos construidos a partir de $\mathcal{P}$, donde $\langle \mathcal{A}_1, h_1 \rangle$ es un argumento conclusivo. Diremos que ocurre un* ataque undermining *de $\langle \mathcal{A}_1, h_1 \rangle$ sobre $\langle \mathcal{A}_2, h_2 \rangle$ si existe un sub-argumento conclusivo $\langle \mathcal{A}, h \rangle$ de $\langle \mathcal{A}_2, h_2 \rangle$ tal que los literales h_1 y h están en desacuerdo y "h" es una presuposición de $\mathcal{P}$ (i. e., $h \prec\ \in \Delta$).*

Dado un ataque undermining de $\langle \mathcal{A}_1, h_1 \rangle$ sobre $\langle \mathcal{A}_2, h_2 \rangle$ diremos que $\langle \mathcal{A}_1, h_1 \rangle$ es un undermine *para $\langle \mathcal{A}_2, h_2 \rangle$. Además, sea $\langle \mathcal{A}, h \rangle$ un sub-argumento de $\langle \mathcal{A}_2, h_2 \rangle$ tal que no existe un sub-argumento propio $\langle \mathcal{B}, q \rangle$ de $\langle \mathcal{A}, h \rangle$ tal que los literales h_1 y q están en desacuerdo. En tal caso, diremos que $\langle \mathcal{A}, h \rangle$ es un* sub-argumento de desacuerdo *en este ataque.*

Ejemplo 5.13 *Consideremos el programa lógico rebatible extendido $\mathcal{P}_{5.11}$ y los argumentos identificados en el Ejemplo 5.11. A partir de escenario, tenemos que el argumento $\langle \mathcal{B}_4, {\sim}c \rangle$ es un rebut para los argumentos $\langle \mathcal{B}_2, a \rangle$ y $\langle \mathcal{B}_3, b \rangle$ y $\langle \mathcal{B}_5, c \rangle$, donde el sub-argumento de desacuerdo en todos los casos es $\langle \mathcal{B}_5, c \rangle$. Luego, dado que "c" es una presuposición de $\mathcal{P}_{5.11}$ (i. e., $c \prec \ \in \Delta_{5.11}$), tenemos que $\langle \mathcal{B}_4, {\sim}c \rangle$ es también un undermine para $\langle \mathcal{B}_2, a \rangle$, $\langle \mathcal{B}_3, b \rangle$ y $\langle \mathcal{B}_5, c \rangle$.*

A continuación introduciremos una representación gráfica para los argumentos y ataques en E-DeLP, la cual será extendida en la Sección 5.5 para considerar la noción de *derrota*. En esta representación, los argumentos serán denotados mediante triángulos. En particular, los triángulos dentro de triángulos mayores corresponderán a sub-argumentos propios. Para el caso de los argumentos conclusivos, la información ubicada en el vértice superior del triángulo corresponderá a la conclusión del argumento. Por otra parte, para los argumentos de backing y undercutting, la información ubicada en el vértice superior del triángulo corresponderá a la regla que está siendo soportada o atacada. Asimismo, para cualquier tipo de argumento, la información situada dentro del triángulo corresponderá a los literales utilizados para su construcción. El símbolo " $\prec$ " representa la conexión establecida entre estos literales a través de las reglas rebatibles pertenecientes al argumento. De manera similar, los símbolos "$\leftarrow\oplus$" y "$\leftarrow\otimes$" representan, respectivamente, el soporte y ataque hacia una regla rebatible de un argumento. En particular, los argumentos cuya conclusión es derivada a partir del conocimiento estricto de un programa E-DeLP se denotarán identificando únicamente el literal que concluyen. Por último, los ataques entre argumentos se denotarán mediante flechas punteadas. Para ilustrar esta notación, consideremos los siguientes ejemplos.

Ejemplo 5.14 *Sea $\mathcal{P}_{5.5}$ el programa lógico rebatible extendido del Ejemplo 5.5. La Figura 5.2 ilustra los argumentos construidos a partir de $\mathcal{P}_{5.5}$, correspondientes a los ejemplos 5.7 y 5.8. Asimismo, las flechas punteadas en la Figura 5.2 corresponden a los ataques identificados en los ejemplos 5.10 y 5.12.*

Nótese que, en este caso, el argumento $\langle \mathcal{A}_5, luz\_encendida(l) \prec llave\_encendida(l) \rangle_u$ es un undercut para el argumento $\langle \mathcal{A}_3, luz\_encendida(l) \rangle$ y su super-argumento $\langle \mathcal{A}_1, habitac\_iluminada(h) \rangle$. En particular, para los dos ataques arriba mencionados el sub-argumento de desacuerdo es $\langle \mathcal{A}_3, luz\_encendida(l) \rangle$. Por otra parte, el argumento $\langle \mathcal{A}_2, {\sim}habitac\_iluminada(h) \rangle$ es un rebut para el argumento

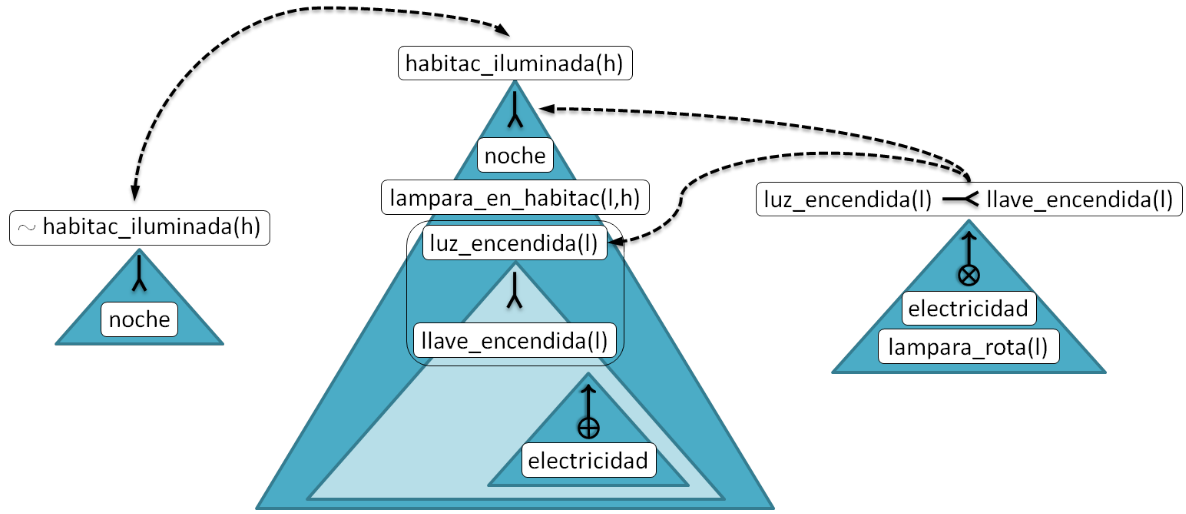


Figura 5.2: Representación gráfica de argumentos y ataques para el Ejemplo 5.14.

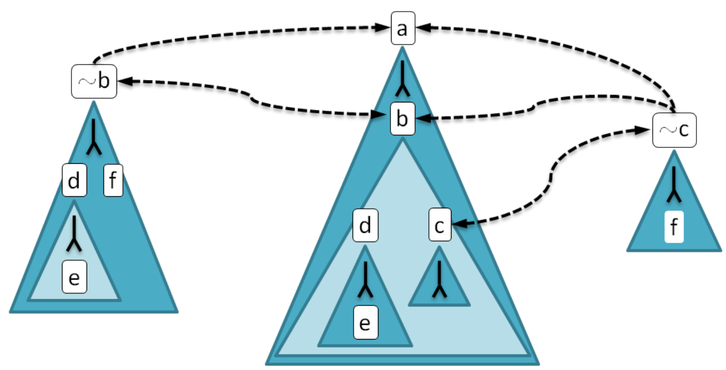


Figura 5.3: Representación gráfica de argumentos y ataques para el Ejemplo 5.15.

$\langle \mathcal{A}_1, habitac\_iluminada(h)\rangle$ *y, análogamente,* $\langle \mathcal{A}_1, habitac\_iluminada(h)\rangle$ *es un rebut para* $\langle \mathcal{A}_2, {\sim}habitac\_iluminada(h)\rangle$.

Ejemplo 5.15 *Sea $\mathcal{P}_{5.11}$ el programa lógico rebatible extendido del Ejemplo 5.11. La Figura 5.3 ilustra los argumentos construidos a partir de $\mathcal{P}_{5.11}$, correspondientes al Ejemplo 5.11. Asimismo, las flechas punteadas en la Figura 5.3 corresponden a los ataques identificados en los ejemplos 5.11 y 5.13.*

5.5. Resolución de Conflictos: derrotas entre argumentos de E-DeLP

La noción de ataque introducida en la sección anterior captura la existencia de un conflicto entre dos argumentos construidos a partir de un programa E-DeLP. Para el caso de los ataques rebutting y undermining este conflicto se desprende del uso de la negación fuerte. Es decir, dado un argumento $\mathcal{A}$ que es un rebut o un undermine para un argumento $\mathcal{B}$, existe un sub-argumento $\mathcal{C}$ de $\mathcal{B}$ tal que las conclusiones de $\mathcal{A}$ y $\mathcal{C}$ están en desacuerdo. Por otra parte, para el caso de los ataques undercutting, el conflicto se origina en el uso de reglas de undercutting. De esta manera, dado un argumento $\mathcal{A}$ que es un undercut para un argumento $\mathcal{B}$, el argumento $\mathcal{B}$ contiene una regla rebatible R, mientras que $\mathcal{A}$ provee razones en contra del uso de la regla R. No obstante esto, la noción de ataque permite expresar conflictos entre argumentos, los cuales no siempre tienen éxito. Como se vio en el Capítulo 2, la resolución de los conflictos expresados mediante ataques conduce a la obtención de *derrotas* entre argumentos. En particular, la obtención de derrotas corresponde al cuarto elemento de la estructura conceptual presentada en la Sección 2.1.

La determinación del éxito de los ataques implica algún tipo de evaluación sobre los argumentos involucrados en el conflicto. En general, los formalismos argumentativos existentes en la literatura realizan esta evaluación mediante el uso de un criterio de comparación entre argumentos (o una relación de preferencia), asignando algún tipo de peso a cada uno de ellos (ver *e. g.*, [GS04, Pra09]). En consecuencia, la noción de derrota captura la intuición de "ataca y no es más débil" (en su versión débil) o "ataca y es más fuerte" (en su versión fuerte). Por ejemplo, adoptando la versión fuerte, la derrota de un argumento $\mathcal{A}$ sobre un argumento $\mathcal{B}$ implica una preferencia del primero por sobre el segundo. En particular, al igual que en DeLP, en E-DeLP se utilizará un criterio de comparación entre argumentos que es modular. De esta manera, el criterio será especificado por el usuario del sistema, permitiendo así emplear aquel que mejor se adapte a cada dominio de aplicación.

Dado el conjunto de argumentos construidos a partir de un programa E-DeLP, la relación de sub-argumento, los ataques identificados entre estos argumentos, y el criterio de comparación adoptado, se especificará un *Marco Argumentativo con Backing y Undercutting (BUAF)* asociado al programa E-DeLP. Luego, a partir del BUAF asociado se obtendrán las derrotas entre argumentos, las cuales serán posteriormente utilizadas para efectuar el cálculo de aceptabilidad sobre los argumentos del sistema. En particular, como se verá en la Sección 5.6, esta aproximación permitirá el uso de técnicas alternativas para el cálculo de aceptabilidad, tales como la construcción de árboles de dialéctica [GS04] o la aplicación de diferentes semánticas de aceptabilidad propuestas en la literatura [Dun95, BCG11].

Recordemos que, como se presentó en el Capítulo 4, un BUAF es una tupla $\langle \mathbb{A}, \mathbb{R}, \mathbb{S}, \preceq \rangle$, donde:

- $\mathbb{A}$ es un conjunto de argumentos;
- $\mathbb{R} \subseteq \mathbb{A} \times \mathbb{A}$ es una relación de ataque entre argumentos, distinguiendo entre ataques de tipo rebutting $\mathbb{R}_b$, undercutting $\mathbb{U}_c$ y undermining $\mathbb{U}_m$ (*i. e.*, $\mathbb{R} = \mathbb{R}_b \cup \mathbb{U}_c \cup \mathbb{U}_m$);
- $\mathbb{S} \subseteq \mathbb{A} \times \mathbb{A}$ es una relación de soporte entre argumentos, distinguiendo entre el soporte brindado por backings $\mathbb{B}_k$ y sub-argumentos $\mathbb{S}_{\sqsubseteq}$ (*i. e.*, $\mathbb{S} = \mathbb{B}_k \cup \mathbb{S}_{\sqsubseteq}$);
- $\preceq \subseteq \mathbb{A} \times \mathbb{A}$ es una relación de preferencia entre argumentos; y
- las relaciones de ataque y soporte son disjuntas (*i. e.*, $\mathbb{R} \cap \mathbb{S} = \emptyset$).

Asimismo, las relaciones de ataque y soporte de un BUAF poseen restricciones adicionales. Por una parte, los conjuntos $\mathbb{R}_b$, $\mathbb{U}_c$ y $\mathbb{U}_m$, correspondientes a los diferentes tipos de ataque dentro de la relación $\mathbb{R}$ del BUAF, deben ser disjuntos de a pares. Es decir, en el contexto de un BUAF un argumento no podrá atacar a otro en más de una forma y, por lo tanto, dado un par $(\mathcal{A}, \mathcal{B}) \in \mathbb{R}$ se tendrá que $(\mathcal{A}, \mathcal{B}) \in \mathbb{R}_b$, $(\mathcal{A}, \mathcal{B}) \in \mathbb{U}_c$ o bien $(\mathcal{A}, \mathcal{B}) \in \mathbb{U}_m$. Por otra parte, la relación de sub-argumento $\mathbb{S}_{\sqsubseteq}$ se corresponde con la noción de sub-argumento propio. Así, para todo par de argumentos $(\mathcal{A}, \mathcal{B}) \in \mathbb{S}_{\sqsubseteq}$ se tendrá que $\mathcal{A} \neq \mathcal{B}$ (*i. e.*, la relación de sub-argumento $\mathbb{S}_{\sqsubseteq}$ de un BUAF es irreflexiva).

A continuación se explicarán las intuiciones detrás del proceso de transformación que nos permitirá obtener, a partir de un programa E-DeLP, la especificación del BUAF asociado a dicho programa. Concretamente, se determinará cómo se obtienen los cuatro componentes del BUAF asociado: conjunto de argumentos $\mathbb{A}$, relación de ataque $\mathbb{R}$, relación de soporte $\mathbb{S}$ y relación de preferencia $\preceq$.

El primer componente del BUAF asociado corresponde al conjunto de argumentos $\mathbb{A}$ utilizados por el sistema. En particular, dado un programa E-DeLP, las definiciones 5.10, 5.11 y 5.12 proveen un mecanismo para obtener todos los argumentos construibles a partir de dicho programa. En consecuencia, el conjunto de argumentos $\mathbb{A}$ del

BUAF asociado será el conjunto de argumentos construidos a partir del programa E-DeLP mediante la aplicación de estas definiciones.

La relación de ataque $\mathbb{R}$ del BUAF asociado será especificada a partir de las definiciones 5.15, 5.16 y 5.17, las cuales caracterizan los ataques ocurridos entre los argumentos de E-DeLP. No obstante esto, se debe considerar la restricción sobre la relación de ataque de los BUAF, la cual establece que un argumento no puede atacar a otro en más de un sentido. Dada esta restricción, es necesario analizar el caso de los ataques undermining identificados en E-DeLP. Esto se debe a que, como se vio en la Definición 5.17, un ataque undermining es un ataque rebutting que, en particular, ataca un literal correspondiente a una presuposición. Luego, al especificar la relación de ataque del BUAF asociado, para cada ataque undermining de E-DeLP deberá optarse por una de las siguientes alternativas: incorporarlo al BUAF como un ataque rebutting o incorporarlo al BUAF como un ataque undermining. En esta tesis se adoptará la segunda alternativa, dado que se considera que los ataques de tipo undermining corresponden a un caso particular de ataque rebutting y, por lo tanto, son un tipo de ataque "más específico" que el rebutting.

Por otra parte, a partir de las definiciones 5.15, 5.16 y 5.17, dado un ataque (rebutting, undercutting o undermining) de un argumento $\mathcal{A}$ sobre un argumento $\mathcal{B}$ también existirá un ataque (del mismo tipo) del argumento $\mathcal{A}$ sobre todo argumento $\mathcal{C}$ tal que $\mathcal{C}$ es un super-argumento de $\mathcal{B}$. Cabe destacar que estos últimos ataques surgen como consecuencia de la relación de sub-argumento entre $\mathcal{B}$ y $\mathcal{C}$. Teniendo en cuenta esto, la especificación de la relación de ataque del BUAF asociado se realizará adoptando la siguiente convención. Dado el ataque de un argumento $\mathcal{A}$ sobre un argumento $\mathcal{B}$, donde $\mathcal{D}$ es un sub-argumento de desacuerdo en este ataque, la relación de ataque del BUAF asociado incluirá únicamente el ataque de $\mathcal{A}$ sobre $\mathcal{D}$. Es decir, se dejarán de lado todos los ataques sobre los super-argumentos de $\mathcal{D}$. Sin embargo, como se verá más adelante, los conflictos entre el argumento atacante $\mathcal{A}$ y los super-argumentos de $\mathcal{D}$ serán capturados por la noción de *derrota indirecta* del BUAF asociado (ver Definición 4.4).

Consideremos ahora la especificación de la relación de soporte $\mathbb{S}$ del BUAF asociado. Como se mencionó anteriormente, la relación de soporte de un BUAF distingue entre el soporte brindado por los backings (subconjunto $\mathbb{B}_k$) y el soporte brindado por los sub-argumentos (subconjunto $\mathbb{S}_{\sqsubseteq}$). Además, la relación de sub-argumento de un BUAF se corresponde con la noción de sub-argumento propio.

Dado un argumento de backing $\mathcal{B}$ para una regla rebatible R y un argumento $\mathcal{A}$ que contiene a la regla R, resulta razonable incluir el par $(\mathcal{B}, \mathcal{A})$ en el subconjunto $\mathbb{B}_k$ de la relación de soporte del BUAF asociado. Nótese que, en particular, todo super-argumento de $\mathcal{A}$ también contendrá a la regla rebatible R. Esto sugiere que, por cada super-argumento $\mathcal{C}$ de $\mathcal{A}$, sería necesario incluir el par $(\mathcal{B}, \mathcal{C})$ en el subconjunto $\mathbb{B}_k$ del BUAF asociado. Sin embargo, la relación existente entre $\mathcal{B}$ y $\mathcal{C}$ surge como consecuencia de la relación de sub-argumento entre $\mathcal{A}$ y $\mathcal{C}$. Teniendo en cuenta esto, la especificación del subconjunto $\mathbb{B}_k$ del BUAF asociado será realizada adoptando la siguiente convención. Dado el argumento de backing $\mathcal{B}$ para la regla rebatible R, el subconjunto $\mathbb{B}_k$ del BUAF asociado incluirá el par $(\mathcal{B}, \mathcal{A})$ si $\mathcal{A}$ contiene a R y no existe un sub-argumento propio $\mathcal{D}$ de $\mathcal{A}$ que contenga a R. Es decir, el subconjunto $\mathbb{B}_k$ incluirá únicamente la relación existente entre un argumento de backing $\mathcal{B}$ para una regla R y los argumentos minimales $\mathcal{A}$ (con respecto a $\subseteq$) que contengan a R. Por otra parte, la relación existente entre $\mathcal{B}$ y los super-argumentos $\mathcal{C}$ de $\mathcal{A}$ se hallará implícita mediante la relación de sub-argumento existente entre $\mathcal{A}$ y $\mathcal{C}$.

Para la determinación de los elementos pertenecientes al subconjunto $\mathbb{S}_{\sqsubseteq}$ se utilizará la noción de sub-argumento propio para E-DeLP caracterizada por la Definición 5.13. Luego, dados dos argumentos $\mathcal{A}$ y $\mathcal{B}$, si $\mathcal{A}$ es un sub-argumento propio de $\mathcal{B}$, resulta razonable incluir el par $(\mathcal{A}, \mathcal{B})$ en el subconjunto $\mathbb{S}_{\sqsubseteq}$ del BUAF asociado. Nótese que, además, para todo super-argumento $\mathcal{C}$ de $\mathcal{B}$ se tiene que $\mathcal{A}$ es un sub-argumento propio de $\mathcal{C}$. Esto sugiere que por cada super-argumento $\mathcal{C}$ de $\mathcal{B}$ sería necesario incluir el par $(\mathcal{A}, \mathcal{C})$ en el subconjunto $\mathbb{S}_{\sqsubseteq}$ del BUAF asociado. Sin embargo, la relación existente entre $\mathcal{A}$ y $\mathcal{C}$ surge como consecuencia de la relación de sub-argumento entre $\mathcal{B}$ y $\mathcal{C}$. Teniendo en cuenta esto, la especificación del subconjunto $\mathbb{S}_{\sqsubseteq}$ del BUAF asociado se realizará adoptando la siguiente convención. Dados dos argumentos $\mathcal{A}$ y $\mathcal{B}$ tales que $\mathcal{A}$ es sub-argumento propio de $\mathcal{B}$, el subconjunto $\mathbb{S}_{\sqsubseteq}$ del BUAF asociado incluirá el par $(\mathcal{A}, \mathcal{B})$ si no existe un sub-argumento propio $\mathcal{D}$ de $\mathcal{B}$ tal que $\mathcal{A}$ es un sub-argumento propio de $\mathcal{D}$. Es decir, el subconjunto $\mathbb{S}_{\sqsubseteq}$ incluirá únicamente los pares correspondientes a la relación de sub-argumento propio *directa*. De esta manera, la relación existente entre $\mathcal{A}$ y los super-argumentos $\mathcal{C}$ de $\mathcal{B}$ se hallará implícita mediante la relación de sub-argumento existente entre $\mathcal{A}$ y $\mathcal{B}$, y aquella existente entre $\mathcal{B}$ y $\mathcal{C}$. En particular, como se vio en el Capítulo 4, esta transitividad de la relación de sub-argumento es capturada en los BUAF mediante cadenas de soporte.

Con respecto a la obtención del subconjunto $\mathbb{S}_{\sqsubseteq}$ del BUAF asociado es necesario efectuar una consideración adicional. Recordemos que, de acuerdo a la Definición 5.13, la noción de sub-argumento propio para E-DeLP relaciona dos argumentos $\mathcal{A}$ y $\mathcal{B}$ tales que el conjunto de reglas rebatibles de $\mathcal{A}$ se halla incluido estrictamente en el conjunto de reglas rebatibles de $\mathcal{B}$. Luego, para todo literal L cuya derivación utiliza únicamente el conocimiento estricto de un programa E-DeLP, existirá un argumento $\langle\emptyset, L\rangle$. Por lo tanto, para todo argumento $\mathcal{C}$ construido a partir de un programa E-DeLP cuya derivación implique el uso de reglas rebatibles se tendrá que $\langle\emptyset, L\rangle$ es un sub-argumento propio de $\mathcal{C}$. No obstante esto, el interés en caracterizar el BUAF asociado a un programa E-DeLP radica en la obtención de las derrotas entre argumentos para posteriormente efectuar el cálculo de aceptabilidad de los mismos. Por lo tanto, para la determinación del subconjunto $\mathbb{S}_{\sqsubseteq}$ del BUAF asociado se considerarán únicamente aquellas relaciones de sub-argumento que sean relevantes. De esta manera, dado un argumento $\langle\emptyset, L\rangle$, se considerará únicamente la relación de sub-argumento propio existente entre este y aquellos otros argumentos que utilicen a L en el proceso de derivación.

Por otra parte, dadas las características de los argumentos de backing de E-DeLP, estos pueden ser sub-argumentos propios de otros argumentos. Luego, dados dos argumentos $\mathcal{A}$ y $\mathcal{B}$ tales que el par $(\mathcal{B}, \mathcal{A})$ pertenece al subconjunto $\mathbb{B}_k$ del BUAF asociado, también se tendrá que $\mathcal{B}$ es un sub-argumento propio de $\mathcal{A}$. Esto sugiere que el par $(\mathcal{B}, \mathcal{A})$ también debería pertenecer al subconjunto $\mathbb{S}_{\sqsubseteq}$ del BUAF asociado. Sin embargo, como los backings son sub-argumentos con características distinguidas, para evitar redundancia en estos casos se optará por incluir el par $(\mathcal{B}, \mathcal{A})$ únicamente en el subconjunto $\mathbb{B}_k$. En consecuencia, los subconjuntos $\mathbb{B}_k$ y $\mathbb{S}_{\sqsubseteq}$ correspondientes a la relación de soporte $\mathbb{S}$ del BUAF asociado serán disjuntos.

Cabe destacar que la especificación del BUAF asociado a un programa E-DeLP cumplir la restricción de que las relaciones de ataque y soporte sean disjuntas. En particular, como se verá a continuación, la caracterización de los diferentes tipos de argumento en E-DeLP provista por las definiciones 5.10, 5.11 y 5.12 asegura que las relaciones de ataque y soporte del BUAF asociado a un programa E-DeLP serán disjuntas.

Consideremos el par $(\mathcal{A}, \mathcal{B})$ perteneciente a la relación de ataque $\mathbb{R}$ del BUAF asociado, donde el ataque de $\mathcal{A}$ sobre $\mathcal{B}$ es de tipo rebutting o undermining (*i. e.*, $(\mathcal{A}, \mathcal{B}) \in \mathbb{R}_b$ o $(\mathcal{A}, \mathcal{B}) \in \mathbb{U}_m$). Esto implica que, por las definiciones 5.15 y 5.17, $\mathcal{A}$ es un argumento conclusivo de E-DeLP. Luego, $\mathcal{A}$ no puede ser un argumento de backing, con lo cual el par $(\mathcal{A}, \mathcal{B})$ no pertenecerá al subconjunto $\mathbb{B}_k$ de la relación de soporte $\mathbb{S}$ del BUAF asociado. De manera similar, si $(\mathcal{A}, \mathcal{B}) \in \mathbb{R}_b$ o $(\mathcal{A}, \mathcal{B}) \in \mathbb{U}_m$ esto implica que $\mathcal{A}$ es un rebut o un undermine para $\mathcal{B}$, donde $\mathcal{B}$ es el sub-argumento de desacuerdo en el ataque. En particular, por las definiciones 5.15 y 5.17, $\mathcal{B}$ será un argumento conclusivo tal que las conclusiones de $\mathcal{A}$ y $\mathcal{B}$ están en desacuerdo. Por lo tanto, no puede ser el caso que $\mathcal{A}$ sea un sub-argumento de $\mathcal{B}$ ya que, de ocurrir esto, el argumento $\mathcal{B}$ violaría la condición de consistencia impuesta por la Definición 5.10.

Consideremos ahora el caso en que el ataque de $\mathcal{A}$ sobre $\mathcal{B}$ en el BUAF asociado es de tipo undercutting (*i. e.*, $(\mathcal{A}, \mathcal{B}) \in \mathbb{U}_c$). En este caso, $\mathcal{A}$ es un argumento de undercutting construido a partir del programa E-DeLP. Por lo tanto, por la Proposición 5.1, $\mathcal{A}$ no puede ser un sub-argumento propio de $\mathcal{B}$. En consecuencia, el par $(\mathcal{A}, \mathcal{B})$ no pertenecerá al subconjunto $\mathbb{S}_{\sqsubseteq}$ de la relación de soporte $\mathbb{S}$ del BUAF asociado. Por otra parte, para que el par $(\mathcal{A}, \mathcal{B})$ pertenezca al subconjunto $\mathbb{B}_k$ de la relación de soporte $\mathbb{S}$ del BUAF asociado, debería ocurrir que $\mathcal{A}$ sea un argumento de backing. Sin embargo, la situación aquí presentada establece que $\mathcal{A}$ es un argumento de undercutting. En consecuencia, $\mathcal{A}$ no es un argumento de backing y, por lo tanto, el par $(\mathcal{A}, \mathcal{B})$ no pertenecerá al subconjunto $\mathbb{B}_k$ de la relación de soporte del BUAF asociado.

A partir del análisis arriba realizado se tiene que las relaciones de ataque y soporte correspondientes a la especificación del BUAF asociado a un programa E-DeLP serán disjuntas. De esta manera, hasta el momento se estableció cómo se obtendrán los primeros tres elementos del BUAF asociado, restando determinar cómo se obtendrá el último de ellos: la relación de preferencia entre argumentos $\preceq$.

Como se mencionó al comienzo de esta sección, en E-DeLP se utiliza un criterio de comparación entre argumentos, el cual establece una preferencia entre los mismos. De esta manera, dados dos argumentos $\mathcal{A}$ y $\mathcal{B}$ construidos a partir de un programa E-DeLP, tendremos una de las siguientes alternativas: (i) $\mathcal{A}$ es estrictamente más preferido que $\mathcal{B}$, (ii) $\mathcal{B}$ es estrictamente más preferido que $\mathcal{A}$, (iii) $\mathcal{A}$ y $\mathcal{B}$ son igualmente preferidos (*i. e.*, equivalentes) de acuerdo al criterio de comparación o (iv) $\mathcal{A}$ y $\mathcal{B}$ son incomparables de acuerdo al criterio de comparación. Teniendo en cuenta esto, dados dos argumentos $\mathcal{A}$ y $\mathcal{B}$, para la alternativa (i) la relación de preferencia $\preceq$ del BUAF asociado incluirá únicamente el par $(\mathcal{B}, \mathcal{A})$. De manera similar, dada la alternativa (ii) se tendrá únicamente que $(\mathcal{A}, \mathcal{B}) \in \preceq$. Para la alternativa (iii) se tendrá que $(\mathcal{A}, \mathcal{B}) \in \preceq$ y $(\mathcal{B}, \mathcal{A}) \in \preceq$. Finalmente, si el criterio de comparación de E-DeLP es tal que ocurre la situación considerada por la alternativa (iv), la relación de preferencia $\preceq$ del BUAF asociado no contendrá información que vincule a los argumentos $\mathcal{A}$ y $\mathcal{B}$.

Por último, con respecto al subconjunto $\mathbb{B}_k$ del BUAF asociado es necesario considerar lo siguiente. Recordemos que un argumento de backing provee soporte para una regla rebatible. Luego, como se mencionó anteriormente, el subconjunto $\mathbb{B}_k$ de la relación de soporte $\mathbb{S}$ del BUAF asociado relacionará a un argumento de backing con un argumento "minimal" que contiene a la regla que el primero soporta. En otras palabras, dado un argumento de backing $\mathcal{B}$ para una regla rebatible r, se establecerá una relación entre $\mathcal{B}$ y un argumento $\mathcal{A}$ cuya conclusión es obtenida a partir de la regla rebatible r. Con el objetivo de establecer esta relación, la siguiente definición introduce la noción de *regla rebatible conclusiva* de un argumento de E-DeLP.

Definición 5.18 (Regla Rebatible Conclusiva) *Sea $\mathcal{P} = (\Pi, \Delta, \Sigma)$ un programa lógico rebatible extendido, $\langle \mathcal{A}, h \rangle$ un argumento conclusivo construido a partir de $\mathcal{P}$, y $r =$ "$h \prec Cuerpo$" una regla rebatible de Δ. Si $r \in \mathcal{A}$, entonces diremos que r es una* regla rebatible conclusiva *del argumento $\langle \mathcal{A}, h \rangle$.*

Cabe destacar que la noción de regla rebatible conclusiva permite identificar aquellos argumentos cuya conclusión es obtenida mediante el uso de una regla rebatible. En consecuencia, aquellos argumentos cuya conclusión corresponde a un hecho del programa o a la cabeza de una regla estricta, no poseerán ninguna regla rebatible conclusiva.

A continuación formalizaremos las intuiciones arriba presentadas, caracterizando un mapeo que, a partir de un programa E-DeLP $\mathcal{P}$, permite obtener un BUAF. Este BUAF será el *marco argumentativo con backing y undercutting asociado a* $\mathcal{P}$, y será denotado como $\mathsf{BUAF}_{\mathcal{P}}$.

Definición 5.19 (BUAF asociado a un Programa E-DeLP) *Sea $\mathcal{P}$ un programa lógico rebatible extendido. El* marco argumentativo con backing y undercutting asociado a $\mathcal{P}$ *es $\mathsf{BUAF}_{\mathcal{P}} = \langle \mathbb{A}_{\mathcal{P}}, \mathbb{R}_{\mathcal{P}}, \mathbb{S}_{\mathcal{P}}, \preceq_{\mathcal{P}} \rangle$, donde:*

- *$\mathbb{A}_{\mathcal{P}}$ es el conjunto de argumentos construidos a partir de $\mathcal{P}$, de acuerdo a las Definiciones 5.10, 5.11 y 5.12.*
- *$\mathbb{R}_{\mathcal{P}} = \mathbb{R}_{b\mathcal{P}} \cup \mathbb{U}_{c\mathcal{P}} \cup \mathbb{U}_{m\mathcal{P}}$ es obtenida a partir de las definiciones 5.15, 5.16 y 5.17, de la siguiente manera:*

- *Sean* $\langle\mathcal{A}_1, h_1\rangle, \langle\mathcal{A}_2, h_2\rangle \in \mathbb{A}_{\mathcal{P}}$ *tales que* $\langle\mathcal{A}_1, h_1\rangle$ *es un rebut para* $\langle\mathcal{A}_2, h_2\rangle$ *y el sub-argumento de desacuerdo en este ataque es* $\langle\mathcal{A}, h\rangle$*. Si* $\langle\mathcal{A}_1, h_1\rangle$ *es además un undermine para* $\langle\mathcal{A}_2, h_2\rangle$*, entonces* $(\langle\mathcal{A}_1, h_1\rangle, \langle\mathcal{A}, h\rangle) \in \mathbb{U}_{m\mathcal{P}}$*. En caso contrario, si* $\langle\mathcal{A}_1, h_1\rangle$ *no es un undermine para* $\langle\mathcal{A}_2, h_2\rangle$*, entonces* $(\langle\mathcal{A}_1, h_1\rangle, \langle\mathcal{A}, h\rangle) \in \mathbb{R}_{b\mathcal{P}}$*.*
- *Sean* $\langle\mathcal{A}, r\rangle, \langle\mathcal{B}, h\rangle \in \mathbb{A}_{\mathcal{P}}$ *tales que* $\langle\mathcal{A}, r\rangle$ *es un undercut para* $\langle\mathcal{B}, h\rangle$ *y el sub-argumento de desacuerdo en este ataque es* $\langle\mathcal{C}, q\rangle$*. En este caso, tenemos que* $(\langle\mathcal{A}, r\rangle, \langle\mathcal{C}, q\rangle) \in \mathbb{U}_{c\mathcal{P}}$*.*

- $\mathbb{S}_{\mathcal{P}} = \mathbb{B}_{k\mathcal{P}} \cup \mathbb{S}_{\sqsubseteq\mathcal{P}}$ *es obtenida a partir de la Definición 5.13 y la Proposición 5.1, de la siguiente manera*[3]*:*

- *Sean* $\langle\mathcal{B}, r\rangle, \langle\mathcal{A}, h\rangle \in \mathbb{A}_{\mathcal{P}}$ *tales que* $\langle\mathcal{B}, r\rangle$ *es un argumento de backing para la regla rebatible* r *y es un sub-argumento propio de* $\langle\mathcal{A}, h\rangle$*. Si* r *es una regla rebatible conclusiva de* $\langle\mathcal{A}, h\rangle$*, entonces* $(\langle\mathcal{B}, r\rangle, \langle\mathcal{A}, h\rangle) \in \mathbb{B}_{k\mathcal{P}}$*.*
- *Sean* $\langle\mathcal{A}, h\rangle, \langle\mathcal{B}, q\rangle \in \mathbb{A}_{\mathcal{P}}$ *tales que* $\langle\mathcal{A}, h\rangle$ *es un sub-argumento propio de* $\langle\mathcal{B}, q\rangle$*. Si* $\langle\mathcal{A}, h\rangle$ *es un argumento conclusivo tal que el literal* h *forma parte de la secuencia de derivación que conduce a la construcción de* $\langle\mathcal{B}, q\rangle$ *y no existe un sub-argumento propio* $\langle\mathcal{C}, p\rangle$ *de* $\langle\mathcal{B}, q\rangle$ *tal que* $\langle\mathcal{A}, h\rangle$ *es un sub-argumento propio de* $\langle\mathcal{C}, p\rangle$*, entonces* $(\langle\mathcal{A}, h\rangle, \langle\mathcal{B}, q\rangle) \in \mathbb{S}_{\sqsubseteq\mathcal{P}}$*.*

- $\preceq_{\mathcal{P}}$ *es obtenida a partir del criterio de comparación de E-DeLP, de la siguiente manera. Sean* $\langle\mathcal{A}_1, h_1\rangle, \langle\mathcal{A}_2, h_2\rangle \in \mathbb{A}_{\mathcal{P}}$*:*

- *Si* $\langle\mathcal{A}_1, h_1\rangle$ *es estrictamente más preferido que* $\langle\mathcal{A}_2, h_2\rangle$ *de acuerdo al criterio de comparación de E-DeLP, entonces* $(\langle\mathcal{A}_2, h_2\rangle, \langle\mathcal{A}_1, h_1\rangle) \in \preceq_{\mathcal{P}}$*.*
- *Si* $\langle\mathcal{A}_2, h_2\rangle$ *es estrictamente más preferido que* $\langle\mathcal{A}_1, h_1\rangle$ *de acuerdo al criterio de comparación de E-DeLP, entonces* $(\langle\mathcal{A}_1, h_1\rangle, \langle\mathcal{A}_2, h_2\rangle) \in \preceq_{\mathcal{P}}$*.*
- *Si* $\langle\mathcal{A}_1, h_1\rangle$ *y* $\langle\mathcal{A}_2, h_2\rangle$ *son igualmente preferidos de acuerdo al criterio de comparación de E-DeLP, entonces* $(\langle\mathcal{A}_2, h_2\rangle, \langle\mathcal{A}_1, h_1\rangle) \in \preceq_{\mathcal{P}}$ *y* $(\langle\mathcal{A}_1, h_1\rangle, \langle\mathcal{A}_2, h_2\rangle) \in \preceq_{\mathcal{P}}$*.*
- *Si* $\langle\mathcal{A}_1, h_1\rangle$ *y* $\langle\mathcal{A}_2, h_2\rangle$ *son incomparables de acuerdo al criterio de comparación de E-DeLP, entonces* $(\langle\mathcal{A}_2, h_2\rangle, \langle\mathcal{A}_1, h_1\rangle) \notin \preceq_{\mathcal{P}}$ *y* $(\langle\mathcal{A}_1, h_1\rangle, \langle\mathcal{A}_2, h_2\rangle) \notin \preceq_{\mathcal{P}}$*.*

[3]Recordemos que, por la Proposición 5.1, todos los sub-argumentos propios de un argumento construido a partir de un programa E-DeLP son argumentos conclusivos o argumentos de backing.

El siguiente ejemplo ilustra la obtención del BUAF asociado a un programa E-DeLP de acuerdo a la transformación caracterizada por la Definición 5.19.

Ejemplo 5.16 *Sea $\mathcal{P}_{5.5}$ el programa lógico rebatible extendido del Ejemplo 5.5. El marco argumentativo con backing y undercutting asociado a $\mathcal{P}_{5.5}$ es* $\mathsf{BUAF}_{\mathcal{P}_{5.5}} = \langle \mathbb{A}_{\mathcal{P}_{5.5}}, \mathbb{R}_{\mathcal{P}_{5.5}}, \mathbb{S}_{\mathcal{P}_{5.5}}, \preceq_{\mathcal{P}_{5.5}} \rangle$*, donde:*

- $$\mathbb{A}_{\mathcal{P}_{5.5}} = \left\{ \begin{array}{l} \langle \mathcal{A}_1, habitac\_iluminada(h) \rangle \\ \langle \mathcal{A}_2, \sim habitac\_iluminada(h) \rangle \\ \langle \mathcal{A}_3, luz\_encendida(l) \rangle \\ \langle \mathcal{A}_4, luz\_encendida(l) \prec llave\_encendida(l) \rangle_b \\ \langle \mathcal{A}_5, luz\_encendida(l) \prec llave\_encendida(l) \rangle_u \\ \langle \emptyset, llave\_encendida(l) \rangle \\ \langle \emptyset, noche \rangle \\ \langle \emptyset, lampara\_en\_habitac(l,h) \rangle \\ \langle \emptyset, electricidad \rangle \\ \langle \emptyset, lampara\_rota(l) \rangle \end{array} \right\}$$

  es tal que

$$\mathcal{A}_1 = \left\{ \begin{array}{l} habitac\_iluminada(h) \prec noche, lampara\_en\_habitac(l,h), luz\_encendida(l) \\ luz\_encendida(l) \prec llave\_encendida(l) \\ [luz\_encendida(l) \prec llave\_encendida(l)] \leftarrow\!\oplus\, [electricidad] \end{array} \right\}$$

$$\mathcal{A}_2 = \left\{ \sim habitac\_iluminada(h) \prec noche \right\}$$

$$\mathcal{A}_3 = \left\{ \begin{array}{l} luz\_encendida(l) \prec llave\_encendida(l) \\ [luz\_encendida(l) \prec llave\_encendida(l)] \leftarrow\!\oplus\, [electricidad] \end{array} \right\}$$

$$\mathcal{A}_4 = \left\{ [luz\_encendida(l) \prec llave\_encendida(l)] \leftarrow\!\oplus\, [electricidad] \right\}$$

$$\mathcal{A}_5 = \left\{ [luz\_encendida(l) \prec llave\_encendida(l)] \leftarrow\!\otimes\, [electricidad, lampara\_rota(l)] \right\}$$

Es decir, el conjunto $\mathbb{A}_{\mathcal{P}_{5.5}}$ contiene a los argumentos identificados en los Ejemplos 5.7 y 5.8.

- $\mathbb{R}_{\mathcal{P}_{5.5}} = \mathbb{R}_{b\mathcal{P}_{5.5}} \cup \mathbb{U}_{c\mathcal{P}_{5.5}} \cup \mathbb{U}_{m\mathcal{P}_{5.5}}$ *es tal que*

$$\mathbb{R}_{b\mathcal{P}_{5.5}} = \left\{ \begin{array}{l} (\langle \mathcal{A}_1, habitac\_iluminada(h)\rangle,\ \langle \mathcal{A}_2, {\sim}habitac\_iluminada(h)\rangle) \\ (\langle \mathcal{A}_2, {\sim}habitac\_iluminada(h)\rangle,\ \langle \mathcal{A}_1, habitac\_iluminada(h)\rangle) \end{array} \right\}$$

$$\mathbb{U}_{c\mathcal{P}_{5.5}} = \left\{ \ (\langle \mathcal{A}_5, luz\_encendida(l) \prec llave\_encendida(l)\rangle_u,\ \langle \mathcal{A}_3, luz\_encendida(l)\rangle) \ \right\}$$

$$\mathbb{U}_{m\mathcal{P}_{5.5}} = \emptyset$$

Por una parte, el subconjunto $\mathbb{R}_{b\mathcal{P}_{5.5}}$ contiene los ataques rebutting identificados en el Ejemplo 5.10. Por otra parte, el subconjunto $\mathbb{U}_{c\mathcal{P}_{5.5}}$ contiene el ataque undercutting identificado en el Ejemplo 5.12.

- $\mathbb{S}_{\mathcal{P}_{5.5}} = \mathbb{B}_{k\mathcal{P}_{5.5}} \cup \mathbb{S}_{\sqsubseteq\mathcal{P}_{5.5}}$ *es tal que*

$$\mathbb{B}_{k\mathcal{P}_{5.5}} = \left\{ \ (\langle \mathcal{A}_4, luz\_encendida(l) \prec llave\_encendida(l)\rangle_b,\ \langle \mathcal{A}_3, luz\_encendida(l)\rangle) \ \right\}$$

$$\mathbb{S}_{\sqsubseteq\mathcal{P}_{5.5}} = \left\{ \begin{array}{l} (\langle \mathcal{A}_3, luz\_encendida(l)\rangle,\ \langle \mathcal{A}_1, habitac\_iluminada(r)\rangle) \\ (\langle \emptyset, llave\_encendida(l)\rangle,\ \langle \mathcal{A}_4, luz\_encendida(l) \prec llave\_encendida(l)\rangle_b) \\ (\langle \emptyset, noche\rangle,\ \langle \mathcal{A}_1, habitac\_iluminada(h)\rangle) \\ (\langle \emptyset, noche\rangle,\ \langle \mathcal{A}_2, {\sim}habitac\_iluminada(h)\rangle) \\ (\langle \emptyset, lampara\_en\_habitac(l,h)\rangle,\ \langle \mathcal{A}_1, habitac\_iluminada(h)\rangle\,)) \\ (\langle \emptyset, electricidad\rangle,\ \langle \mathcal{A}_4, luz\_encendida(l) \prec llave\_encendida(l)\rangle_b) \\ (\langle \emptyset, electricidad\rangle,\ \langle \mathcal{A}_5, luz\_encendida(l) \prec llave\_encendida(l)\rangle_u) \\ (\langle \emptyset, lampara\_rota(l)\rangle,\ \langle \mathcal{A}_5, luz\_encendida(l) \prec llave\_encendida(l)\rangle_u) \end{array} \right\}$$

En este caso, el subconjunto $\mathbb{B}_{k\mathcal{P}_{5.5}}$ modela la relación existente entre el argumento de backing $\langle \mathcal{A}_4, luz\_encendida(l) \prec llave\_encendida(l)\rangle_b$ y el argumento conclusivo $\langle \mathcal{A}_3, luz\_encendida(l)\rangle$, por ser este último el super-argumento propio de $\langle \mathcal{A}_4, luz\_encendida(l) \prec llave\_encendida(l)\rangle_b$ cuya regla rebatible conclusiva es $luz\_encendida(l) \prec llave\_encendida(l)$. Luego, el subconjunto $\mathbb{S}_{\sqsubseteq\mathcal{P}_{5.5}}$ modela la relación de sub-argumento existente entre $\langle \mathcal{A}_3, luz\_encendida(l)\rangle$ y $\langle \mathcal{A}_1, habitac\_iluminada(h)\rangle$, así como también aquellas resultantes de considerar los argumentos para los hechos del programa $\mathcal{P}_{5.5}$.

- *Supongamos que el criterio de comparación de argumentos para E-DeLP es tal que $\langle \mathcal{A}_1, habitac\_iluminada(h)\rangle$ es estrictamente más preferido que*

$$
\begin{tikzcd}
\mathcal{A}rg_2 \arrow[r, tail, shift left, "R_b"] & \mathcal{A}rg_1 \arrow[l, tail, shift left, "R_b"] & \mathcal{A}rg_3 \arrow[l, Rightarrow, "\sqsubseteq"'] & \mathcal{A}rg_4 \arrow[l, Rightarrow, "\sqsubseteq"'] & \mathcal{A}rg_6 \arrow[l, Rightarrow, "\sqsubseteq"'] \\
\mathcal{A}rg_7 \arrow[u, Rightarrow, "\sqsubseteq"] \arrow[ur, Rightarrow, "\sqsubseteq"] & \mathcal{A}rg_8 \arrow[u, Rightarrow, "\sqsubseteq"] & \mathcal{A}rg_5 \arrow[u, tail, "U_c"] & \mathcal{A}rg_9 \arrow[l, Rightarrow, "\sqsubseteq"'] \arrow[u, Rightarrow, "\sqsubseteq"] & \\
 & & \mathcal{A}rg_{10} \arrow[u, Rightarrow, "\sqsubseteq"] & &
\end{tikzcd}
$$

Figura 5.4: Representación gráfica del $\mathsf{BUAF}_{\mathcal{P}_{5.5}}$ asociado al programa E-DeLP $\mathcal{P}_{5.5}$, correspondiente al Ejemplo 5.16.

$\langle \mathcal{A}_2, \sim habitac\_iluminada(h)\rangle$ *y* $\langle \mathcal{A}_5, luz\_encendida(l) \prec llave\_encendida(l)\rangle_u$ *es estrictamente más preferido que* $\langle \mathcal{A}_4, luz\_encendida(l) \prec llave\_encendida(l)\rangle_b$. *Además, supongamos que dicho criterio no establece otras relaciones de preferencia sobre los argumentos de* $\mathcal{P}_{5.5}$. *Entonces, la relación* $\preceq_{\mathcal{P}_{5.5}}$ *de* $\mathsf{BUAF}_{\mathcal{P}_{5.5}}$ *se define como sigue:*

$$
\preceq_{\mathcal{P}_{5.5}} = \{(\langle \mathcal{A}_2, \sim habitac\_iluminada(h)\rangle, \langle \mathcal{A}_1, habitac\_iluminada(h)\rangle), \\
(\langle \mathcal{A}_4, luz\_encendida(l) \prec llave\_encendida(l)\rangle_b, \langle \mathcal{A}_5, luz\_encendida(l) \prec llave\_encendida(l)\rangle_u)\}
$$

La Figura 5.4 ilustra la representación gráfica del $\mathsf{BUAF}_{\mathcal{P}_{5.5}}$, *utilizando la notación introducida en el Capítulo 4. Recordemos que, en dicha representación, la relación de ataque de un BUAF es denotada mediante flechas sólidas con cola '*$\rightarrowtail$*', respectivamente etiquetadas con '*R_b*', '*U_c*' y '*U_m*' según el tipo de ataque. De manera similar, la relación de soporte del BUAF se denota mediante flechas dobles '*$\Longrightarrow$*', respectivamente etiquetadas con '*b*' y '*$\sqsubseteq$*' para distinguir entre el soporte brindado por los backings y aquel provisto por los subargumentos. En este caso, para incrementar la legibilidad, los argumentos pertenecientes al conjunto* $\mathbb{A}_{\mathcal{P}_{5.5}}$ *son notados en la Figura 5.4 como* $\mathcal{A}rg_1, \mathcal{A}rg_2, \ldots, \mathcal{A}rg_n$, *donde:*

$$
\begin{aligned}
\mathcal{A}rg_1 &= \langle \mathcal{A}_1, habitac\_iluminada(h)\rangle \\
\mathcal{A}rg_2 &= \langle \mathcal{A}_2, \sim habitac\_iluminada(h)\rangle \\
\mathcal{A}rg_3 &= \langle \mathcal{A}_3, luz\_encendida(l)\rangle \\
\mathcal{A}rg_4 &= \langle \mathcal{A}_4, luz\_encendida(l) \prec llave\_encendida(l)\rangle_b \\
\mathcal{A}rg_5 &= \langle \mathcal{A}_5, luz\_encendida(l) \prec llave\_encendida(l)\rangle_u \\
\mathcal{A}rg_6 &= \langle \emptyset, llave\_encendida(l)\rangle \\
\mathcal{A}rg_7 &= \langle \emptyset, noche\rangle \\
\mathcal{A}rg_8 &= \langle \emptyset, lampara\_en\_habitac(l, h)\rangle
\end{aligned}
$$

$$\mathcal{A}rg_9 = \langle \emptyset, electricidad \rangle$$
$$\mathcal{A}rg_{10} = \langle \emptyset, lampara\_rota(l) \rangle$$

Dada la transformación caracterizada por la Definición 5.19, el BUAF asociado a un programa E-DeLP $\mathcal{P}$, llamado BUAF$_{\mathcal{P}}$, poseerá características especiales. Por una parte, como muestra la siguiente proposición, todo argumento conclusivo de BUAF$_{\mathcal{P}}$ poseerá una única regla rebatible conclusiva.

Proposición 5.2 *Sea* $\mathcal{P} = (\Pi, \Delta, \Sigma)$ *un programa lógico rebatible extendido y* $\langle \mathcal{A}, h \rangle$ *un argumento conclusivo construido a partir de* $\mathcal{P}$*. Si* r *es una regla rebatible conclusiva de* $\langle \mathcal{A}, h \rangle$*, entonces* $\nexists r' \in \Delta$ *tal que* $r' \neq r$ *y* r' *es una regla rebatible conclusiva de* $\langle \mathcal{A}, h \rangle$*.*

Demostración: ver Apéndice A.

Luego, la Proposición 5.3 establece que todo argumento del BUAF$_{\mathcal{P}}$ poseerá a lo sumo un argumento de backing.

Proposición 5.3 *Sea* $\mathcal{P}$ *un programa lógico rebatible extendido y sea* $BUAF_{\mathcal{P}} = \langle \mathbb{A}_{\mathcal{P}}, \mathbb{R}_{\mathcal{P}}, \mathbb{S}_{\mathcal{P}}, \preceq_{\mathcal{P}} \rangle$ *el marco argumentativo con backing y undercutting asociado a* $\mathcal{P}$*. Para todo argumento* $\langle \mathcal{A}, h \rangle \in \mathbb{A}_{\mathcal{P}}$ *vale que: si* $\exists \langle \mathcal{B}, r \rangle \in \mathbb{A}_{\mathcal{P}}$ *tal que* $(\langle \mathcal{B}, r \rangle, \langle \mathcal{A}, h \rangle) \in \mathbb{B}_{k\mathcal{P}}$*, entonces* $\nexists \langle \mathcal{C}, r' \rangle \in \mathbb{A}_{\mathcal{P}}$ *tal que* $\langle \mathcal{C}, r' \rangle \neq \langle \mathcal{B}, r \rangle$ *y* $(\langle \mathcal{C}, r' \rangle, \langle \mathcal{A}, h \rangle) \in \mathbb{B}_{k\mathcal{P}}$*.*

Demostración: ver Apéndice A.

Como se mencionó al comienzo de esta sección, las derrotas entre los argumentos construidos a partir de un programa E-DeLP serán obtenidas a partir de su BUAF asociado. De esta manera, una derrota entre argumentos del BUAF asociado constituirá una derrota en el marco del correspondiente programa E-DeLP.

Definición 5.20 (Derrota entre argumentos de E-DeLP) *Sea* $\mathcal{P}$ *un programa lógico rebatible extendido y sea* $BUAF_{\mathcal{P}} = \langle \mathbb{A}_{\mathcal{P}}, \mathbb{R}_{\mathcal{P}}, \mathbb{S}_{\mathcal{P}}, \preceq_{\mathcal{P}} \rangle$ *el marco argumentativo con backing y undercutting asociado a* $\mathcal{P}$*. Para todo par de argumentos* $\mathcal{A}, \mathcal{B} \in \mathbb{A}_{\mathcal{P}}$ *tales que* $\mathcal{A}$ *derrota a* $\mathcal{B}$ *en* $BUAF_{\mathcal{P}}$ *se tiene que* $\mathcal{A}$ derrota *a* $\mathcal{B}$ *en* $\mathcal{P}$*.*

Recordemos que, como se presentó en el Capítulo 4, la Definición 4.5 engloba los diferentes tipos de derrota que pueden ocurrir entre los argumentos de un BUAF: derrota primaria, derrota implícita y derrota indirecta (ver definiciones 4.2, 4.3 y 4.4). En particular, como se verá en la Sección 5.6.2, la naturaleza de las derrotas obtenidas a partir del BUAF asociado será tenida en cuenta para el cálculo de aceptabilidad de los argumentos mediante la construcción de árboles de dialéctica.

Ejemplo 5.17 *Sea $\mathcal{P}_{5.5}$ el programa lógico rebatible extendido del Ejemplo 5.5 y $\mathsf{BUAF}_{\mathcal{P}_{5.5}}$ el marco argumentativo con backing y undercutting asociado a $\mathcal{P}_{5.7}$, correspondiente al Ejemplo 5.16. A partir de las relaciones de ataque, soporte y preferencia de $\mathsf{BUAF}_{\mathcal{P}_{5.5}}$ se obtienen las siguientes derrotas entre argumentos de $\mathcal{P}_{5.5}$, las cuales se corresponden con derrotas ocurridas en $\mathsf{BUAF}_{\mathcal{P}_{5.5}}$:*

- $\langle \mathcal{A}_1, habitac\_iluminada(h)\rangle$ *derrota al argumento* $\langle \mathcal{A}_2, {\sim}habitac\_iluminada(h)\rangle$ *en* $\mathcal{P}_{5.5}$*, dado que existe una derrota primaria de* $\langle \mathcal{A}_1, habitac\_iluminada(h)\rangle$ *sobre* $\langle \mathcal{A}_2, {\sim}habitac\_iluminada(h)\rangle$ *en* $\mathsf{BUAF}_{\mathcal{P}_{5.5}}$.

- $\langle \mathcal{A}_5, luz\_encendida(l) \prec llave\_encendida(l)\rangle_u$ *derrota al argumento* $\langle \mathcal{A}_4, luz\_encendida(l) \prec llave\_encendida(l)\rangle_b$ *en* $\mathcal{P}_{5.5}$*, dado que existe una derrota implícita de* $\langle \mathcal{A}_5, luz\_encendida(l) \prec llave\_encendida(l)\rangle_u$ *sobre* $\langle \mathcal{A}_4, luz\_encendida(l) \prec llave\_encendida(l)\rangle_b$ *en* $\mathsf{BUAF}_{\mathcal{P}_{5.5}}$.

- $\langle \mathcal{A}_5, luz\_encendida(l) \prec llave\_encendida(l)\rangle_u$ *derrota al argumento* $\langle \mathcal{A}_3, luz\_encendida(l)\rangle$ *en* $\mathcal{P}_{5.5}$*, dado que existe una derrota primaria (e indirecta) de* $\langle \mathcal{A}_5, luz\_encendida(l) \prec llave\_encendida(l)\rangle_u$ *sobre* $\langle \mathcal{A}_3, luz\_encendida(l)\rangle$ *en* $\mathsf{BUAF}_{\mathcal{P}_{5.5}}$.

- $\langle \mathcal{A}_5, luz\_encendida(l) \prec llave\_encendida(l)\rangle_u$ *derrota al argumento* $\langle \mathcal{A}_1, habitac\_iluminada(h)\rangle$ *en* $\mathcal{P}_{5.5}$*, dado que existe una derrota indirecta de* $\langle \mathcal{A}_5, luz\_encendida(l) \prec llave\_encendida(l)\rangle_u$ *sobre* $\langle \mathcal{A}_1, habitac\_iluminada(h)\rangle$ *en* $\mathsf{BUAF}_{\mathcal{P}_{5.5}}$.

La Figura 5.5 ilustra el grafo de derrotas de $\mathsf{BUAF}_{\mathcal{P}_{5.5}}$, donde los nodos son los argumentos del BUAF y los arcos representan las derrotas entre ellos. En particular, la Figura 5.5 adopta la notación simplificada para los argumentos introducida en el Ejemplo 5.16.

$$\begin{tikzcd}
\mathcal{A}rg_2 & \mathcal{A}rg_1 \arrow[l] & \mathcal{A}rg_3 & \mathcal{A}rg_4 & \mathcal{A}rg_6 \\
\mathcal{A}rg_7 & \mathcal{A}rg_8 & \mathcal{A}rg_5 \arrow[ul] \arrow[u] \arrow[ur] & \mathcal{A}rg_9 & \\
 & & \mathcal{A}rg_{10} & &
\end{tikzcd}$$

Figura 5.5: Grafo de derrotas de $\mathsf{BUAF}_{\mathcal{P}_{5.5}}$, correspondiente al Ejemplo 5.17.

Habiendo caracterizado la noción de derrota en E-DeLP, podemos extender la representación gráfica introducida en la Sección 5.4 para considerar las derrotas entre los argumentos construidos a partir de un programa E-DeLP. En particular, siguiendo con la convención adoptada en esta tesis (y en consonancia con la representación gráfica utilizada por los BUAF), las derrotas entre argumentos serán denotadas mediante flechas sólidas. De esta manera, por ejemplo, la derrota de un argumento $\mathcal{A}$ sobre un argumento $\mathcal{B}$ será notada como $\mathcal{A} \longrightarrow \mathcal{B}$. Por último, para aquellos ataques que resultan en derrotas entre argumentos de E-DeLP, la representación gráfica incluirá únicamente las flechas sólidas correspondientes a las derrotas. Es decir, si un argumento $\mathcal{A}$ ataca a un argumento $\mathcal{B}$ (notado $\mathcal{A} \dashrightarrow \mathcal{B}$) y, en particular, lo derrota (notado $\mathcal{A} \longrightarrow \mathcal{B}$), la representación gráfica incluirá únicamente la flecha sólida de $\mathcal{A}$ hacia $\mathcal{B}$. Para ilustrar esta representación extendida, consideremos el siguiente ejemplo.

Ejemplo 5.18 *Sea $\mathcal{P}_{5.5}$ el programa lógico rebatible extendido del Ejemplo 5.5. Los argumentos construidos a partir de $\mathcal{P}_{5.5}$ y los ataques entre ellos se encuentran ilustrados en la Figura 5.2 del Ejemplo 5.14. Dadas las derrotas identificadas en el Ejemplo 5.17, obtenemos la representación gráfica extendida para los argumentos, ataques y derrotas de $\mathcal{P}_{5.5}$, la cual se ilustra en la Figura 5.6.*

Nótese que, por la Definición 5.20, las derrotas entre argumentos de E-DeLP no necesariamente surgen de ataques entre argumentos. Tal es el caso de la derrota de $\langle \mathcal{A}_5, luz\_encendida(l) \prec llave\_encendida(l) \rangle_u$ sobre $\langle \mathcal{A}_4, luz\_encendida(l) \prec llave\_encendida(l) \rangle_b$, la cual se corresponde con una derrota implícita en el $\mathsf{BUAF}_{\mathcal{P}_{5.5}}$ asociado a $\mathcal{P}_{5.5}$.

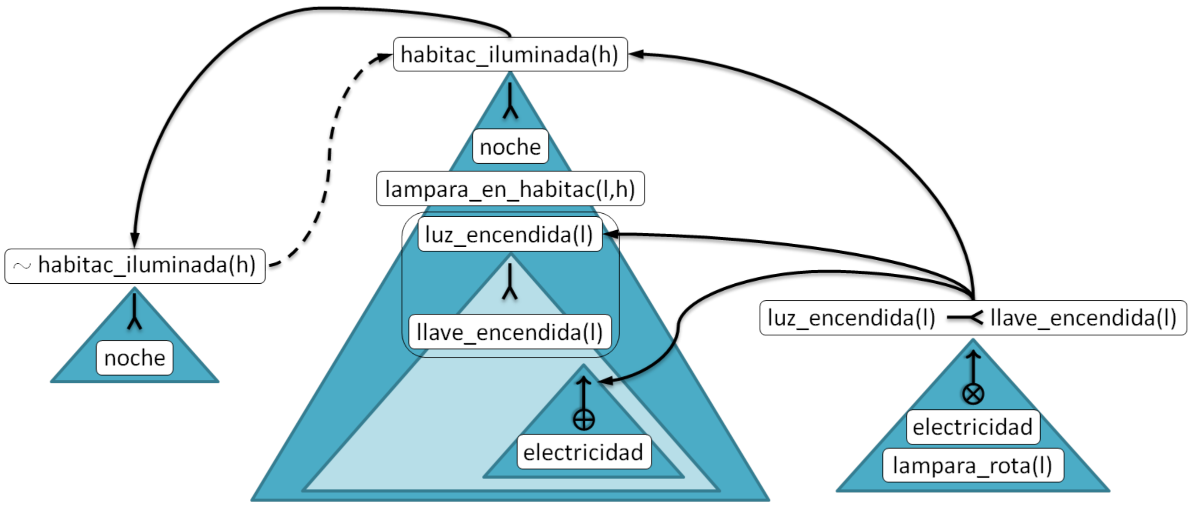


Figura 5.6: Representación gráfica de argumentos, ataques y derrotas del Ejemplo 5.18.

5.6. Aceptabilidad en E-DeLP: argumentos aceptados y literales garantizados

En esta sección se abordará el cálculo de aceptabilidad para los argumentos de E-DeLP, correspondiente al quinto elemento de la estructura conceptual presentada en la Sección 2.1. Luego, una vez identificados los argumentos aceptados, será posible determinar cuáles son las inferencias del sistema. Estas inferencias corresponderán a literales que son conclusiones de argumentos aceptados, a los cuales denominaremos *literales garantizados* a partir de un programa E-DeLP.

Como se vio en el Capítulo 2, en la literatura existen diferentes propuestas para efectuar el cálculo de aceptabilidad de los argumentos de un sistema argumentativo, tales como la definición de semánticas de aceptabilidad [Dun95] o la especificación de procedimientos de prueba dialécticos [GS04]. No obstante esto, independientemente de la aproximación adoptada, las bases para efectuar cálculo de aceptabilidad residen en la consideración de una noción de derrota. Por lo tanto, para determinar el estado de aceptabilidad de los argumentos de E-DeLP será necesario considerar las derrotas entre dichos argumentos, las cuales, como se vio en la Sección 5.5, son obtenidas a partir del BUAF asociado a un programa E-DeLP.

Cabe destacar que la obtención de las derrotas entre los argumentos de E-DeLP mediante la construcción de su BUAF asociado conlleva un importante beneficio. Como se vio en el Capítulo 4, los BUAF son marcos generales y abstractos sobre los cuales es posible aplicar cualquier tipo de semántica de aceptabilidad entre aquellas existentes en la literatura de marcos argumentativos abstractos [Dun95, BCG11]. Por lo tanto, esto brinda la posibilidad de resolver el cálculo de aceptabilidad para E-DeLP utilizando el mecanismo definido para los BUAF en el Capítulo 4.

En esta sección se abordará el cálculo de aceptabilidad para los argumentos de E-DeLP siguiendo dos aproximaciones alternativas. Por una parte, aprovechando los desarrollos del Capítulo 4, se efectuará el cálculo de aceptabilidad en E-DeLP aplicando semánticas de aceptabilidad sobre el BUAF asociado a un programa lógico rebatible extendido. Cabe destacar que esta alternativa permitirá obtener todos los conjuntos de argumentos aceptados con respecto a cada semántica. Es decir, dado un programa E-DeLP $\mathcal{P}$ y una semántica de aceptabilidad s, se identificarán todos los argumentos construidos a partir de $\mathcal{P}$ que pertenecen a las extensiones de $\mathcal{P}$ bajo la semántica s. Por otra parte, siguiendo el enfoque orientado a consultas de DeLP (ver Sección 2.2.2), se presentará una aproximación procedural que define un modelo de prueba dialéctico para E-DeLP. Siguiendo esta aproximación, el cálculo de aceptabilidad para los argumentos de E-DeLP será efectuado por demanda, en función de las consultas realizadas, mediante la construcción de árboles de dialéctica. Por lo tanto, este proceso no involucrará el cálculo de aceptabilidad para todos los argumentos del sistema, sino que solamente determinará el estado de aceptabilidad de los argumentos asociados a la consulta realizada. Finalmente, independientemente de la aproximación adoptada, una vez identificados los argumentos aceptados será posible determinar las inferencias del sistema. En particular, estas se identificarán mediante los *literales garantizados*, los cuales corresponderán a conclusiones de los argumentos aceptados.

5.6.1. Cálculo de Aceptabilidad utilizando Semánticas

En esta sección abordaremos el cálculo de aceptabilidad para los argumentos construidos a partir de un programa E-DeLP, valiéndonos del BUAF asociado a dicho programa. Como se mencionó anteriormente, la obtención de las derrotas entre los argumentos resulta esencial para determinar el estado de aceptabilidad de los mismos. Como se verá a continuación, dado un programa E-DeLP $\mathcal{P}$, a partir del $\mathsf{BUAF}_{\mathcal{P}}$ asociado a $\mathcal{P}$ y, en particular, de las derrotas obtenidas sobre los argumentos de $\mathsf{BUAF}_{\mathcal{P}}$, será posible computar la aceptabilidad de los argumentos de $\mathcal{P}$.

En la Sección 5.5 se mostró que dado un programa E-DeLP $\mathcal{P}$ y su $\mathsf{BUAF}_{\mathcal{P}}$ asociado, los argumentos de $\mathsf{BUAF}_{\mathcal{P}}$ son, en particular, los argumentos construidos a partir de $\mathcal{P}$. Luego, dada esta correspondencia, se mostró que las derrotas entre los argumentos de $\mathcal{P}$ corresponden a derrotas sobre los argumentos de $\mathsf{BUAF}_{\mathcal{P}}$. Por lo tanto, dado que los argumentos y las derrotas entre argumentos de $\mathcal{P}$ y $\mathsf{BUAF}_{\mathcal{P}}$ coinciden, el estado de aceptabilidad de los argumentos de $\mathcal{P}$ se hallará determinado por su correspondiente estado de aceptabilidad en el $\mathsf{BUAF}_{\mathcal{P}}$ asociado.

Siguiendo la aproximación presentada en el Capítulo 4, consideraremos las cuatro semánticas de aceptabilidad propuestas por Dung en [Dun95]: *completa*, *preferida*, *estable* y *grounded*. Cada una de estas semánticas caracteriza un conjunto de *extensiones*, las cuales corresponden a conjuntos de argumentos que verifican ciertas propiedades y pueden ser aceptados conjuntamente. De esta manera, un argumento construido a partir de un programa E-DeLP $\mathcal{P}$ estará aceptado con respecto a una determinada semántica si y sólo si el argumento pertenece a una de las extensiones del $\mathsf{BUAF}_{\mathcal{P}}$ asociado a $\mathcal{P}$, de acuerdo a esa semántica. Esta intuición se halla formalizada por la siguiente definición.

Definición 5.21 (Extensiones) *Sea $\mathcal{P}$ un programa lógico rebatible extendido, $\mathsf{BUAF}_{\mathcal{P}}$ el marco argumentativo con backing y undercutting asociado a $\mathcal{P}$, y S un conjunto de argumentos construidos a partir de $\mathcal{P}$:*

- *S es una* extensión completa *de $\mathcal{P}$ si y sólo si S es una extensión completa de $\mathsf{BUAF}_{\mathcal{P}}$.*
- *S es una* extensión preferida *de $\mathcal{P}$ si y sólo si S es una extensión preferida de $\mathsf{BUAF}_{\mathcal{P}}$.*
- *S es una* extensión estable *de $\mathcal{P}$ si y sólo si S es una extensión estable de $\mathsf{BUAF}_{\mathcal{P}}$.*
- *S es la* extensión grounded *de $\mathcal{P}$ si y sólo si S es la extensión grounded de $\mathsf{BUAF}_{\mathcal{P}}$.*

En otras palabras, los conjuntos de argumentos construidos a partir de un programa E-DeLP $\mathcal{P}$ que se aceptarán en forma conjunta serán aquellos correspondientes a su marco argumentativo con backing y undercutting asociado $\mathsf{BUAF}_{\mathcal{P}}$. Para ilustrar esto, consideremos el siguiente ejemplo.

Ejemplo 5.19 *Sea $\mathcal{P}_{5.5}$ el programa lógico rebatible extendido del Ejemplo 5.5 y $\mathsf{BUAF}_{\mathcal{P}_{5.5}}$ su marco argumentativo con backing y undercutting asociado, correspondiente al Ejemplo 5.16. Dadas las derrotas identificadas entre los argumentos de $\mathcal{P}_{5.5}$ (las cuales coinciden con las derrotas entre los argumentos de $\mathsf{BUAF}_{\mathcal{P}_{5.5}}$), ilustradas en el Ejemplo 5.17, podemos obtener las siguientes extensiones de $\mathcal{P}_{5.5}$:*

- *la extensión completa $\{\mathcal{A}rg_2, \mathcal{A}rg_5, \mathcal{A}rg_6, \mathcal{A}rg_7, \mathcal{A}rg_8, \mathcal{A}rg_9, \mathcal{A}rg_{10}\}$;*
- *la extensión preferida $\{\mathcal{A}rg_2, \mathcal{A}rg_5, \mathcal{A}rg_6, \mathcal{A}rg_7, \mathcal{A}rg_8, \mathcal{A}rg_9, \mathcal{A}rg_{10}\}$;*
- *la extensión estable $\{\mathcal{A}rg_2, \mathcal{A}rg_5, \mathcal{A}rg_6, \mathcal{A}rg_7, \mathcal{A}rg_8, \mathcal{A}rg_9, \mathcal{A}rg_{10}\}$; y*
- *la extensión grounded $\{\mathcal{A}rg_2, \mathcal{A}rg_5, \mathcal{A}rg_6, \mathcal{A}rg_7, \mathcal{A}rg_8, \mathcal{A}rg_9, \mathcal{A}rg_{10}\}$.*

Nótese que la extensión obtenida bajo cualquiera de estas cuatro semánticas es única y, en particular, es la misma. Esto se debe a que las derrotas entre los argumentos de $\mathcal{P}_{5.5}$ y $\mathsf{BUAF}_{\mathcal{P}_{5.5}}$ son tal que $\mathcal{A}rg_5$, quien no posee derrotadores, defiende a $\mathcal{A}rg_2$ de la derrota originada por $\mathcal{A}rg_1$. Asimismo, $\mathcal{A}rg_5$ también derrota a $\mathcal{A}rg_3$ y $\mathcal{A}rg_4$, motivo por el cual los argumentos $\mathcal{A}rg_1$, $\mathcal{A}rg_3$ y $\mathcal{A}rg_4$ no pertenecen a la extensión. Finalmente, utilizando la notación expandida para los argumentos de $\mathcal{P}_{5.5}$, tenemos que la única extensión completa, estable, preferida y grounded de $\mathcal{P}_{5.5}$ es

$$\mathsf{E}_{5.5} = \left\{ \begin{array}{l} \langle \mathcal{A}_2, \sim habitac\_iluminada(h) \rangle \\ \langle \mathcal{A}_5, luz\_encendida(l) \prec llave\_encendida(l) \rangle_u \\ \langle \emptyset, llave\_encendida(l) \rangle \\ \langle \emptyset, noche \rangle \\ \langle \emptyset, lampara\_en\_habitac(l, h) \rangle \\ \langle \emptyset, electricidad \rangle \\ \langle \emptyset, lampara\_rota(l) \rangle \end{array} \right\}, \textit{ donde}$$

$$\mathcal{A}_2 = \left\{ \ \sim habitac\_iluminada(h) \prec noche \ \right\}$$

$$\mathcal{A}_5 = \left\{ \ [luz\_encendida(l) \prec llave\_encendida(l)] \leftarrow\!\otimes [electricidad, lampara\_rota(l)] \ \right\}$$

Como se vio en la Sección 4.4, la aplicación de semánticas de aceptabilidad alternativas puede conducir a la obtención de diferentes conjuntos de extensiones. Por una parte, un conjunto de argumentos E puede constituir una extensión bajo una semántica s, pero no bajo otra semántica s'. Más aún, puede ocurrir que un argumento $\mathcal{A}$ pertenezca a una extensión E obtenida bajo una semántica s, pero no pertenezca a otra extensión E' obtenida bajo la misma semántica s (ver pág. 138, Ejemplo 4.12). Por este motivo, siguiendo la aproximación presentada en la Sección 4.4, nos interesará distinguir aquellos argumentos que, dada una semántica s: *i*) pertenecen a *toda* extensión obtenida bajo s, *ii*) pertenecen a *alguna pero no a toda* extensión obtenida bajo s, o *iii*) no pertenecen a *ninguna* extensión obtenida bajo s.

Definición 5.22 (Argumento Aceptado y Argumento Rechazado) *Sea $\mathcal{P}$ un programa lógico rebatible extendido, $\mathsf{s} \in \{$completa, preferida, estable, grounded$\}$ una semántica de aceptabilidad, y $\mathcal{A}$ un argumento construido a partir de $\mathcal{P}$:*

- *El argumento $\mathcal{A}$ está* escépticamente aceptado *con respecto a s si y sólo si pertenece a toda extensión de $\mathcal{P}$ bajo la semántica s.*
- *El argumento $\mathcal{A}$ está* crédulamente aceptado *con respecto a s si y sólo si pertenece a alguna extensión de $\mathcal{P}$ (pero no a todas) bajo la semántica s.*
- *El argumento $\mathcal{A}$ está* rechazado *con respecto a s si y sólo si no pertenece a ninguna extensión de $\mathcal{P}$ bajo la semántica s.*

Ejemplo 5.20 *Sea $\mathcal{P}_{5.5}$ el programa lógico rebatible extendido del Ejemplo 5.5. En el Ejemplo 5.19 se mostró que $\mathcal{P}_{5.5}$ posee una única extensión, la cual corresponde a las semánticas completa, preferida, estable y grounded. Luego, como dicha extensión es única, tendremos que todo argumento perteneciente a ella estará escépticamente aceptado con respecto a cualquiera de estas cuatro semánticas.*

Finalmente, dado un conjunto de argumentos aceptados a partir de un programa E-DeLP con respecto a una determinada semántica, será posible determinar qué literales son inferidos a partir de dicho programa. Intuitivamente, dado un programa E-DeLP $\mathcal{P}$ tal que el argumento conclusivo $\langle \mathcal{A}, h \rangle$ está aceptado a partir de $\mathcal{P}$ con respecto a una determinada semántica, el literal "h" formará parte de las inferencias de $\mathcal{P}$ bajo esa semántica. No obstante esto, resulta necesario considerar los diferentes tipos de aceptación caracterizados por la Definición 5.22, para así evitar inferir literales en desacuerdo. Para ilustrar esto, consideremos el siguiente ejemplo.

Ejemplo 5.21 *Sea* $\mathcal{P}_{5.21} = (\Pi_{5.21}, \Delta_{5.21}, \emptyset)$ *un programa lógico rebatible extendido, donde* $\Pi_{5.21} = \{a, b\}$ *y* $\Delta_{5.21} = \{(l \prec a),\ (\sim l \prec b)\}$*. A partir de* $\mathcal{P}_{5.21}$ *es posible construir los siguientes argumentos:*

- $\langle \emptyset, a \rangle$*;*
- $\langle \emptyset, b \rangle$*;*
- $\langle \mathcal{A}_1, l \rangle$*, con* $\mathcal{A}_1 = \{l \prec a\}$*; y*
- $\langle \mathcal{A}_2, \sim l \rangle$*, con* $\mathcal{A}_2 = \{\sim l \prec b\}$*.*

En este caso, $\langle \mathcal{A}_1, l \rangle$ *es un rebut para* $\langle \mathcal{A}_2, \sim l \rangle$ *y viceversa. Supongamos que el criterio de comparación adoptado es tal que* $\langle \mathcal{A}_1, l \rangle$ *y* $\langle \mathcal{A}_2, \sim l \rangle$ *son igualmente preferidos. Luego, tenemos que los argumentos* $\langle \mathcal{A}_1, l \rangle$ *y* $\langle \mathcal{A}_2, \sim l \rangle$ *se derrotan entre sí, siendo estas las únicas derrotas entre los argumentos de* $\mathcal{P}_{5.21}$*. Finalmente, considerando la semántica preferida, obtenemos las siguientes extensiones de* $\mathcal{P}_{5.21}$*:* $\mathsf{E}_{5.21} = \{\langle \emptyset, a \rangle, \langle \emptyset, b \rangle, \langle \mathcal{A}_1, l \rangle\}$ *y* $\mathsf{E}'_{5.21} = \{\langle \emptyset, a \rangle, \langle \emptyset, b \rangle, \langle \mathcal{A}_2, \sim l \rangle\}$*. En consecuencia, los argumentos* $\langle \emptyset, a \rangle$ *y* $\langle \emptyset, b \rangle$ *están escépticamente aceptados con respecto a la semántica preferida, mientras que* $\langle \mathcal{A}_1, l \rangle$ *y* $\langle \mathcal{A}_2, \sim l \rangle$ *están crédulamente aceptados con respecto a dicha semántica.*

A partir de la situación ilustrada en el Ejemplo 5.21, si consideráramos las inferencias del sistema de acuerdo a la aceptación crédula, los literales en desacuerdo "l" y "$\sim l$" (los cuales son, en particular, complementarios) formarían parte de las inferencias de $\mathcal{P}_{5.21}$, lo cual claramente es una situación indeseada.

La siguiente definición formaliza las condiciones bajo las cuales se infiere un literal en E-DeLP. De esta manera, se introduce la noción de *literal garantizado* a partir de un programa E-DeLP bajo una determinada semántica de aceptabilidad.

Definición 5.23 (Literal Garantizado bajo una Semántica) *Sea* $\mathcal{P}$ *un programa lógico rebatible extendido,* L *un literal y* $\mathsf{s} \in \{$*completa, preferida, estable, grounded*$\}$ *una semántica de aceptabilidad.* L *es un* literal garantizado a partir de $\mathcal{P}$ bajo la semántica s *si y sólo si existe un argumento conclusivo* $\langle \mathcal{A}, L \rangle$ *construido a partir de* $\mathcal{P}$ *tal que* $\langle \mathcal{A}, L \rangle$ *está escépticamente aceptado con respecto a la semántica* s*.*

Ejemplo 5.22 *Sea $\mathcal{P}_{5.5}$ el programa lógico rebatible extendido del Ejemplo 5.5. Como se mostró en el Ejemplo 5.19, la única extensión completa, preferida, estable y grounded de $\mathcal{P}_{5.5}$ es*

$$\mathsf{E}_{5.5} = \left\{ \begin{array}{l} \langle \mathcal{A}_2, {\sim}habitac\_iluminada(h) \rangle \\ \langle \mathcal{A}_5, luz\_encendida(l) \prec llave\_encendida(l) \rangle_u \\ \langle \emptyset, llave\_encendida(l) \rangle \\ \langle \emptyset, noche \rangle \\ \langle \emptyset, lampara\_en\_habitac(l, h) \rangle \\ \langle \emptyset, electricidad \rangle \\ \langle \emptyset, lampara\_rota(l) \rangle \end{array} \right\}$$

Asimismo, como se mostró en el Ejemplo 5.20, todos los argumentos pertenecientes a $\mathsf{E}_{5.5}$ se hallan escépticamente aceptados con respecto a las semánticas completa, preferida, estable y grounded. Por lo tanto, los literales garantizados a partir de $\mathcal{P}_{5.5}$ bajo las semánticas completa, preferida, estable y grounded son los siguientes:

- $\sim habitac\_iluminada(h)$;
- $llave\_encendida(l)$;
- *noche;*
- $lampara\_en\_habitac(l, h)$;
- *electricidad; y*
- $lampara\_rota(l)$

Finalmente, a continuación se presentará una serie de propiedades satisfechas por E-DeLP, las cuales corresponden a características deseables de un sistema argumentativo. Estas propiedades conciernen la aceptación de los argumentos y las inferencias del sistema. En primer lugar, se mostrará que los argumentos construidos a partir del conocimiento estricto de un programa E-DeLP estarán aceptados y, en consecuencia, sus conclusiones estarán garantizadas. Luego, se mostrará que no se aceptarán simultáneamente argumentos que se derroten entre sí. Finalmente, se mostrará que las inferencias obtenidas a partir de un programa E-DeLP son consistentes, dado que no se garantizarán literales en desacuerdo.

La siguiente proposición muestra que, dado un programa E-DeLP, todo argumento construido a partir de una derivación estricta estará aceptado bajo cualquier semántica de aceptabilidad.

Proposición 5.4 *Sea $\mathcal{P}$ un programa lógico rebatible extendido y $\mathsf{s} \in \{$completa, preferida, estable, grounded$\}$ una semántica de aceptabilidad. Si $\langle \emptyset, L \rangle$ es un argumento construido a partir de $\mathcal{P}$, entonces $\langle \emptyset, L \rangle$ está escépticamente aceptado con respecto a la semántica s.*

Demostración: ver Apéndice A.

La siguiente proposición hace extensivo el resultado anterior hacia las conclusiones de los argumentos. De este modo, todo literal derivado a partir del conocimiento estricto de un programa E-DeLP estará garantizado bajo cualquier semántica de aceptabilidad.

Proposición 5.5 *Sea $\mathcal{P}$ un programa lógico rebatible extendido y $\mathsf{s} \in \{$completa, preferida, estable, grounded$\}$ una semántica de aceptabilidad. Si $\langle \emptyset, L \rangle$ es un argumento construido a partir de $\mathcal{P}$, entonces L es un literal garantizado a partir de $\mathcal{P}$ bajo la semántica s.*

Demostración: ver Apéndice A.

Otra propiedad satisfecha por E-DeLP, es que no se aceptarán simultáneamente argumentos que se derroten entre sí.

Proposición 5.6 *Sea $\mathcal{P}$ un programa lógico rebatible extendido y $\mathsf{s} \in \{$completa, preferida, estable, grounded$\}$ una semántica de aceptabilidad. Si $\langle \mathcal{A}, h \rangle$ es un argumento de $\mathcal{P}$ aceptado escépticamente con respecto a la semántica s, entonces para todo argumento $\langle \mathcal{B}, q \rangle$ tal que $\langle \mathcal{A}, h \rangle$ derrota a $\langle \mathcal{B}, q \rangle$ en $\mathcal{P}$ o $\langle \mathcal{B}, q \rangle$ derrota a $\langle \mathcal{A}, h \rangle$ en $\mathcal{P}$, se verifica que $\langle \mathcal{B}, q \rangle$ está rechazado con respecto a la semántica s.*

Demostración: ver Apéndice A.

El siguiente resultado corresponde a una propiedad fundamental acerca de las inferencias en E-DeLP. Concretamente, el siguiente teorema muestra que a partir de un programa E-DeLP no se garantizarán literales en desacuerdo.

Teorema 5.1 *Sea $\mathcal{P}$ un programa lógico rebatible extendido y* $\mathsf{s} \in \{$*completa, preferida, estable, grounded*$\}$ *una semántica de aceptabilidad. Si L es un literal garantizado a partir de $\mathcal{P}$ bajo la semántica s, entonces para todo literal L' tal que L y L' están en desacuerdo se verifica que L' no es un literal garantizado a partir de $\mathcal{P}$ bajo la semántica s.*

Demostración: ver Apéndice A.

5.6.2. Cálculo de Aceptabilidad utilizando Árboles de Dialéctica

El modelo para el cálculo de aceptabilidad en E-DeLP presentado en la sección anterior sigue la aproximación extensional de [Dun95]. Como se vio en el Capítulo 2, existen aproximaciones alternativas para determinar el estado de aceptación de los argumentos de un sistema. En particular, en esta sección se presentará una aproximación dialéctica para el cálculo de aceptabilidad en E-DeLP, tomando como base la propuesta de DeLP presentada en la Sección 2.2.2.6.

En la Sección 2.2.2.6 se introdujeron los árboles de dialéctica, los cuales proveen un mecanismo procedural para efectuar el cálculo de aceptabilidad sobre los argumentos. En general, este procedimiento consiste en la simulación de todas las líneas de discusión entre un proponente y un oponente enfrentados. El proponente comienza proponiendo un argumento $\mathcal{A}$, para el cual se determinará el estado de aceptación, y luego tanto oponente como proponente van presentando sucesivamente argumentos derrotadores en respuesta a los argumentos de su contrincante. En particular, esta discusión da lugar a una estructura arbórea que relaciona argumentos y derrotas. Finalmente, el estado de aceptación de $\mathcal{A}$ es determinado luego de analizar la discusión en forma exhaustiva.

A continuación se introducirá un procedimiento para determinar el estado de aceptabilidad de los argumentos construidos a partir de un programa E-DeLP. Este procedimiento involucrará la construcción de *árboles de dialéctica*, reflejando de esta manera la naturaleza dialéctica del proceso de razonamiento argumentativo. Finalmente, al igual que en la aproximación presentada en la Sección 5.6.1, una vez identificados los argumentos aceptados a partir de un programa E-DeLP será posible determinar las inferencias del sistema, correspondientes a las conclusiones de los argumentos aceptados.

En primer lugar introduciremos la noción de *línea argumentativa* en E-DeLP. Al igual que en DeLP, una línea argumentativa en E-DeLP representa una línea de discusión entre un proponente y un oponente para un determinado argumento $\mathcal{A}$. Concretamente, se trata de una secuencia de argumentos, comenzando por $\mathcal{A}$, donde cada argumento que aparece en una posición impar (incluido $\mathcal{A}$) se considera como presentado por el proponente y, por lo tanto, a favor de $\mathcal{A}$. Análogamente, cada argumento que ocupa una posición par en la secuencia se considera como presentado por el oponente y, por lo tanto, en contra de $\mathcal{A}$. En este sentido, los argumentos presentados por el proponente conforman el *conjunto de argumentos de apoyo* de la línea argumentativa, mientras que los argumentos presentados por el oponente conforman el *conjunto de argumentos de interferencia* de la misma. Finalmente, la incorporación de argumentos de apoyo e interferencia en la línea argumentativa surge de la consideración de la relación de derrota: cada argumento perteneciente a la secuencia es un derrotador de su predecesor.

Definición 5.24 (Línea Argumentativa) *Sea $\mathcal{P}$ un programa lógico rebatible extendido y $\langle \mathcal{A}_0, h_0 \rangle$ un argumento construido a partir de $\mathcal{P}$. Una* línea argumentativa *para $\langle \mathcal{A}_0, h_0 \rangle$ es una secuencia de argumentos construidos a partir de $\mathcal{P}$, denotada $\Lambda = [\langle \mathcal{A}_0, h_0 \rangle, \langle \mathcal{A}_1, h_1 \rangle, \langle \mathcal{A}_2, h_2 \rangle, \langle \mathcal{A}_3, h_3 \rangle, \ldots]$, donde cada argumento $\langle \mathcal{A}_i, h_i \rangle$ perteneciente a la secuencia es un derrotador del argumento predecesor $\langle \mathcal{A}_{i-1}, h_{i-1} \rangle$.*

En particular, el conjunto $\Lambda_A = \{\langle \mathcal{A}_0, h_0 \rangle, \langle \mathcal{A}_2, h_2 \rangle, \ldots\}$ es llamado conjunto de argumentos de apoyo de Λ*, mientras que el conjunto $\Lambda_I = \{\langle \mathcal{A}_1, h_1 \rangle, \langle \mathcal{A}_3, h_3 \rangle, \ldots\}$ corresponde al* conjunto de argumentos de interferencia de Λ*.*

Para evitar secuencias de argumentos que pueden conducir a cadenas de razonamiento falaz, se introduce una serie de restricciones sobre las líneas argumentativas, caracterizando así a las *líneas argumentativas aceptables*. Siguiendo la aproximación presentada en la Sección 2.2.2.6, una línea argumentativa será aceptable si satisface las siguientes restricciones: es una secuencia finita, no contiene argumentos o sub-argumentos repetidos, los conjuntos de argumentos de apoyo y de interferencia son consistentes, y no posee dos derrotas por bloqueo consecutivas.

A continuación analizaremos cómo caracterizar estas restricciones sobre las líneas argumentativas aceptables en E-DeLP. Con respecto a la primer restricción, al especificar el último elemento de una línea argumentativa se está determinando que la misma posee longitud finita. Por otra parte, la noción de sub-argumento caracterizada en la Definición 5.13 permite establecer la segunda restricción sobre las líneas argumentativas de E-DeLP. En este sentido, se evitará introducir en una línea argumentativa aceptable aquellos argumentos que sean sub-argumentos de argumentos que aparecen con anterioridad en la línea.

Con respecto a la restricción de consistencia de los conjuntos de apoyo e interferencia, se extenderá la restricción presentada en la Sección 2.2.2.6 con el objetivo de considerar la existencia de reglas de backing y undercutting. En este sentido, el conjunto de argumentos de apoyo (respectivamente, de interferencia) será consistente la si unión de sus reglas es un conjunto no contradictorio y no contiene reglas de backing y undercutting para la misma regla rebatible. Por último, consideremos la restricción que establece que una línea argumentativa aceptable no debe incluir argumentos que den lugar a dos derrotas por bloqueo consecutivas.

Como se vio en la Sección 2.2.2.5, la noción de derrota en DeLP es definida de manera tal que caracteriza los ataques exitosos, distinguiendo además entre *derrota propia* y *derrota por bloqueo*. En particular, esta distinción es establecida en función de las preferencias entre los argumentos involucrados en el ataque, de acuerdo al criterio de comparación adoptado. Dado que DeLP considera únicamente ataques de tipo rebutting (ver Definición 2.22), el criterio de comparación es utilizado para determinar la preferencia del argumento atacante por sobre el argumento atacado. Luego, dado un ataque de un argumento $\mathcal{A}$ por sobre un argumento $\mathcal{B}$, donde $\mathcal{C}$ es el sub-argumento atacado de $\mathcal{B}$ (*i. e.*, el sub-argumento de desacuerdo), tenemos dos casos: si $\mathcal{A}$ es estrictamente preferido sobre $\mathcal{C}$, ocurre una derrota propia de $\mathcal{A}$ sobre $\mathcal{B}$; en contraste, si $\mathcal{A}$ y $\mathcal{C}$ son igualmente preferidos o resultan incomparables de acuerdo al criterio de comparación, ocurre una derrota por bloqueo de $\mathcal{A}$ sobre $\mathcal{B}$.

En contraste a lo realizado por DeLP, las derrotas en E-DeLP son obtenidas a partir de la especificación del BUAF asociado a un programa E-DeLP (ver definiciones 5.19 y 5.20). Es por esto que, en primera instancia, no es posible identificar cuáles de estas derrotas corresponden a derrotas propias y cuáles corresponden a derrotas por bloqueo. No obstante esto, como se mostrará a continuación, a partir de la especificación del $\mathsf{BUAF}_{\mathcal{P}}$ asociado a un programa E-DeLP $\mathcal{P}$ y de los diferentes tipos de derrota obtenidos sobre el $\mathsf{BUAF}_{\mathcal{P}}$ (derrota primaria, derrota implícita y derrota indirecta), será posible distinguir las derrotas propias y las derrotas por bloqueo entre los argumentos de $\mathcal{P}$ (los cuales son, en particular, argumentos de $\mathsf{BUAF}_{\mathcal{P}}$).

En la Sección 4.3 se presentaron tres tipos de derrota que pueden ocurrir entre los argumentos de un BUAF: derrota primaria (Definición 4.2), derrota implícita (Definición 4.3) y derrota indirecta (Definición 4.4). La derrota primaria determina el éxito de un ataque rebutting, undercutting o undermining especificado en la relación de ataque del BUAF. Las derrotas implícitas capturan el conflicto existente entre argumentos de backing y undercutting para un mismo argumento. Por último, las derrotas implícitas tienen por objetivo la propagación de otras derrotas a través de las relaciones de backing y sub-argumento del BUAF.

En el caso de la derrota primaria, para determinar si se trata de una derrota propia o una derrota por bloqueo, debemos analizar el tipo del ataque que dio origen a la misma. Si se trata de un ataque rebutting o undermining, bastará con analizar la preferencia entre el argumento atacante y el argumento atacado, de acuerdo a lo establecido por la Definición 4.2. De esta manera, dados dos argumentos $\mathcal{A}$ y $\mathcal{B}$ tales que $\mathcal{A}$ es un rebut o un undermine para $\mathcal{B}$ en el BUAF, si la relación de preferencia del BUAF es tal que $\mathcal{A}$ es estrictamente preferido sobre $\mathcal{B}$ (*i. e.*, $\mathcal{B} \prec \mathcal{A}$), tendremos que $\mathcal{A}$ es un *derrotador propio* de $\mathcal{B}$. Similarmente, si la relación de preferencia del BUAF es tal que $\mathcal{A}$ y $\mathcal{B}$ son igualmente preferidos o incomparables (*i. e.*, $\mathcal{A} \equiv \mathcal{B}$ o $\mathcal{A} \bowtie \mathcal{B}$), tendremos que $\mathcal{A}$ es un *derrotador por bloqueo* de $\mathcal{B}$. Por otra parte, si la derrota primaria fue originada por un ataque de tipo undercutting, deberá considerarse la preferencia entre el argumento atacante y los posibles argumentos de backing para el argumento atacado, dando así lugar a dos alternativas. Dados dos argumentos $\mathcal{A}$ y $\mathcal{B}$ tales que $\mathcal{A}$ es un undercut para $\mathcal{B}$ y no existen argumentos de backing para $\mathcal{B}$, tendremos que $\mathcal{A}$ es un *derrotador propio* de $\mathcal{B}$. En contraste, si existen argumentos de backing para $\mathcal{B}$, la naturaleza de la derrota quedará determinada por la relación de preferencia entre el undercut $\mathcal{A}$ y el backing $\mathcal{C}$ de $\mathcal{B}$ que dio origen a la derrota[4]. Luego, si $\mathcal{A}$ es estrictamente preferido sobre $\mathcal{C}$ (*i. e.*, $\mathcal{C} \prec \mathcal{A}$), tendremos que $\mathcal{A}$ es un *derrotador propio* de $\mathcal{B}$. Asimismo, si $\mathcal{A}$ y $\mathcal{C}$ son igualmente preferidos o incomparables (*i. e.*, $\mathcal{A} \equiv \mathcal{C}$ o $\mathcal{A} \bowtie \mathcal{C}$), tendremos que $\mathcal{A}$ es un *derrotador por bloqueo* de $\mathcal{B}$.

Consideremos ahora la caracterización de las derrotas implícitas del BUAF, de acuerdo a lo establecido por la Definición 4.3. Como se mencionó anteriormente, este tipo de derrota captura el conflicto existente entre argumentos de backing y undercutting para un mismo argumento. Por lo tanto, dados dos argumentos $\mathcal{A}$ y $\mathcal{B}$ tales que $\mathcal{A}$ derrota implícitamente a $\mathcal{B}$ en el BUAF, para determinar si se trata de una derrota propia o una derrota por bloqueo bastará con analizar la preferencia entre dichos argumentos. Luego, si $\mathcal{A}$ es estrictamente preferido sobre $\mathcal{B}$ de acuerdo a la relación de preferencia del BUAF

[4]Recordemos que, de acuerdo a lo establecido por la Definición 4.2, en este caso, el ataque undercutting de $\mathcal{A}$ sobre $\mathcal{B}$ resultará en derrota si existe un backing $\mathcal{C}$ de $\mathcal{B}$ tal que $\mathcal{C}$ no es estrictamente preferido sobre $\mathcal{A}$ (*i. e.*, $\mathcal{A} \not\prec \mathcal{C}$).

(*i. e.*, $\mathcal{B} \prec \mathcal{A}$), tendremos que $\mathcal{A}$ es un *derrotador propio* de $\mathcal{B}$. Análogamente, si $\mathcal{A}$ y $\mathcal{B}$ son igualmente preferidos o incomparables (*i. e.*, $\mathcal{A} \equiv \mathcal{B}$ o $\mathcal{A} \bowtie \mathcal{B}$), tendremos que $\mathcal{A}$ es un *derrotador por bloqueo* de $\mathcal{B}$.

Para el caso de las derrotas indirectas del BUAF, caracterizadas por la Definición 4.4, el análisis resulta más sencillo. Dado que este tipo de derrota tiene por objetivo propagar las derrotas sobre los backings y sub-argumentos hacia los argumentos que estos soportan, la naturaleza de una derrota indirecta quedará determinada por la naturaleza de la derrota que la originó. Teniendo en cuenta esto, consideremos los argumentos $\mathcal{A}, \mathcal{B}$ y $\mathcal{C}$ tales que $\mathcal{A}$ derrota a $\mathcal{C}$, $\mathcal{C}$ soporta a $\mathcal{B}$ (ya sea mediante la relación de backing o sub-argumento) y, en consecuencia, $\mathcal{A}$ derrota indirectamente a $\mathcal{B}$ en el BUAF. Luego, si $\mathcal{A}$ es un derrotador propio de $\mathcal{C}$, tendremos que $\mathcal{A}$ es un *derrotador propio* de $\mathcal{B}$. En caso contrario, si $\mathcal{A}$ es un derrotador por bloqueo de $\mathcal{C}$, tendremos que $\mathcal{A}$ es un *derrotador por bloqueo* de $\mathcal{B}$.

Las intuiciones arriba presentadas se encuentran formalizadas en la siguiente definición, la cual caracteriza las nociones de derrota propia y derrota por bloqueo en el contexto de un programa E-DeLP.

Definición 5.25 (Derrota Propia y Derrota por Bloqueo) *Sea $\mathcal{P}$ un programa lógico rebatible extendido y $BUAF_{\mathcal{P}} = \langle \mathbb{A}_{\mathcal{P}}, \mathbb{R}_{\mathcal{P}}, \mathbb{S}_{\mathcal{P}}, \preceq_{\mathcal{P}} \rangle$ el marco argumentativo con backing y undercutting asociado a $\mathcal{P}$. Dados dos argumentos $\mathcal{A}$ y $\mathcal{B}$ tales que $\mathcal{A}$ derrota a $\mathcal{B}$ en $\mathcal{P}$:*

- *Si $\mathcal{A}$ derrota primariamente a $\mathcal{B}$ en $BUAF_{\mathcal{P}}$ y $(\mathcal{A}, \mathcal{B}) \in \mathbb{R}_{b\mathcal{P}} \cup \mathbb{U}_{m\mathcal{P}}$: si $\mathcal{B} \prec_{\mathcal{P}} \mathcal{A}$, entonces $\mathcal{A}$ es un* derrotador propio *de $\mathcal{B}$; en caso contrario, si $\mathcal{A} \equiv \mathcal{B}$ o $\mathcal{A} \bowtie \mathcal{B}$, entonces $\mathcal{A}$ es un* derrotador por bloqueo *de $\mathcal{B}$.*

- *Si $\mathcal{A}$ derrota primariamente a $\mathcal{B}$, $(\mathcal{A}, \mathcal{B}) \in \mathbb{U}_{c\mathcal{P}}$ y $\nexists \mathcal{C} \in \mathbb{A}_{\mathcal{P}}$ tal que $(\mathcal{C}, \mathcal{B}) \in \mathbb{B}_{k\mathcal{P}}$, entonces $\mathcal{A}$ es un* derrotador propio *de $\mathcal{B}$.*

- *Si $\mathcal{A}$ derrota primariamente a $\mathcal{B}$, $(\mathcal{A}, \mathcal{B}) \in \mathbb{U}_{c\mathcal{P}}$, y $\exists \mathcal{C} \in \mathbb{A}_{\mathcal{P}}$ tal que $(\mathcal{C}, \mathcal{B}) \in \mathbb{B}_{k\mathcal{P}}$: si $\mathcal{C} \prec_{\mathcal{P}} \mathcal{A}$, entonces $\mathcal{A}$ es un* derrotador propio *de $\mathcal{B}$; en caso contrario, si $\mathcal{A} \equiv \mathcal{C}$ o $\mathcal{A} \bowtie \mathcal{C}$, entonces $\mathcal{A}$ es un* derrotador por bloqueo *de $\mathcal{B}$.*

- *Si $\mathcal{A}$ derrota implícitamente a $\mathcal{B}$: si $\mathcal{B} \prec_{\mathcal{P}} \mathcal{A}$, entonces $\mathcal{A}$ es un* derrotador propio *de $\mathcal{B}$; en caso contrario, si $\mathcal{A} \equiv \mathcal{B}$ o $\mathcal{A} \bowtie \mathcal{B}$, entonces $\mathcal{A}$ es un* derrotador por bloqueo *de $\mathcal{B}$.*

- *Si $\mathcal{A}$ derrota indirectamente a $\mathcal{B}$ y $\exists \mathbb{C} \in \mathbb{A}_{\mathcal{P}}$ tal que $(\mathcal{C}, \mathcal{B}) \in \mathbb{S}_{\mathcal{P}}$ y $\mathcal{A}$ derrota a $\mathcal{C}$: si $\mathcal{A}$ es un derrotador propio de $\mathcal{C}$, entonces $\mathcal{A}$ es un* derrotador propio *de $\mathcal{B}$; en caso contrario, si $\mathcal{A}$ es un derrotador por bloqueo de $\mathcal{C}$, entonces $\mathcal{A}$ es un derrotador por bloqueo de $\mathcal{B}$.*

Cabe destacar que, dada la Definición 5.25, podría ocurrir que un argumento $\mathcal{A}$ sea, simultáneamente, un derrotador propio y por bloqueo de un argumento $\mathcal{B}$. En particular, si consideramos el último caso de la Definición 5.25 donde $\mathcal{A}$ derrota indirectamente a $\mathcal{B}$, pueden existir dos (o más) argumentos $\mathcal{C}_1$ y $\mathcal{C}_2$ que soportan a $\mathcal{B}$ y son derrotados por $\mathcal{A}$. Luego, podría ocurrir que $\mathcal{A}$ sea un derrotador propio de $\mathcal{C}_1$, mientras que $\mathcal{A}$ sea un derrotador por bloqueo de $\mathcal{C}_2$. En consecuencia, tendríamos que $\mathcal{A}$ es tanto un derrotador propio como un derrotador por bloqueo de $\mathcal{B}$. Ante situaciones como esta, se considerará que la derrota propia de $\mathcal{A}$ sobre $\mathcal{B}$ prevalece ante la derrota por bloqueo. Esto se debe a que, de acuerdo con el *principio del enlace más débil* (en inglés, *weakest link principle*) [Pol01], un argumento no puede ser más fuerte que el más débil de sus sub-argumentos. Esta intuición se verá reflejada en la siguiente definición, al especificar la restricción que impide incluir argumentos que den lugar a dos derrotas por bloqueo consecutivas en las líneas argumentativas aceptables de E-DeLP.

Definición 5.26 (Línea Argumentativa Aceptable) *Sea $\mathcal{P} = (\Pi, \Delta, \Sigma)$ un programa lógico rebatible extendido y $\Lambda = [\langle \mathcal{A}_0, h_0 \rangle, \ldots, \langle \mathcal{A}_n, h_n \rangle]$ una línea argumentativa para $\langle \mathcal{A}_0, h_0 \rangle$ obtenida a partir de $\mathcal{P}$. Diremos que Λ es una* línea argumentativa aceptable *si y sólo si satisface las siguientes condiciones:*

1. *Λ es una secuencia finita;*

2. *el conjunto Λ_A de argumentos de apoyo de Λ y el conjunto Λ_I de argumentos de interferencia de Λ son tal que $R_{\Lambda_A} = \Pi \cup \{\mathcal{A} \mid \langle \mathcal{A}, p \rangle \in \Lambda_A\}$ y $R_{\Lambda_I} = \Pi \cup \{\mathcal{B} \mid \langle \mathcal{B}, q \rangle \in \Lambda_I\}$ son conjuntos no contradictorios y verifican lo siguiente: si R_{Λ_A} (respectivamente, R_{Λ_I}) contiene una regla rebatible R o una regla de backing para R, entonces no contiene una regla de undercutting para R;*

3. *ningún argumento $\langle \mathcal{A}_k, h_k \rangle$ de Λ es un sub-argumento de un argumento $\langle \mathcal{A}_j, h_j \rangle$ que aparece previamente en Λ $(j < k)$; y*

4. para todo argumento $\langle \mathcal{A}_i, h_i \rangle$ de Λ tal que $\langle \mathcal{A}_i, h_i \rangle$ es un derrotador por bloqueo de $\langle \mathcal{A}_{i-1}, h_{i-1} \rangle$ y no es un derrotador propio de $\langle \mathcal{A}_{i-1}, h_{i-1} \rangle$, si existe $\langle \mathcal{A}_{i+1}, h_{i+1} \rangle$ en Λ, entonces $\langle \mathcal{A}_{i+1}, h_{i+1} \rangle$ es un derrotador propio de $\langle \mathcal{A}_i, h_i \rangle$.

La noción de línea argumentativa aceptable en E-DeLP se halla ilustrada por el siguiente ejemplo.

Ejemplo 5.23 *Sea $\mathcal{P}_{5.5}$ el programa lógico rebatible extendido del Ejemplo 5.5. El Ejemplo 5.17 ilustra las derrotas existentes entre los argumentos construidos a partir de $\mathcal{P}_{5.5}$. En particular, todas estas derrotas son derrotas propias, de acuerdo a lo establecido por la Definición 5.25. A continuación se listan todas las líneas argumentativas aceptables para los argumentos de $\mathcal{P}_{5.5}$:*

- *Para el argumento $\langle \mathcal{A}_1, habitac\_iluminada(h) \rangle$:*
  - $\Lambda_{1.1} = [\langle \mathcal{A}_1, habitac\_iluminada(h) \rangle]$
  - $\Lambda_{1.2} = [\langle \mathcal{A}_1, habitac\_iluminada(h) \rangle, \langle \mathcal{A}_5, luz\_encendida(l) \prec llave\_encendida(l) \rangle_u]$
- *Para el argumento $\langle \mathcal{A}_2, {\sim} habitac\_iluminada(h) \rangle$:*
  - $\Lambda_{2.1} = [\langle \mathcal{A}_2, {\sim} habitac\_iluminada(h) \rangle]$
  - $\Lambda_{2.2} = [\langle \mathcal{A}_2, {\sim} habitac\_iluminada(h) \rangle, \langle \mathcal{A}_1, habitac\_iluminada(h) \rangle]$
  - $\Lambda_{2.3} = [\langle \mathcal{A}_2, {\sim} habitac\_iluminada(h) \rangle, \langle \mathcal{A}_1, habitac\_iluminada(h) \rangle,$
    $\langle \mathcal{A}_5, luz\_encendida(l) \prec llave\_encendida(l) \rangle_u]$
- *Para el argumento $\langle \mathcal{A}_3, luz\_encendida(l) \rangle$:*
  - $\Lambda_{3.1} = [\langle \mathcal{A}_3, luz\_encendida(l) \rangle]$
  - $\Lambda_{3.2} = [\langle \mathcal{A}_3, luz\_encendida(l) \rangle, \langle \mathcal{A}_5, luz\_encendida(l) \prec llave\_encendida(l) \rangle_u]$
- *Para el argumento $\langle \mathcal{A}_4, luz\_encendida(l) \prec llave\_encendida(l) \rangle_b$:*
  - $\Lambda_{4.1} = [\langle \mathcal{A}_4, luz\_encendida(l) \prec llave\_encendida(l) \rangle_b$
  - $\Lambda_{4.2} = [\langle \mathcal{A}_4, luz\_encendida(l) \prec llave\_encendida(l) \rangle_b,$
    $\langle \mathcal{A}_5, luz\_encendida(l) \prec llave\_encendida(l) \rangle_u]$
- *Para el argumento $\langle \mathcal{A}_5, luz\_encendida(l) \prec llave\_encendida(l) \rangle_u$:*

- $\Lambda_5 = [\langle \mathcal{A}_5, luz\_encendida(l) \multimap llave\_encendida(l) \rangle_u]$

- *Para el argumento* $\langle \emptyset, llave\_encendida(l) \rangle$*:*

  - $\Lambda_6 = [\langle \emptyset, llave\_encendida(l) \rangle]$

- *Para el argumento* $\langle \emptyset, noche \rangle$*:*

  - $\Lambda_7 = [\langle \emptyset, noche \rangle]$

- *Para el argumento* $\langle \emptyset, lampara\_en\_habitac(l, h) \rangle$*:*

  - $\Lambda_8 = [\langle \emptyset, lampara\_en\_habitac(l, h) \rangle]$

- *Para el argumento* $\langle \emptyset, electricidad \rangle$*:*

  - $\Lambda_9 = [\langle \emptyset, electricidad \rangle]$

- *Para el argumento* $\langle \emptyset, lampara\_rota(l) \rangle$*:*

  - $\Lambda_{10} = [\langle \emptyset, lampara\_rota(l) \rangle]$

A partir de la noción de línea argumentativa aceptable, la siguiente definición formaliza la noción de *árbol de dialéctica* en E-DeLP. Básicamente, al igual que en DeLP, un árbol de dialéctica considera simultáneamente y exhaustivamente todas las líneas de discusión para un argumento de E-DeLP. Es decir, dado un argumento $\mathcal{A}$ construido a partir de un programa E-DeLP, el árbol de dialéctica para $\mathcal{A}$ agrupa todas las líneas argumentativas aceptables maximales para $\mathcal{A}$.

Definición 5.27 (Árbol de Dialéctica) *Sea $\mathcal{P}$ un programa lógico rebatible extendido y $\langle \mathcal{A}_0, h_0 \rangle$ un argumento construido a partir de $\mathcal{P}$. Un* árbol de dialéctica para $\langle \mathcal{A}_0, h_0 \rangle$, *notado $\mathcal{T}_{\langle \mathcal{A}_0, h_0 \rangle}$, es un árbol tal que los nodos son argumentos construidos a partir de $\mathcal{P}$, los arcos representan derrotas entre los nodos, y se satisfacen las siguientes condiciones:*

1. *$\langle \mathcal{A}_0, h_0 \rangle$ es la raíz de $\mathcal{T}_{\langle \mathcal{A}_0, h_0 \rangle}$; y*

2. *si $\Lambda = [\langle \mathcal{A}_0, h_0 \rangle, \langle \mathcal{A}_1, h_1 \rangle, \ldots, \langle \mathcal{A}_n, h_n \rangle]$ es una línea argumentativa aceptable para $\langle \mathcal{A}_0, h_0 \rangle$ y no existe $\langle \mathcal{A}_m, h_m \rangle$ tal que $\Lambda' = [\langle \mathcal{A}_0, h_0 \rangle, \langle \mathcal{A}_1, h_1 \rangle, \ldots, \langle \mathcal{A}_n, h_n \rangle, \langle \mathcal{A}_m, h_m \rangle]$ es una línea argumentativa aceptable para $\langle \mathcal{A}_0, h_0 \rangle$, entonces $\langle \mathcal{A}_n, h_n \rangle$ es una hoja de $\mathcal{T}_{\langle \mathcal{A}_0, h_0 \rangle}$ y Λ es un camino desde la raíz $\langle \mathcal{A}_0, h_0 \rangle$ hasta la hoja $\langle \mathcal{A}_n, h_n \rangle$.*

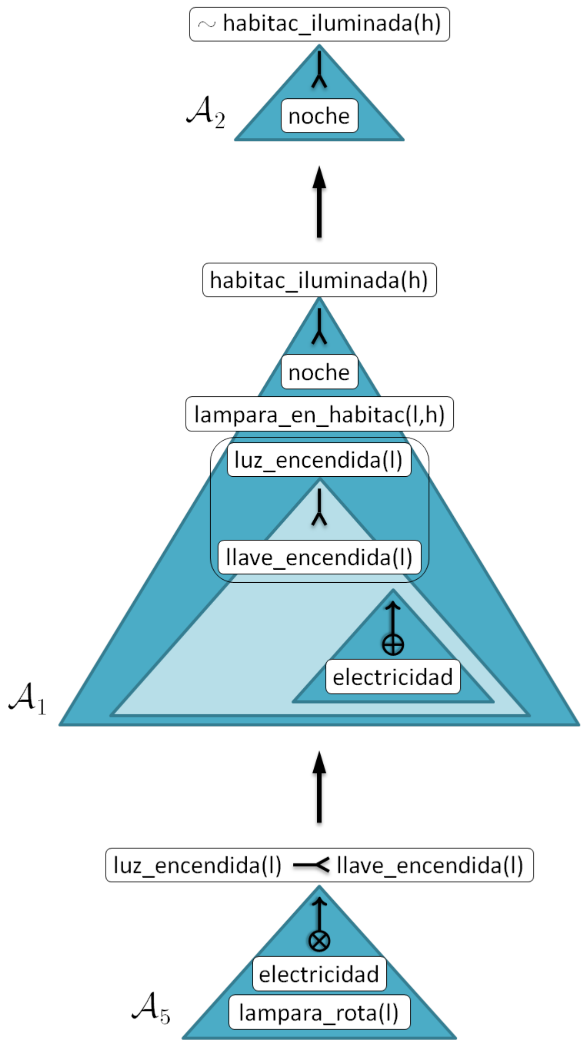


Figura 5.7: Árbol de dialéctica $\mathcal{T}_{\langle \mathcal{A}_2, \sim habitac\_iluminada(h) \rangle}$ del Ejemplo 5.24.

Ejemplo 5.24 *Sea $\mathcal{P}_{5.5}$ el programa lógico rebatible extendido del Ejemplo 5.5. El Ejemplo 5.23 lista todas las líneas argumentativas aceptables para los argumentos de $\mathcal{P}_{5.5}$. En particular, si consideramos las líneas aceptables para el argumento $\langle \mathcal{A}_2, \sim habitac\_iluminada(h) \rangle$, obtenemos el árbol de dialéctica $\mathcal{T}_{\langle \mathcal{A}_2, \sim habitac\_iluminada(h) \rangle}$ que se ilustra en la Figura 5.7. Observando las líneas argumentativas aceptables para $\langle \mathcal{A}_2, \sim habitac\_iluminada(h) \rangle$, ilustradas en el Ejemplo 5.23, podemos notar que la línea $\Lambda_{2.3}$ contiene a las líneas $\Lambda_{2.1}$ y $\Lambda_{2.2}$. Por lo tanto, el árbol de dialéctica $\mathcal{T}_{\langle \mathcal{A}_2, \sim habitac\_iluminada(h) \rangle}$ coincide con la línea argumentativa aceptable $\Lambda_{2.3}$.*

Siguiendo la aproximación presentada en la Sección 2.2.2.6, la siguiente definición caracteriza el proceso de marcado de un árbol de dialéctica en E-DeLP. De este modo, será posible determinar el estado de aceptabilidad del argumento correspondiente a la raíz del árbol.

Definición 5.28 (Árbol de Dialéctica Marcado) *Sea $\mathcal{P}$ un programa lógico rebatible extendido y $\mathcal{T}_{\langle \mathcal{A}, h\rangle}$ un árbol de dialéctica para el argumento $\langle \mathcal{A}, h\rangle$ construido a partir de $\mathcal{P}$. El* árbol de dialéctica marcado para $\langle \mathcal{A}, h\rangle$, *notado $\mathcal{T}^*_{\langle \mathcal{A}, h\rangle}$, se obtiene marcando cada nodo $\mathcal{N}$ de $\mathcal{T}_{\langle \mathcal{A}, h\rangle}$ como* **D** *(derrotado) o* **U** *(no derrotado) de acuerdo al siguiente criterio:*

- *Si $\mathcal{N}$ es un nodo hoja de $\mathcal{T}_{\langle \mathcal{A}, h\rangle}$, entonces $\mathcal{N}$ es marcado como* **U** *en $\mathcal{T}^*_{\langle \mathcal{A}, h\rangle}$.*
- *Si $\mathcal{N}$ es un nodo interno de $\mathcal{T}_{\langle \mathcal{A}, h\rangle}$ y posee algún hijo marcado como* **U** *en $\mathcal{T}^*_{\langle \mathcal{A}, h\rangle}$, entonces $\mathcal{N}$ es marcado como* **D** *en $\mathcal{T}^*_{\langle \mathcal{A}, h\rangle}$.*
- *Si $\mathcal{N}$ es un nodo interno de $\mathcal{T}_{\langle \mathcal{A}, h\rangle}$ y todos sus hijos son marcados como* **D** *en $\mathcal{T}^*_{\langle \mathcal{A}, h\rangle}$, entonces $\mathcal{N}$ es marcado como* **U** *en $\mathcal{T}^*_{\langle \mathcal{A}, h\rangle}$.*

Ejemplo 5.25 *Consideremos el árbol de dialéctica $\mathcal{T}_{\langle \mathcal{A}_2, \sim habitac\_iluminada(h)\rangle}$ del Ejemplo 5.24, el cual se halla ilustrado en la Figura 5.7. El árbol de dialéctica marcado para $\langle \mathcal{A}_2, \sim habitac\_iluminada(h)\rangle$ es $\mathcal{T}^*_{\langle \mathcal{A}_2, \sim habitac\_iluminada(h)\rangle}$, ilustrado en la Figura 5.8. En este caso, el argumento $\langle \mathcal{A}_5, luz\_encendida(l) \prec llave\_encendida(l)\rangle_u$ está marcado como* **U***, el argumento $\langle \mathcal{A}_1, habitac\_iluminada(h)\rangle$ está marcado como* **D***, y el argumento $\langle \mathcal{A}_2, \sim habitac\_iluminada(h)\rangle$ está marcado como* **U***.*

Dado un argumento $\langle \mathcal{A}, h\rangle$ construido a partir de un programa E-DeLP $\mathcal{P}$, el árbol de dialéctica marcado $\mathcal{T}^*_{\langle \mathcal{A}, h\rangle}$ representa el análisis dialéctico que considera todos los argumentos relevantes de $\mathcal{P}$ a fin de decidir el estado de aceptabilidad de $\langle \mathcal{A}, h\rangle$. Por lo tanto, a partir del marcado de la raíz de $\mathcal{T}^*_{\langle \mathcal{A}, h\rangle}$ es posible determinar el estado de aceptabilidad de $\langle \mathcal{A}, h\rangle$, como se indica en la siguiente definición.

Definición 5.29 (Argumento Aceptado y Argumento Rechazado) *Sea $\mathcal{P}$ un programa lógico rebatible extendido y $\langle \mathcal{A}, h\rangle$ un argumento construido a partir de $\mathcal{P}$. Diremos que $\langle \mathcal{A}, h\rangle$ es un* argumento aceptado *de $\mathcal{P}$ si y sólo si la raíz del árbol de dialéctica marcado $\mathcal{T}^*_{\langle \mathcal{A}, h\rangle}$ está marcada como* **U***. En caso contrario, diremos que $\langle \mathcal{A}, h\rangle$ es un* argumento rechazado *de $\mathcal{P}$.*

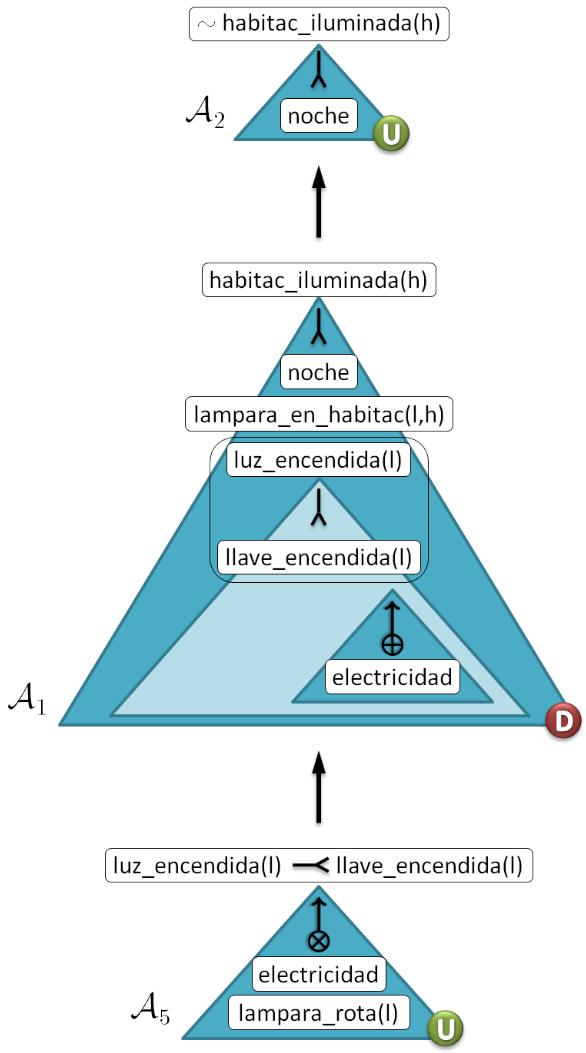


Figura 5.8: Árbol de dialéctica marcado $\mathcal{T}^*_{\langle \mathcal{A}_2, \sim habitac\_iluminada(h)\rangle}$ del Ejemplo 5.25.

Una vez identificados los argumentos aceptados entre aquellos construidos a partir de un programa E-DeLP, podremos determinar cuáles son las inferencias del programa. Como se vio en la Sección 5.3, a partir de un programa E-DeLP es posible construir argumentos de diferente tipo: argumentos conclusivos, argumentos de backing y argumentos de undercutting. En particular, las inferencias de un programa E-DeLP estarán determinadas por los argumentos conclusivos que están aceptados. Para esto, siguiendo la aproximación presentada en la Sección 2.2.2.6, definiremos la noción de *literal garantizado* a partir de un programa E-DeLP. Intuitivamente, estos serán aquellos literales correspondientes a las conclusiones de los argumentos conclusivos aceptados.

Definición 5.30 (Literal Garantizado) *Sea $\mathcal{P}$ un programa lógico rebatible extendido y h un literal. Diremos que h es un* literal garantizado a partir de $\mathcal{P}$ *si y sólo si existe un argumento conclusivo $\langle \mathcal{A}, h\rangle$ tal que es un argumento aceptado de $\mathcal{P}$.*

Para ilustrar las nociones presentadas en las definiciones 5.29 y 5.30, consideremos el siguiente ejemplo.

Ejemplo 5.26 *Sea $\mathcal{P}_{5.5}$ el programa lógico rebatible extendido del Ejemplo 5.5 y $\mathcal{T}^*_{\langle \mathcal{A}_2, \sim habitac\_iluminada(h)\rangle}$ el árbol de dialéctica marcado del Ejemplo 5.25. Como se ilustra en la Figura 5.8, la raíz de $\mathcal{T}^*_{\langle \mathcal{A}_2, \sim habitac\_iluminada(h)\rangle}$ está marcada como* U*. Por lo tanto, $\langle \mathcal{A}_2, \sim habitac\_iluminada(h)\rangle$ es un argumento aceptado de $\mathcal{P}_{5.5}$. Luego, esto implica que "$\sim habitac\_iluminada(h)$" es un literal garantizado a partir de $\mathcal{P}_{5.5}$.*

En la Sección 5.6.1 se introdujo una serie de propiedades satisfechas por E-DeLP con respecto a la aceptación de los argumentos y las inferencias del sistema. Para concluir, se mostrará que dichas propiedades se verifican para la aproximación introducida en esta sección, la cual emplea árboles de dialéctica para determinar el estado de aceptación de los argumentos en E-DeLP.

Proposición 5.7 *Sea $\mathcal{P}$ un programa lógico rebatible extendido. Si $\langle \emptyset, L\rangle$ es un argumento construido a partir de $\mathcal{P}$, entonces $\langle \emptyset, L\rangle$ es un argumento aceptado de $\mathcal{P}$.*

Demostración: ver Apéndice A.

Proposición 5.8 *Sea $\mathcal{P}$ un programa lógico rebatible extendido. Si $\langle \emptyset, L\rangle$ es un argumento construido a partir de $\mathcal{P}$, entonces L es un literal garantizado a partir de $\mathcal{P}$.*

Demostración: ver Apéndice A.

Para mostrar la propiedad que establece que no se aceptarán simultáneamente argumentos que se derrotan entre sí utilizaremos el siguiente resultado intermedio. Resumidamente, la Proposición 5.9 muestra que, dado un programa E-DeLP $\mathcal{P}$, si $\mathcal{A}$ es un argumento aceptado de $\mathcal{P}$, entonces para todo argumento $\mathcal{B}$ que es derrotado por $\mathcal{A}$ se verifica que $\mathcal{A}$ es un argumento no derrotado en el árbol de dialéctica de $\mathcal{B}$.

Proposición 5.9 *Sea $\mathcal{P}$ un programa lógico rebatible extendido. Si $\langle\mathcal{A}, h\rangle$ es un argumento aceptado de $\mathcal{P}$, entonces para todo argumento $\langle\mathcal{B}, q\rangle$ tal que $\langle\mathcal{A}, h\rangle$ derrota a $\langle\mathcal{B}, q\rangle$ en $\mathcal{P}$ se verifica que $\langle\mathcal{A}, h\rangle$ está marcado como* **U** *en el árbol de dialéctica marcado $\mathcal{T}^*_{\langle\mathcal{B}, q\rangle}$.*

Demostración: ver Apéndice A.

Utilizando el resultado anterior, la siguiente proposición formaliza la propiedad de que, dado un programa E-DeLP $\mathcal{P}$, no se aceptarán argumentos que se encuentran vinculados por una derrota en $\mathcal{P}$.

Proposición 5.10 *Sea $\mathcal{P}$ un programa lógico rebatible extendido. Si $\langle\mathcal{A}, h\rangle$ es un argumento aceptado de $\mathcal{P}$, entonces para todo argumento $\langle\mathcal{B}, q\rangle$ tal que $\langle\mathcal{A}, h\rangle$ derrota a $\langle\mathcal{B}, q\rangle$ en $\mathcal{P}$ o $\langle\mathcal{B}, q\rangle$ derrota a $\langle\mathcal{A}, h\rangle$ en $\mathcal{P}$, se verifica que $\langle\mathcal{B}, q\rangle$ es un argumento rechazado de $\mathcal{P}$.*

Demostración: ver Apéndice A.

Finalmente, de manera análoga a lo realizado en la Sección 5.6.1, el siguiente teorema muestra que a partir de un programa E-DeLP no se garantizarán literales en desacuerdo.

Teorema 5.2 *Sea $\mathcal{P}$ un programa lógico rebatible extendido. Si L es un literal garantizado a partir de $\mathcal{P}$, entonces para todo literal L' tal que L y L' están en desacuerdo se verifica que L' no es un literal garantizado a partir de $\mathcal{P}$.*

Demostración: ver Apéndice A.

5.7. Trabajo Relacionado

En este capítulo se presentó un sistema argumentativo basado en reglas que extiende al formalismo de Programación en Lógica Rebatible propuesto en [GS04]. El formalismo aquí desarrollado recibe el nombre de *Programación en Lógica Rebatible Extendida (E-DeLP)* y se halla inspirado por las ideas de Toulmin [Tou58] y Pollock [Pol87]. E-DeLP permite expresar ataque y soporte para reglas rebatibles mediante la incorporación de reglas de backing y undercutting. De esta manera, el lenguaje extendido de E-DeLP permite la representación de los backings de Toulmin y los undercutting defeaters de Pollock;

nociones que se hallan ausentes en DeLP. Asimismo, dado que E-DeLP mantiene los elementos que conforman el lenguaje de representación de DeLP, el formalismo propuesto en este capítulo constituye efectivamente una extensión de aquel propuesto en [GS04]. En este sentido, es posible especificar programas DeLP en E-DeLP, simplemente considerando un conjunto vacío de reglas de backing y undercutting.

En [Nut88, Nut94] Nute presentó d-Prolog, el primer formalismo de programación en lógica en considerar reglas estrictas y rebatibles. El lenguaje de representación de d-Prolog también brinda la posibilidad de expresar reglas derrotadoras (en inglés, *defeater rules*) las cuales no pueden ser utilizadas para obtener conclusiones, sino que representan excepciones sobre reglas rebatibles. Si bien pareciera que las reglas derrotadoras de d-Prolog pueden ser utilizadas para simular los undercutting defeaters de Pollock, este tipo de regla no constituye una razón en contra de una regla rebatible sino que provee una razón en contra de un literal. Por ejemplo, consideremos la regla derrotadora "$enfermo(X) \rightarrow \sim vuela(X)$" que expresa que "*si X está enfermo entonces no vuela*" y la regla rebatible "$ave(X) \Rightarrow vuela(X)$" que expresa que "*si X es un ave entonces existen razones para creer que vuela*". Dado el literal "$ave(pepe)$", la regla derrotadora no permite derivar el literal "$\sim vuela(pepe)$", sino que puede bloquear la inferencia del literal "$vuela(pepe)$" obtenida mediante la regla rebatible arriba mencionada. En particular, una gran diferencia entre d-Prolog y E-DeLP es que en el primero no existe la noción de argumento y, por lo tanto, no existen derrotas entre argumentos. En este sentido, para decidir entre conclusiones contradictorias, en d-Prolog simplemente se efectúa una comparación entre reglas. Finalmente, en [Gar00, GS04] los autores realizan una comparación entre d-Prolog y DeLP analizando similitudes y diferencias entre estos formalismos. Por lo tanto, dado que el sistema propuesto en este capítulo es una extensión de DeLP, los resultados de la comparación entre d-Prolog y DeLP son también aplicables al considerar el formalismo de E-DeLP.

Como se mencionó en los capítulos 2 y 4, Verheij fue uno de los primeros en retomar el estudio de la noción de soporte en argumentación en la última década. En particular, en [Ver05, Ver09] Verheij efectuó una reconstrucción de las ideas de Toulmin utilizando DefLog [Ver03a], que permite la representación de los elementos del esquema de Toulmin, así como también los vínculos de soporte entre ellos. En [Ver03a] el autor caracteriza a DefLog como una teoría de argumentación dialéctica basada en sentencias, en lugar de basada en argumentos. Esto se debe a que DefLog se enfoca en el uso suposiciones *prima* *facie*, en lugar de en los argumentos obtenidos en términos de ellas.

Resumidamente, el lenguaje lógico de DefLog posee dos conectivos: $\times$ (negación dialéctica) y $\leadsto$ (implicación primitiva). La negación dialéctica $\times S$ de una sentencia S expresa que la sentencia S está derrotada. La negación dialéctica es inherentemente "dirigida" en el siguiente sentido: si $\times S$ está justificada, entonces S está derrotada; sin embargo, no siempre será el caso que si S está justificada, entonces $\times S$ está derrotada. Por otra parte, la implicación primitiva $\leadsto$, en contraste con la implicación material de la lógica clásica, tiene por objetivo expresar relaciones condicionales elementales que existen contingentemente en el mundo. Es un conectivo binario utilizado para expresar que una sentencia soporta a otra, y permite obtener otras sentencias mediante la aplicación de *modus ponens*. Finalmente, esto posibilita la combinación y anidamiento de los conectivos $\times$ y $\leadsto$ para obtener sentencias más complejas como $A \leadsto (B \leadsto C)$ y $A \leadsto \times B$.

Otra diferencia entre DefLog y E-DeLP, es que DefLog es un formalismo basado en sentencias, mientras que E-DeLP es un formalismo argumentativo basado en programación en lógica. Adicionalmente, los argumentos en DefLog son conjuntos de sentencias, mientras que en E-DeLP los argumentos constituyen un conjunto específico de reglas (rebatibles y/o de backing). Como se mencionó anteriormente, en DefLog es posible combinar y anidar los conectivos $\times$ y $\leadsto$, permitiendo así representar tanto los backings de Toulmin [Tou58] como los undercutting defeaters de Pollock [Pol87]. No obstante esto, dado que la negación dialéctica implica derrota, un argumento para una sentencia $\times S$ siempre será preferido ante un argumento para una sentencia S. Por lo tanto, en DefLog no es posible expresar ataque sin derrota; es decir, las nociones de ataque y derrota son equivalentes en la aproximación de Verheij. En contraste, en E-DeLP es posible representar tanto la noción de ataque como la de derrota, dado que la existencia de un ataque sobre un argumento no necesariamente conduce a su derrota.

En [Pra09] se introdujo ASPIC+, un formalismo argumentativo que instancia la aproximación de [Dun95] especificando parcialmente la estructura interna de los argumentos, e identificando diferentes tipos de ataque entre argumentos. Una diferencia central entre ASPIC+ y E-DeLP es que el primero propone una noción general de argumento, mientras que el segundo realiza una distinción entre diferentes tipos de argumento, incluyendo argumentos de backing y undercutting. Por otra parte, al igual que en ASPIC+, en E-DeLP se distinguen tres tipos de ataque entre argumentos: rebutting, undermining y undercutting, cada uno de los cuales conduce a su correspondiente tipo de derrota. Para la determinación del éxito de los ataques, en la mayoría de los casos, ASPIC+ utiliza un orden parcial correspondiente a una relación de preferencia entre argumentos. Sin embargo, para los ataques de tipo undercutting, en ASPIC+ no se efectúa comparación alguna entre argumentos, sino que dichos ataques siempre resultan exitosos. En contraste, en E-DeLP, la determinación del éxito de los undercuts involucra una comparación entre el argumento atacante y los argumentos de backing correspondientes. En este sentido, de manera similar a lo ocurrido con otros tipos de ataque, los undercuts no siempre darán lugar a derrotas entre argumentos de E-DeLP.

5.8. Conclusiones

En este capítulo se presentó un Sistema Argumentativo Basado en Reglas (SABR) llamado *Programación en Lógica Rebatible Extendida (E-DeLP)*. De manera similar a la propuesta desarrollada en el Capítulo 4, este formalismo aborda las ideas de Toulmin [Tou58] y Pollock [Pol87], pero en un entorno más concreto. En particular, E-DeLP constituye una extensión a la Programación en Lógica Rebatible (DeLP) [GS04] que incorpora reglas de backing y undercutting para modelar ataque y soporte para inferencias. A partir de esto, E-DeLP brinda la posibilidad de construir argumentos conclusivos, argumentos de backing y argumentos de undercutting; en particular, los dos últimos se hallan estrechamente relacionados con las ideas de Toulmin y Pollock.

Dada la caracterización de los argumentos en E-DeLP, es posible identificar tres tipos de ataque entre los mismos: rebutting, undercutting y undermining. Mientras que el primer y tercer tipo de ataque tienen por objeto disputar la conclusión de un argumento, el segundo constituye una razón en contra sobre una regla rebatible utilizada por otro argumento. En particular, los ataques de tipo undercutting son originados por los argumentos homónimos. Luego, los argumentos de backing tienen por objetivo impedir el éxito de los undercuts, proveyendo una defensa para la regla rebatible que se halla en disputa.

Como se mencionó anteriormente, el objetivo final de E-DeLP radica en la obtención de las inferencias del sistema, correspondientes a los literales garantizados a partir de los programas lógicos rebatibles extendidos. Para esto, en primera medida, es necesario determinar el éxito de los ataques, dando lugar así a las derrotas entre los argumentos de E-DeLP. Como se vio en la Sección 5.5, dichas derrotas son obtenidas a partir de la caracterización del BUAF asociado a un programa E-DeLP. Luego, a partir de estas derrotas, es posible efectuar el cálculo de aceptabilidad sobre los argumentos de E-DeLP.

En la Sección 5.6 se presentaron dos alternativas para determinar la aceptación o rechazo de los argumentos construidos a partir de un programa E-DeLP. La primera de ellas, introducida en la Sección 5.6.1, comprende el uso de semánticas de aceptabilidad sobre el BUAF asociado al programa E-DeLP en cuestión. Por otra parte, en la Sección 5.6.2 se propuso una aproximación procedural para el cálculo de aceptabilidad, siguiendo el modelo de prueba dialéctico de DeLP presentado en la Sección 2.2.2. Para ambas aproximaciones, se mostró una serie de propiedades satisfechas por E-DeLP con respecto a la aceptación de los argumentos y a las inferencias del sistema, las cuales corresponden a características deseables para todo sistema argumentativo.

Al igual que E-DeLP, existen otras aproximaciones en la literatura que abordan el uso de soporte. No obstante esto, como se vio en el Capítulo 3, gran parte de los desarrollos existentes corresponden al área de argumentación abstracta. En particular, en [Ver05, Ver09] Verheij efectuó una reconstrucción de las ideas de Toulmin, admitiendo también el modelado de los undercuts de Pollock. Sin embargo, dicha reconstrucción fue llevada a cabo mediante el uso de DefLog [Ver03a], una teoría de argumentación dialéctica basada en sentencias. Por lo tanto, E-DeLP constituye el primer SABR en formalizar las ideas de Toulmin y Pollock en forma conjunta.

Finalmente, dada su formalización como un SABR, E-DeLP posee características que favorecen su implementación computacional. Asimismo, E-DeLP permite la instanciación del marco argumentativo abstracto propuesto en el Capítulo 4. En consecuencia, los desarrollos correspondientes a los capítulos 4 y 5 de esta tesis proveen un entorno combinado para el estudio de las ideas de Toulmin y Pollock. Por una parte, dado su nivel de abstracción, el BUAF permite el estudio de propiedades de alto nivel. Por otra parte, los programas E-DeLP involucran especificaciones concretas, proveyendo entonces una herramienta computacional para asistir a la toma de decisiones utilizando argumentación.

Capítulo 6

Conclusiones y Trabajo a Futuro

En esta tesis se desarrollaron dos formalismos argumentativos que incorporan la noción de *soporte*. Uno de ellos corresponde a un formalismo de argumentación abstracta, mientras que el otro constituye un sistema argumentativo basado en reglas. Estos formalismos extienden a otras aproximaciones existentes en la literatura, de manera tal que permiten expresar una relación de soporte entre los elementos del sistema. En particular, los desarrollos realizados en esta tesis se hallan motivados por una crítica usualmente realizada sobre los formalismos argumentativos ya existentes. Esta crítica radica en que determinados patrones de razonamiento argumentativo estudiados en áreas como el razonamiento legal, la teoría de la comunicación y la filosofía [Tou58, Pol87], y que constituyen aportes a la argumentación, no han sido considerados por los formalismos desarrollados hasta el momento. Entre estos patrones de razonamiento, podemos identificar la presencia de una interacción positiva entre argumentos, la cual se halla capturada por la noción de soporte.

La noción de soporte ha estado presente en la literatura de argumentación desde sus trabajos fundacionales en la filosofía. En particular, en su libro "*The Uses of Argument*" [Tou58], Toulmin promulgó la idea de que los argumentos deberían ser analizados utilizando un formato más rico que la dicotomía de premisas y conclusiones empleada en la lógica formal. Por lo tanto, propuso un modelo en el que una sentencia en forma de regla general que provee la conexión entre premisas y conclusión de un argumento (*warrant*) puede estar soportada por un *backing*, donde este último provee la justificación por la cual el warrant correspondiente es válido.

Estudios subsiguientes en el área de argumentación dejaron de lado la noción de soporte para concentrarse en la noción de *derrota*, la cual constituye un elemento clave en el cálculo de aceptabilidad de los argumentos de un sistema. Entre los trabajos con estas características podemos destacar el realizado por Dung en [Dun95] (ver Sección 2.2.1), muy influyente en los desarrollos de argumentación abstracta. Dung propone un marco argumentativo abstracto (AF) considerando un conjunto de argumentos y una relación de derrota entre ellos, la cual es utilizada para identificar los argumentos aceptados del sistema, sin contemplar la existencia de soporte. Siguiendo la línea de trabajo de Dung, muchos trabajos en la literatura se han dedicado al desarrollo de formalismos argumentativos cuyo foco se encuentra en la relación de derrota (ver *e. g.*, [Mod09, BCGG11, BCG11]).

En la última década, el estudio de la noción de soporte recobró atención entre los investigadores del área de argumentación. Como resultado de estas investigaciones, se han desarrollado diversas propuestas que adoptan diferentes interpretaciones para la noción de soporte [CGGS13b, CLS13]. En el Capítulo 3 se analizaron diferentes aproximaciones existentes en la literatura que contemplan el uso de soporte en sistemas argumentativos abstractos. En particular, se analizaron diversas *interpretaciones* para la noción de soporte, efectuando un estudio comparativo entre ellas. En primer lugar, se analizó la propuesta de [CLS05] (ver Sección 3.1), donde se utiliza una relación de *soporte general* que simplemente representa una interacción positiva entre argumentos, es decir, sin imponer restricciones sobre los argumentos que relaciona. Seguidamente, se analizó la interpretación de *soporte deductivo* [BGvdTV10] (ver Sección 3.2), la cual intenta capturar la siguiente intuición: si un argumento $\mathcal{A}$ soporta a un argumento $\mathcal{B}$, entonces la aceptación de $\mathcal{A}$ implica la aceptación de $\mathcal{B}$ y, análogamente, la no aceptación de $\mathcal{B}$ implica la no aceptación de $\mathcal{A}$. En la Sección 3.3 se estudió la interpretación de *soporte por necesidad* introducida en [NR10]. El soporte por necesidad refuerza la siguiente restricción: si un argumento $\mathcal{A}$ soporta a un argumento $\mathcal{B}$, esto significa que $\mathcal{A}$ es *necesario* para $\mathcal{B}$. Por lo tanto, la aceptación de $\mathcal{B}$ implica la aceptación de $\mathcal{A}$ y, de manera similar, la no aceptación de $\mathcal{A}$ implica la no aceptación de $\mathcal{B}$. Luego, se analizó la interpretación de *soporte evidencial* [ON08] (ver Sección 3.4), la cual permite distinguir entre argumentos *prima facie* y argumentos estándar. Los argumentos *prima facie* representan la noción de evidencia y no requieren del soporte de otros argumentos para subsistir, mientras que los argumentos estándar no pueden ser aceptados a menos que estén soportados por evidencia.

En el Capítulo 3, se mostró que cada interpretación de soporte de las allí analizadas conduce a la consideración de una serie de restricciones sobre la aceptabilidad de los argumentos que relaciona. En la mayoría de los casos, estas restricciones resultan en la obtención de nuevas derrotas entre argumentos, identificadas como *derrotas complejas*, las cuales se infieren a partir de la combinación de derrotas ya existentes y de la relación de soporte entre argumentos. En particular, se mostró que estas derrotas complejas cobran sentido dependiendo de la interpretación adoptada por la relación de soporte, dado que refuerzan las restricciones de aceptabilidad impuestas por cada una de estas interpretaciones. De esta manera, para una interpretación particular de soporte será necesario considerar el conjunto de derrotas complejas correspondiente a la misma. Como conclusión, no sería posible proveer un mecanismo formal que compute estas derrotas abstrayéndose de la interpretación adoptada para la relación de soporte de un sistema argumentativo ya que, si estas derrotas fueran consideradas de manera arbitraria, las inferencias del sistema podrían conducir a resultados incorrectos.

A partir del análisis de las similitudes y diferencias entre las interpretaciones de soporte estudiadas, podemos concluir que algunas de ellas se encuentran estrechamente relacionadas. En particular, se mostró que las interpretaciones de soporte deductivo y soporte por necesidad (presentadas en las secciones 3.2 y 3.3) son duales en el siguiente sentido: un argumento $\mathcal{A}$ soporta deductivamente a un argumento $\mathcal{B}$ si y sólo si $\mathcal{B}$ es necesario para el argumento $\mathcal{A}$. De manera similar, se mostró que existe una correspondencia directa entre la relación de sub-argumento presentada en la Sección 3.5.1 y la relación de soporte por necesidad. Por otra parte, se mostró que existe una característica esencial que diferencia al soporte evidencial (presentado en la Sección 3.4) de las otras aproximaciones estudiadas: un argumento aceptado a partir de un sistema argumentativo evidencial debe estar soportado por evidencia. Por lo tanto, dado un sistema argumentativo evidencial, podría ocurrir que un argumento del sistema que no es objeto de derrotas no pertenezca al conjunto de argumentos aceptados dado que no se encuentra soportado por evidencia.

En el Capítulo 4 se propuso un nuevo formalismo de argumentación abstracta llamado *Marco Argumentativo con Backing y Undercutting (BUAF)*. El BUAF extiende a los marcos argumentativos abstractos de [Dun95] permitiendo modelar tanto ataque como soporte a inferencias en el contexto de argumentación abstracta. Por una parte, el BUAF considera diferentes tipos de ataque entre argumentos: *rebutting*, *undercutting* y *undermining*. El primero de estos constituye un ataque sobre la conclusión de un argumento; el segundo representa un ataque sobre una premisa de un argumento; y el último constituye un ataque sobre una regla de inferencia de un argumento. Por otra parte, el BUAF incorpora una relación de soporte especializada que combina las nociones de *backing* [Tou58] y *sub-argumento*. De esta manera, los BUAF permiten la representación de los *backings* de Toulmin [Tou58] y los *undercuts* de Pollock [Pol87], dos importantes nociones en el área de argumentación.

Si bien las nociones de backing y undercutting han sido abordadas con anterioridad por diferentes aproximaciones existentes en la literatura, ninguna de las propuestas de argumentación abstracta desarrolladas previamente las había considerado en forma conjunta. Por ejemplo, en [Pra09] y [BGvdTV10] se proveen formalizaciones de argumentación abstracta que permiten modelar la noción de undercut propuesta por Pollock. En particular, la propuesta de [BGvdTV10] también considera una relación de soporte entre argumentos (ver Sección 3.2). No obstante esto, como se mencionó anteriormente, en [BGvdTV10] se adopta una interpretación deductiva para la noción de soporte. De esta manera, el BUAF es el primer formalismo en proveer un marco unificado para el modelado de ataque y soporte a inferencias en un contexto de argumentación abstracta.

Al igual que los BUAF, muchas otras aproximaciones que proponen una relación de soporte entre argumentos adoptan una interpretación de soporte particular. Tal es el caso de las interpretaciones de soporte deductivo, soporte por necesidad, o soporte evidencial estudiadas en el Capítulo 3. Asimismo, todas estas aproximaciones definen un conjunto de derrotas complejas, las cuales refuerzan las restricciones de aceptabilidad impuestas por la correspondiente interpretación de soporte. Para el caso de los BUAF, en el Capítulo 4 se identificaron dos tipos de derrota compleja: derrota implícita y derrota indirecta. En particular, se mostró que existe una relación entre las derrotas implícitas y las derrotas obtenidas a partir de un ataque de tipo undercutting (Proposición 4.1).

Se observó que las derrotas complejas definidas por el BUAF coinciden con algunas de las derrotas complejas de los formalismos estudiados en el Capítulo 3. Por una parte, se mostró que la derrota indirecta definida para los BUAF se corresponde con la derrota secundaria de los BAF presentados en la Sección 3.1, así como también con el primer caso de derrota extendida de los AFN introducidos en la Sección 3.3. Esta correspondencia es de esperarse ya que, como se vio7 en el Capítulo 4, la relación de soporte de los BUAF combina las nociones de sub-argumento y backing. Por una parte, tomando en cuenta el principio de composicionalidad propuesto en [Vre97] (y como fue mencionado en la Sección 3.5.1), los sub-argumentos son considerados como *necesarios* para sus super-argumentos. Por otra parte, dado que una de las responsabilidades que un backing posee con respecto al argumento que soporta es establecer las circunstancias bajo las cuales la conexión entre sus premisas y conclusión (*i. e.*, su warrant asociado) es válida, podemos considerar que los argumentos de backing son *necesarios* para los argumentos que soportan.

Con respecto a la derrota implícita de los BUAF, esta fue definida de manera tal que considera los conflictos existentes entre un argumento de backing $\mathcal{B}$ y un argumento de undercutting $\mathcal{A}$ para un argumento $\mathcal{C}$. Dada esta situación, es posible que $\mathcal{A}$ derrote implícitamente a $\mathcal{B}$ *y/o* $\mathcal{B}$ derrote implícitamente a $\mathcal{A}$. Para el primer caso, se mostró que la derrota implícita se corresponde con la derrota mediada de los d-BAF presentados en la Sección 3.2. Por otra parte, se observó que el segundo caso de derrota implícita no se corresponde con ninguna de las derrotas complejas estudiadas en el Capítulo 3. En particular, esto se debe a que las derrotas implícitas se hallan estrechamente relacionadas con la interpretación de soporte adoptada por los BUAF. Es decir, en un contexto en el que pueden existir argumentos de backing y undercutting para un mismo argumento, las derrotas implícitas tienen por objetivo prevenir la aceptación conjunta de los argumentos de backing y undercutting.

Dada la coexistencia de argumentos de ataque y soporte para inferencias en el BUAF, se caracterizó la noción de *coherencia externa* para conjuntos que no atacan y soportan simultáneamente a un mismo argumento, correspondiente a una propiedad deseable para todo conjunto de argumentos aceptados a partir del BUAF. En este sentido, se mostró que la noción de *libertad de conflictos* es lo suficientemente fuerte como para asegurar las restricciones impuestas por la coherencia externa (Proposición 4.2). Por otra parte, con respecto a la aceptabilidad de los argumentos, se mostró que si un argumento $\mathcal{A}$ es aceptable con respecto a un conjunto C, entonces todo argumento que soporta a $\mathcal{A}$ también será aceptable con respecto a dicho conjunto (Proposición 4.3).

Como resultado del análisis efectuado en el Capítulo 4, podemos concluir que algunos aspectos de la relación de soporte de los BUAF se corresponden con aspectos de otras interpretaciones de soporte propuestas en al literatura. Sin embargo, se mostró que la relación de backing de los BUAF impone restricciones adicionales sobre la aceptabilidad de los argumentos, las cuales se ven reflejadas en la definición de las derrotas implícitas. Por lo tanto, dada la alta dependencia de las restricciones con respecto a la interpretación de soporte adoptada, estas no pueden ser capturadas por otros formalismos como los estudiados en el Capítulo 3.

Adicionalmente, dado que el BUAF se encuentra formalizado como un marco argumentativo abstracto, esto le posibilita la utilización de la gran variedad de semánticas de aceptabilidad propuestas en la literatura (ver *e. g.*, [Dun95, BCG11]). Más aún, se mostró que, a partir de los argumentos y las derrotas entre argumentos de un BUAF, es posible caracterizar un AF que acepta exactamente los mismos argumentos que el BUAF (Proposición 4.4). Asimismo, la incorporación de relaciones de soporte y preferencia, como también la distinción entre diferentes tipos de ataque, sitúa a los BUAF en un nivel intermedio entre los AF y sistemas argumentativos más concretos, tales como el sistema argumentativo basado en reglas propuesto en el Capítulo 5.

En el Capítulo 5 se desarrolló un Sistema Argumenativo Basado en Reglas (SABR) llamado *Programación en Lógica Rebatible Extendida (E-DeLP)*. En particular, E-DeLP constituye una extensión de la Programación en Lógica Rebatible (DeLP) (presentada en la Sección 2.2.2), incorporando reglas de backing y undercutting que permiten modelar ataque y soporte para inferencias. En consecuencia, de manera similar a los BUAF presentados en el Capítulo 4, E-DeLP aborda las ideas de Toulmin [Tou58] y Pollock [Pol87] en forma conjunta. No obstante esto, dada su caracterización como un SABR, el tratamiento de E-DeLP sobre las nociones de backing y undercutting es efectuado sobre un entorno más concreto que el de los BUAF.

En E-DeLP se definieron tres tipos de argumento: argumentos conclusivos, argumentos de backing y argumentos de undercutting; estando los dos últimos estrechamente relacionados con las ideas de Toulmin y Pollock. A partir de esta clasificación, y dada la existencia de sub-argumentos, se mostró que todo sub-argumento propio de un argumento será, en particular, un argumento conclusivo o un argumento de backing (Proposición 5.1). Luego, dados los argumentos construidos a partir de un programa E-DeLP, se identificaron tres tipos de ataque entre ellos: ataque rebutting, ataque undermining y ataque undercutting. Por una parte, los primeros dos tipos de ataque tienen por objeto disputar la conclusión de un argumento (para el caso del ataque undermining siendo la conclusión, en particular, una premisa). Por otra parte, el último tipo de ataque representa la existencia de una razón en contra para una regla rebatible utilizada por otro argumento y, en consecuencia, es originado por un argumento de undercutting. Teniendo en cuenta esto, el rol de los argumentos de backing resultó fundamental ante la existencia de ataques de este tipo, dado que proveen una defensa para la regla rebatible atacada con el objetivo de impedir el éxito de los undercuts.

A partir de la identificación de los argumentos de backing y los diferentes tipos de ataque entre los argumentos de E-DeLP, resultó necesario proveer un mecanismo para obtener las derrotas entre los argumentos. En particular, como se vio en la Sección 5.5, para esto se utilizaron los desarrollos del Capítulo 4, brindando así una especificación del BUAF asociado a un programa E-DeLP. Se mostró que dicho BUAF posee características especiales, entre las cuales se destaca el hecho de que todo argumento del BUAF posee a lo sumo un argumento de backing (Proposición 5.3). Luego, a partir de la especificación del BUAF asociado al programa E-DeLP, se determinaron las derrotas entre los argumentos del sistema, las cuales constituyen un elemento clave en el cálculo de aceptabilidad para los argumentos de E-DeLP.

Cabe destacar que, como se mostró en la Sección 5.6, la utilización del BUAF asociado a un programa E-DeLP posibilitó la exploración de diferentes alternativas para el cálculo de aceptabilidad de los argumentos del sistema. Por una parte, aprovechando los desarrollos del Capítulo 4, se propuso una aproximación extensional para la determinación de los argumentos aceptados de E-DeLP mediante la aplicación de semánticas de aceptabilidad sobre el BUAF asociado a un programa lógico rebatible extendido. En particular, un aspecto característico de esta aproximación es que el cálculo de aceptabilidad concierne a todos los argumentos construidos a partir de un programa E-DeLP. Es decir, dado un programa E-DeLP $\mathcal{P}$ y una semántica de aceptabilidad s, la aproximación extensional identifica todos los argumentos de $\mathcal{P}$ que se encuentran aceptados bajo la semántica s.

Por otra parte, dado que E-DeLP extiende el formalismo de DeLP, se presentó una aproximación alternativa para el cálculo de aceptabilidad, de manera similar a lo realizado por DeLP (ver Sección 2.2.2), proveyendo de esta manera un modelo de prueba dialéctico para E-DeLP. En particular, esta alternativa adopta el enfoque orientado a consultas de DeLP, donde el estado de aceptabilidad de los argumentos es determinado mediante la construcción de árboles de dialéctica, en función de las consultas realizadas por los usuarios del sistema. En este sentido, el interés radica en la determinación de los literales garantizados a partir de un programa E-DeLP, los cuales corresponden a las conclusiones de los argumentos aceptados. Por lo tanto, en contraste con la aproximación extensional, este proceso no involucra el cálculo de aceptabilidad para todos los argumentos del sistema, sino únicamente para aquellos argumentos vinculados a las consultas realizadas.

Para ambas aproximaciones propuestas en la Sección 5.6, se mostró una serie de propiedades satisfechas por E-DeLP con respecto a la aceptación de los argumentos y a las inferencias del sistema. En primera medida, se mostró que los argumentos construidos a partir del conocimiento estricto de un programa E-DeLP estarán aceptados (Proposiciones 5.4 y 5.7) y, en consecuencia, sus conclusiones estarán garantizadas (Proposiciones 5.5 y 5.8). Luego, se mostró que a partir de un programa E-DeLP no se aceptan simultáneamente argumentos que se derroten entre sí (Proposiciones 5.6 y 5.10). Finalmente, se mostró una propiedad fundamental acerca de las inferencias de E-DeLP, la cual establece que las mismas son consistentes ya que no se garantizan literales en desacuerdo (Teoremas 5.1 y 5.2).

Al igual que E-DeLP, existen otras aproximaciones en la literatura que abordan el uso de soporte. No obstante esto, como se vio en el Capítulo 3, gran parte de los desarrollos existentes corresponden al área de argumentación abstracta. En particular, en [Ver05, Ver09] Verheij efectuó una reconstrucción de las ideas de Toulmin, admitiendo también el modelado de los undercuts de Pollock. Sin embargo, dicha reconstrucción fue llevada a cabo mediante el uso de DefLog [Ver03a], una teoría de argumentación dialéctica basada en sentencias. Por lo tanto, E-DeLP constituye el primer SABR en formalizar las ideas de Toulmin y Pollock en forma conjunta.

Es importante notar que los desarrollos de esta tesis proveen un marco combinado para el estudio de las ideas de Toulmin y Pollock en forma conjunta. Por una parte, el BUAF desarrollado en el Capítulo 4 permite la representación de ataque y soporte a inferencias en entornos abstractos, posibilitando el estudio de propiedades de alto nivel. Por otra parte, cabe destacar que el formalismo de E-DeLP permite la instanciación del BUAF desarrollado en el Capítulo 4, favoreciendo su implementación computacional. En este sentido, dado que los programas E-DeLP involucran especificaciones concretas, el formalismo desarrollado en el Capítulo 5 brinda una herramienta computacional para asistir a la toma de decisiones utilizando argumentación.

Finalmente, podemos concluir que el estudio de la relación de soporte entre argumentos constituye un prometedor tópico de investigación para la comunidad de argumentación dentro del área de Inteligencia Artificial. Gran parte de los trabajos existentes en la literatura se concentran en el estudio de marcos argumentativos abstractos como el propuesto por Dung en [Dun95], donde únicamente se consideran derrotas entre argumentos. Sin embargo, la incorporación de un relación de soporte en este tipo de sistemas ha demostrado ser de gran utilidad. En particular, la incorporación de la noción de soporte en los formalismos argumentativos incrementa las capacidades de representación en este tipo de sistemas. En consecuencia, para aquellos dominios de aplicación que actualmente utilizan argumentación como mecanismo de razonamiento, será posible brindar especificaciones más granulares, resultando en un modelado más preciso de los problemas abordados. Asimismo, esto contribuirá al uso de argumentación en otros dominios de aplicación en los que resulte imprescindible la consideración de una noción de soporte.

Trabajo a Futuro

A continuación se incluyen algunas líneas de investigación sobre las cuales se planea seguir trabajando, y que se encuentran relacionadas con los desarrollos de esta tesis.

A partir del análisis realizado en el Capítulo 3, se observó que las interpretaciones de soporte consideradas por otros formalismos existentes sólo han sido estudiadas a nivel de argumentación abstracta. Por lo tanto, se buscará estudiar dichas interpretaciones de soporte en un nivel más concreto (considerando sistemas argumentativos basados en reglas) y se buscará establecer una relación con la formalización de soporte provista por los marcos argumentativos correspondientes.

Como otra línea de trabajo a futuro, se abordará el estudio de interpretaciones de soporte alternativas a aquellas ya consideradas en la literatura. Asimismo, el estudio de interpretaciones alternativas para la relación de soporte puede conducir a la exploración de otras relaciones entre argumentos que aún no hayan sido consideradas por los formalismos argumentativos desarrollados hasta el momento. En particular, tomando como motivación los trabajos de [Mod09] y [BCGG09, BCGG11], donde se contempla la existencia de derrotas recursivas en un contexto de argumentación abstracta, como línea de investigación futura se buscará expandir el estudio acerca de la noción de soporte para considerar, además del soporte entre argumentos, una relación de soporte capaz de proveer soporte sobre sí misma o incluso sobre la relación de derrota. Análogamente, se analizará la posibilidad de permitir también derrotas sobre la relación de soporte. Con lo cual, como parte de esta línea de investigación, se buscará definir sistemas argumentativos que permitan representar, de la forma más general posible, derrotas sobre otras derrotas, derrotas sobre soportes, soportes sobre derrotas, y soportes sobre soportes.

Apéndice A

Demostraciones

Este apéndice contiene todas las propiedades (proposiciones y teoremas) introducidas en esta tesis, incluyendo además sus correspondientes demostraciones.

Proposición 4.1 [pág. 125] *Sea* $\langle \mathbb{A}, \mathbb{R}, \mathbb{S}, \preceq \rangle$ *un* BUAF *y* $\mathcal{A}, \mathcal{B}, \mathcal{C} \in \mathbb{A}$ *tal que* $(\mathcal{A}, \mathcal{C}) \in \mathbb{U}_c$ *y* $(\mathcal{B}, \mathcal{C}) \in \mathbb{B}_k$*. Si* $\mathcal{A}$ *derrota implícitamente a* $\mathcal{B}$*, entonces* $\mathcal{A}$ *derrota primariamente a* $\mathcal{C}$*.*

Demostración: Si $\mathcal{A}$ *derrota implícitamente a* $\mathcal{B}$*, entonces, por Definición 4.3, la relación de preferencia es tal que* $\mathcal{A} \not\prec \mathcal{B}$*. Luego, por Definición 4.2 se tiene que* $\mathcal{A}$ *derrota primariamente a* $\mathcal{C}$ *ya que existe un backing* $\mathcal{B}$ *para* $\mathcal{C}$ *tal que* $\mathcal{A} \not\prec \mathcal{B}$*.* □

Proposición 4.2 [pág. 135] *Sea* $\langle \mathbb{A}, \mathbb{R}, \mathbb{S}, \preceq \rangle$ *un* BUAF *y* $S \subseteq \mathbb{A}$*. Si* S *es un conjunto libre de conflictos, entonces* S *es externamente coherente.*

Demostración: Supongamos por el absurdo que el conjunto S *es libre de conflictos y no es externamente coherente. Entonces, por Definición 4.6,* $\exists \mathcal{A}, \mathcal{B} \in S$*,* $\exists \mathcal{C} \in \mathbb{A}$ *tal que* $(\mathcal{A}, \mathcal{C}) \in \mathbb{U}_c$ *y* $(\mathcal{B}, \mathcal{C}) \in \mathbb{B}_k$*. Luego, por Definición 4.3 esto implica que* $\mathcal{A} \longrightarrow \mathcal{B}$*,* $\mathcal{B} \longrightarrow \mathcal{A}$*, o bien* $\mathcal{A} \longrightarrow \mathcal{B}$ *y* $\mathcal{B} \longrightarrow \mathcal{A}$*. En cualquiera de estos casos, existe al menos una derrota que involucra a los argumentos* $\mathcal{A}$ *y* $\mathcal{B}$*, contradiciendo la hipótesis de que* S *es un conjunto libre de conflictos.* □

Proposición 4.3 [pág. 137] *Sea* $\langle \mathbb{A}, \mathbb{R}, \mathbb{S}, \preceq \rangle$ *un* BUAF*,* $\mathcal{A} \in \mathbb{A}$ *y* $S \subseteq \mathbb{A}$*. Si* $\mathcal{A}$ *es aceptable con respecto a* S*, entonces* $\forall \mathcal{B} \in \mathbb{A}$ *tal que* $(\mathcal{B}, \mathcal{A}) \in \mathbb{S}$ *se verifica que* $\mathcal{B}$ *es aceptable con respecto a* S*.*

Demostración: Supongamos por el absurdo que $\mathcal{A}$ es aceptable con respecto a S y que existe un argumento $\mathcal{B} \in \mathbb{A}$ tal que $(\mathcal{B}, \mathcal{A}) \in \mathbb{S}$ y $\mathcal{B}$ no es aceptable con respecto a S. Entonces, por Definición 4.8, $\exists \mathcal{C} \in \mathbb{A}$ tal que $\mathcal{C} \longrightarrow \mathcal{B}$ y S no defiende a $\mathcal{B}$ ante la derrota de $\mathcal{C}$. Luego, por Definición 4.4, tenemos que $\mathcal{C} \longrightarrow \mathcal{A}$ dado que $\mathcal{C} \longrightarrow \mathcal{B}$ y $(\mathcal{B}, \mathcal{A}) \in \mathbb{S}$. En particular, como S no defiende a $\mathcal{B}$ ante la derrota de $\mathcal{C}$, tampoco puede defender a $\mathcal{A}$ de esta derrota. Por lo tanto, el argumento $\mathcal{A}$ no es aceptable con respecto a S, lo que contradice la hipótesis. □

Proposición 4.4 [pág. 139] *Sea $\mathsf{BUAF} = \langle \mathbb{A}, \mathbb{R}, \mathbb{S}, \preceq \rangle$ un marco argumentativo con backing y undercutting. Existe un marco argumentativo abstracto $\mathsf{AF} = \langle \mathbb{A}, \longrightarrow \rangle$ tal que para toda semántica $\mathsf{s} \in \{completa,\ preferida,\ estable,\ grounded\}$ las extensiones de BUAF y AF bajo s coinciden.*

Demostración: Directa a partir de las definiciones 2.2, 2.3, 4.8, 4.9 y 4.10. □

Proposición 5.1 [pág. 171] *Sea $\mathcal{P}$ un programa lógico rebatible extendido y $\langle \mathcal{A}, h \rangle$ un argumento construido a partir de $\mathcal{P}$. Todo sub-argumento propio de $\langle \mathcal{A}, h \rangle$ es un argumento conclusivo o un argumento de backing.*

Demostración: Sea $\mathcal{P} = (\Pi, \Delta, \Sigma)$. Consideremos los siguientes casos: (a) $\langle \mathcal{A}, h \rangle$ es un argumento conclusivo, (b) $\langle \mathcal{A}, h \rangle$ es un argumento de backing, o (c) $\langle \mathcal{A}, h \rangle$ es un argumento de undercutting.

- (a) *Si $\langle \mathcal{A}, h \rangle$ es un argumento conclusivo para h, por Definición 5.10, existe una derivación rebatible para h a partir de $\Pi \cup \mathcal{A}$. Además, por Definición 5.7, la derivación de h no emplea reglas de undercutting. Luego, por la condición de minimalidad de la Definición 5.10, el conjunto $\mathcal{A}$ no contiene reglas de undercutting. Por lo tanto, ningún subconjunto propio $\mathcal{B}$ de $\mathcal{A}$ contendrá reglas de undercutting por lo que, en consecuencia, todo sub-argumento propio $\langle \mathcal{B}, q \rangle$ de $\langle \mathcal{A}, h \rangle$ será un argumento conclusivo o un argumento de backing.*

- (b) *Si $\langle \mathcal{A}, h \rangle$ es un argumento de backing para h, por Definición 5.11, $\mathcal{A} = \{[h] \leftarrow\!\oplus [L_1, \ldots, L_n]\} \cup \mathcal{A}'$ y existe una derivación rebatible para todo literal L_i $(1 \leq i \leq n)$ a partir de $\Pi \cup \mathcal{A}'$. Nuevamente, por Definición 5.7, la derivación de los literales L_i no emplea reglas de undercutting. Luego, por la condición de minimalidad de la Definición 5.11, el conjunto $\mathcal{A}$ no contiene reglas* *de undercutting. Luego, ningún subconjunto propio $\mathcal{B}$ de $\mathcal{A}$ contendrá reglas de undercutting por lo que, en consecuencia, todo sub-argumento propio $\langle \mathcal{B}, q\rangle$ de $\langle \mathcal{A}, h\rangle$ será un argumento conclusivo o un argumento de backing.*

(c) *Si $\langle \mathcal{A}, h\rangle$ es un argumento de undercutting, por Definición 5.12, $\mathcal{A} = \{[h] \leftarrow\!\otimes [L_1, \ldots, L_n]\} \cup \mathcal{A}'$ y existe una derivación rebatible para todo literal L_i $(1 \leq i \leq n)$ a partir de $\Pi \cup \mathcal{A}'$. Nuevamente, por Definición 5.7, la derivación de los literales L_i no emplea reglas de undercutting. Luego, por la condición de minimalidad de la Definición 5.12, el conjunto $\mathcal{A}$ contendrá exactamente una regla de undercutting: la regla $[h] \leftarrow\!\otimes [L_1, \ldots, L_n]$. Supongamos por el absurdo que existe un subconjunto propio $\mathcal{B}$ de $\mathcal{A}$ tal que $\langle \mathcal{B}, q\rangle$ es un argumento de undercutting. En particular, para ser un argumento de undercutting, el conjunto $\mathcal{B}$ debe contener una regla de undercutting. Luego, como la única regla de undercutting perteneciente al conjunto $\mathcal{A}$ es $[h] \leftarrow\!\otimes [L_1, \ldots, L_n]$, $\langle \mathcal{B}, q\rangle$ debe ser un argumento de undercutting para la regla rebatible h (y, por lo tanto, $q = h$). En consecuencia, existiría un subconjunto propio $\mathcal{B}$ de $\mathcal{A}$ tal que verifica las condiciones impuestas por la Definición 5.12, contradiciendo el hecho de que $\langle \mathcal{A}, h\rangle$ es un argumento de undercutting para h (porque $\mathcal{A}$ no sería minimal). Finalmente, para todo sub-argumento propio $\langle \mathcal{B}, q\rangle$ de $\langle \mathcal{A}, h\rangle$, debe ser el caso que $\langle \mathcal{B}, q\rangle$ es un argumento conclusivo o un argumento de backing.*

$\square$

Proposición 5.2 [pág. 189] *Sea $\mathcal{P} = (\Pi, \Delta, \Sigma)$ un programa lógico rebatible extendido y $\langle \mathcal{A}, h\rangle$ un argumento conclusivo construido a partir de $\mathcal{P}$. Si r es una regla rebatible conclusiva de $\langle \mathcal{A}, h\rangle$, entonces $\nexists r' \in \Delta$ tal que $r' \neq r$ y r' es una regla rebatible conclusiva de $\langle \mathcal{A}, h\rangle$.*

Demostración: Dado el argumento conclusivo $\langle \mathcal{A}, h\rangle$, supongamos por el absurdo que $\exists r, r' \in \Delta$ tal que $r' \neq r$ y tanto r como r' son una regla rebatible final de $\langle \mathcal{A}, h\rangle$. Entonces, por Definición 5.18 tenemos que $r =$ "$h \prec Cuerpo$", $r' =$ "$h \prec Cuerpo'$", $r \in \mathcal{A}$ y $r' \in \mathcal{A}$. Es decir, $\langle \mathcal{A}, h\rangle$ contiene dos reglas rebatibles diferentes para una misma conclusión. Luego, $\langle \mathcal{A}, h\rangle$ no es minimal y, por Definición 5.10, no es un argumento conclusivo, lo cual contradice la hipótesis. $\square$

Proposición 5.3 [pág. 189] *Sea $\mathcal{P}$ un programa lógico rebatible extendido y sea $\mathsf{BUAF}_{\mathcal{P}} = \langle \mathbb{A}_{\mathcal{P}}, \mathbb{R}_{\mathcal{P}}, \mathbb{S}_{\mathcal{P}}, \preceq_{\mathcal{P}} \rangle$ el marco argumentativo con backing y undercutting asociado a $\mathcal{P}$. Para todo argumento $\langle \mathcal{A}, h \rangle \in \mathbb{A}_{\mathcal{P}}$ vale que: si $\exists \langle \mathcal{B}, r \rangle \in \mathbb{A}_{\mathcal{P}}$ tal que $(\langle \mathcal{B}, r \rangle, \langle \mathcal{A}, h \rangle) \in \mathbb{B}_{k\mathcal{P}}$, entonces $\nexists \langle \mathcal{C}, r' \rangle \in \mathbb{A}_{\mathcal{P}}$ tal que $\langle \mathcal{C}, r' \rangle \neq \langle \mathcal{B}, r \rangle$ y $(\langle \mathcal{C}, r' \rangle, \langle \mathcal{A}, h \rangle) \in \mathbb{B}_{k\mathcal{P}}$.*

Demostración: Sean $\langle \mathcal{A}, h \rangle, \langle \mathcal{B}, r \rangle, \langle \mathcal{C}, r' \rangle \in \mathbb{A}_{\mathcal{P}}$ tal que $(\langle \mathcal{B}, r \rangle, \langle \mathcal{A}, h \rangle) \in \mathbb{B}_{k\mathcal{P}}$ y $(\langle \mathcal{C}, r' \rangle, \langle \mathcal{A}, h \rangle) \in \mathbb{B}_{k\mathcal{P}}$. Entonces, por Definición 5.19, tanto r como r' son una regla rebatible final de $\langle \mathcal{A}, h \rangle$, el cual es un argumento conclusivo. Luego, por Proposición 5.2, se tiene que la regla rebatible final de un argumento conclusivo es única. Por lo tanto, debe ser el caso que $r = r'$. Por otra parte, por Definición 5.19, $\langle \mathcal{B}, r \rangle$ y $\langle \mathcal{C}, r' \rangle$ son sub-argumentos propios de $\langle \mathcal{A}, h \rangle$. Finalmente, como $r = r'$ y, por Definición 5.10, $\langle \mathcal{A}, h \rangle$ es minimal, debe ser el caso que $\mathcal{B} = \mathcal{C}$ y, por lo tanto, $\langle \mathcal{B}, r \rangle = \langle \mathcal{C}, r' \rangle$. □

Proposición 5.4 [pág. 199] *Sea $\mathcal{P}$ un programa lógico rebatible extendido y $\mathsf{s} \in \{$completa, preferida, estable, grounded$\}$ una semántica de aceptabilidad. Si $\langle \emptyset, L \rangle$ es un argumento construido a partir de $\mathcal{P}$, entonces $\langle \emptyset, L \rangle$ está escépticamente aceptado con respecto a la semántica s.*

Demostración: Sea $\mathcal{P} = (\Pi, \Delta, \Sigma)$. Dado el argumento $\langle \emptyset, L \rangle$, por Definición 5.10, tenemos que $\langle \emptyset, L \rangle$ es un argumento conclusivo. Luego, dado que $\langle \emptyset, L \rangle$ posee un conjunto vacío de reglas, por Definiciones 5.10 y 5.8, existe una derivación estricta de L a partir de Π. Por definiciones 5.10, 5.11 y 5.12, para todo argumento $\langle \mathcal{A}, h \rangle$ construido a partir de $\mathcal{P}$ se verifica que $\mathcal{A} \cup \Pi$ es un conjunto no contradictorio. Por lo tanto, no será posible construir un argumento $\langle \mathcal{B}, q \rangle$ tal que los literales L y q están en desacuerdo. A partir de esto, por definiciones 5.15 y 5.17, no existirán ataques rebutting o undermining sobre $\langle \emptyset, L \rangle$. Por otra parte, como $\langle \emptyset, L \rangle$ no posee reglas rebatibles, no existirán ataques undercutting sobre $\langle \emptyset, L \rangle$. En consecuencia, por definiciones 5.19 y 5.20, no existirán argumentos que derroten a $\langle \emptyset, L \rangle$ en $\mathcal{P}$. Finalmente, por definiciones 4.10, 5.21 y 5.22, $\langle \emptyset, L \rangle$ pertenecerá a toda extensión de $\mathcal{P}$ bajo la semántica s y, por lo tanto, estará escépticamente aceptado con respecto a la semántica s. □

Proposición 5.5 [pág. 199] *Sea $\mathcal{P}$ un programa lógico rebatible extendido y $\mathsf{s} \in \{$completa, preferida, estable, grounded$\}$ una semántica de aceptabilidad. Si $\langle \emptyset, L \rangle$ es un argumento construido a partir de $\mathcal{P}$, entonces L es un literal garantizado a partir de $\mathcal{P}$ bajo la* *semántica* s.

Demostración: Directa por Proposición 5.4 y Definición 5.23. □

Proposición 5.6 [pág. 199] *Sea* $\mathcal{P}$ *un programa lógico rebatible extendido y* $\mathsf{s} \in \{$*completa, preferida, estable, grounded*$\}$ *una semántica de aceptabilidad. Si* $\langle \mathcal{A}, h \rangle$ *es un argumento de* $\mathcal{P}$ *aceptado escépticamente con respecto a la semántica* s*, entonces para todo argumento* $\langle \mathcal{B}, q \rangle$ *tal que* $\langle \mathcal{A}, h \rangle$ *derrota a* $\langle \mathcal{B}, q \rangle$ *en* $\mathcal{P}$ *o* $\langle \mathcal{B}, q \rangle$ *derrota a* $\langle \mathcal{A}, h \rangle$ *en* $\mathcal{P}$*, se verifica que* $\langle \mathcal{B}, q \rangle$ *está rechazado con respecto a la semántica* s.

Demostración: Si $\langle \mathcal{A}, h \rangle$ *está escépticamente aceptado bajo la semántica* s*, por Definición 5.22, tenemos que* $\langle \mathcal{A}, h \rangle$ *pertenece a toda extensión de* $\mathcal{P}$ *bajo la semántica* s*. Además, por Definición 5.21, las extensiones de* $\mathcal{P}$ *bajo una semántica* s *coinciden con las extensiones del* $\mathsf{BUAF}_{\mathcal{P}}$ *asociado a* $\mathcal{P}$ *bajo la semántica* s*. Luego, por Definición 4.10, toda extensión del* $\mathsf{BUAF}_{\mathcal{P}}$ *es un conjunto libre de conflictos. Esto implica que, por Definición 4.7, para toda extensión* E *del* $\mathsf{BUAF}_{\mathcal{P}}$ *bajo* s*,* $\nexists \mathcal{A}, \mathcal{B} \in \mathsf{E}$ *tal que* $\mathcal{A}$ *derrota a* $\mathcal{B}$ *en el* $\mathsf{BUAF}_{\mathcal{P}}$*. Por Definición 5.20, tenemos que las derrotas entre argumentos de* $\mathcal{P}$ *coinciden con las derrotas en el* $\mathsf{BUAF}_{\mathcal{P}}$*. Por lo tanto, para toda extensión* E *de* $\mathcal{P}$ *bajo* s*,* $\nexists \mathcal{A}, \mathcal{B} \in \mathsf{E}$ *tal que* $\mathcal{A}$ *derrota a* $\mathcal{B}$ *en* $\mathcal{P}$*. En consecuencia, por Definición 5.22, como* $\langle \mathcal{A}, h \rangle$ *pertenece a toda extensión de* $\mathcal{P}$ *bajo la semántica* s*, para todo argumento* $\langle \mathcal{B}, q \rangle$ *tal que* $\langle \mathcal{A}, h \rangle$ *derrota a* $\langle \mathcal{B}, q \rangle$ *o* $\langle \mathcal{B}, q \rangle$ *derrota a* $\langle \mathcal{A}, h \rangle$ *en* $\mathcal{P}$ *se verifica que* $\langle \mathcal{B}, q \rangle$ *está rechazado con respecto a la semántica* s*.* □

Teorema 5.1 [pág. 200] *Sea* $\mathcal{P}$ *un programa lógico rebatible extendido y* $\mathsf{s} \in \{$*completa, preferida, estable, grounded*$\}$ *una semántica de aceptabilidad. Si* L *es un literal garantizado a partir de* $\mathcal{P}$ *bajo la semántica* s*, entonces para todo literal* L' *tal que* L *y* L' *están en desacuerdo se verifica que* L' *no es un literal garantizado a partir de* $\mathcal{P}$ *bajo la semántica* s.

Demostración: Si L *es un literal garantizado a partir de* $\mathcal{P}$ *bajo la semántica* s*, entonces, por Definición 5.23, existe un argumento* $\langle \mathcal{A}, L \rangle$ *tal que* $\langle \mathcal{A}, L \rangle$ *está escépticamente aceptado con respecto a la semántica* s*. Por otra parte, para que* L' *sea un literal garantizado a partir de* $\mathcal{P}$ *bajo la semántica* s *debe existir un argumento* $\langle \mathcal{B}, L' \rangle$ *tal que* $\langle \mathcal{B}, L' \rangle$ *está escépticamente aceptado con respecto a* s*. Supongamos que existe un argumento* $\langle \mathcal{B}, L' \rangle$ *tal que los literales* L *y* L' *están en desacuerdo. Entonces, por definiciones 5.15* *y 5.17, tendremos que $\langle \mathcal{A}, L\rangle$ es un rebut o un undermine para $\langle \mathcal{B}, L'\rangle$, y $\langle \mathcal{B}, L'\rangle$ es un rebut o un undermine para $\langle \mathcal{A}, L\rangle$. Luego, por Definición 5.19, en el $\mathsf{BUAF}_{\mathcal{P}}$ asociado a $\mathcal{P}$ existirá un ataque rebutting o undermining de $\langle \mathcal{A}, L\rangle$ hacia $\langle \mathcal{B}, L'\rangle$, y existirá un ataque rebutting o undermining de $\langle \mathcal{B}, L'\rangle$ hacia $\langle \mathcal{A}, L\rangle$. Por Definición 4.2, esto dará lugar a las siguientes alternativas: $\langle \mathcal{A}, L\rangle$ derrota a $\langle \mathcal{B}, L'\rangle$ en $\mathsf{BUAF}_{\mathcal{P}}$, $\langle \mathcal{B}, L'\rangle$ derrota a $\langle \mathcal{A}, L\rangle$ en $\mathsf{BUAF}_{\mathcal{P}}$, o $\langle \mathcal{A}, L\rangle$ y $\langle \mathcal{B}, L'\rangle$ se derrotan mutuamente en $\mathsf{BUAF}_{\mathcal{P}}$. Es decir, existirá en $\mathsf{BUAF}_{\mathcal{P}}$ al menos una derrota involucrando a los argumentos $\langle \mathcal{A}, L\rangle$ y $\langle \mathcal{B}, L'\rangle$. Por otra parte, por Definición 5.20, tenemos que las derrotas en $\mathcal{P}$ coinciden con las derrotas en $\mathsf{BUAF}_{\mathcal{P}}$. En consecuencia, existirá en $\mathcal{P}$ al menos una derrota involucrando a los argumentos $\langle \mathcal{A}, L\rangle$ y $\langle \mathcal{B}, L'\rangle$. Finalmente, como $\langle \mathcal{A}, L\rangle$ está escépticamente aceptado con respecto a la semántica s, por Proposición 5.6 tenemos que $\langle \mathcal{B}, L'\rangle$ está rechazado con respecto a la semántica s. Por lo tanto, por Definición 5.23, L' no es un literal garantizado a partir de $\mathcal{P}$ bajo la semántica s.* □

Proposición 5.7 [pág. 211] *Sea $\mathcal{P}$ un programa lógico rebatible extendido. Si $\langle \emptyset, L\rangle$ es un argumento construido a partir de $\mathcal{P}$, entonces $\langle \emptyset, L\rangle$ es un argumento aceptado de $\mathcal{P}$.*

Demostración: Sea $\mathcal{P} = (\Pi, \Delta, \Sigma)$. Dado el argumento $\langle \emptyset, L\rangle$, por Definición 5.10, tenemos que $\langle \emptyset, L\rangle$ es un argumento conclusivo. Luego, dado que $\langle \emptyset, L\rangle$ posee un conjunto vacío de reglas, por Definiciones 5.10 y 5.8, existe una derivación estricta de L a partir de Π. Por definiciones 5.10, 5.11 y 5.12, para todo argumento $\langle \mathcal{A}, h\rangle$ construido a partir de $\mathcal{P}$ se verifica que $\mathcal{A} \cup \Pi$ es un conjunto no contradictorio. Por lo tanto, no será posible construir un argumento $\langle \mathcal{B}, q\rangle$ tal que los literales L y q están en desacuerdo. A partir de esto, por definiciones 5.15 y 5.17, no existirán ataques rebutting o undermining sobre $\langle \emptyset, L\rangle$. Por otra parte, como $\langle \emptyset, L\rangle$ no posee reglas rebatibles, no existirán ataques undercutting sobre $\langle \emptyset, L\rangle$. En consecuencia, por definiciones 5.19 y 5.20, no existirán argumentos que derroten a $\langle \emptyset, L\rangle$ en $\mathcal{P}$. Luego, por definiciones 5.24, 5.26 y 5.27, el árbol de dialéctica $\mathcal{T}_{\langle \emptyset, L\rangle}$ poseerá únicamente el argumento $\langle \emptyset, L\rangle$. Luego, por Definición 5.28 la raíz del árbol de dialéctica marcado $\mathcal{T}^*_{\langle \emptyset, L\rangle}$ estará marcada como $\mathbf{U}$. Finalmente, por Definición 5.29, $\langle \emptyset, L\rangle$ es un argumento aceptado de $\mathcal{P}$. □

Proposición 5.8 [pág. 211] *Sea $\mathcal{P}$ un programa lógico rebatible extendido. Si $\langle \emptyset, L\rangle$ es un argumento construido a partir de $\mathcal{P}$, entonces L es un literal garantizado a partir de $\mathcal{P}$.*

Demostración: Directa por Proposición 5.7 y Definición 5.30. □

Proposición 5.9 [pág. 212] *Sea $\mathcal{P}$ un programa lógico rebatible extendido. Si $\langle\mathcal{A}, h\rangle$ es un argumento aceptado de $\mathcal{P}$, entonces para todo argumento $\langle\mathcal{B}, q\rangle$ tal que $\langle\mathcal{A}, h\rangle$ derrota a $\langle\mathcal{B}, q\rangle$ en $\mathcal{P}$ se verifica que $\langle\mathcal{A}, h\rangle$ está marcado como* **U** *en el árbol de dialéctica marcado $\mathcal{T}^*_{\langle\mathcal{B}, q\rangle}$.*

Demostración: Sea $\mathcal{P} = (\Pi, \Delta, \Sigma)$. Si $\langle\mathcal{A}, h\rangle$ es un argumento aceptado de $\mathcal{P}$, entonces, por Definición 5.29, $\langle\mathcal{A}, h\rangle$ está marcado como **U** *en el árbol de dialéctica marcado $\mathcal{T}^*_{\langle\mathcal{A}, h\rangle}$. Si $\langle\mathcal{A}, h\rangle$ no posee derrotadores en $\mathcal{T}^*_{\langle\mathcal{A}, h\rangle}$, entonces tampoco poseerá derrotadores en $\mathcal{T}^*_{\langle\mathcal{B}, q\rangle}$ y, por lo tanto, estará marcado como* **U** *en $\mathcal{T}^*_{\langle\mathcal{B}, q\rangle}$.*
Consideremos ahora el caso en que $\langle\mathcal{A}, h\rangle$ posee derrotadores en $\mathcal{T}^*_{\langle\mathcal{A}, h\rangle}$ y supongamos que $\langle\mathcal{A}, h\rangle$ está marcado como **D** *en $\mathcal{T}^*_{\langle\mathcal{B}, q\rangle}$. Entonces, debe ser el caso que algún argumento $\langle\mathcal{C}, p\rangle$ de apoyo para $\langle\mathcal{A}, h\rangle$ que está marcado como* **U** *en $\mathcal{T}^*_{\langle\mathcal{A}, h\rangle}$ no aparezca en $\mathcal{T}^*_{\langle\mathcal{B}, q\rangle}$. Esto conduce a una de las siguientes alternativas:*

- *El argumento $\langle\mathcal{D}, t\rangle$ al que derrota $\langle\mathcal{C}, p\rangle$ no aparece en $\mathcal{T}^*_{\langle\mathcal{B}, q\rangle}$. Luego, la ausencia de $\langle\mathcal{C}, p\rangle$ no afecta el marcado de $\langle\mathcal{A}, h\rangle$ y, por lo tanto, $\langle\mathcal{A}, h\rangle$ también estará marcado como* **U** *en $\mathcal{T}^*_{\langle\mathcal{B}, q\rangle}$, lo cual contradice la suposición.*

- *El argumento $\langle\mathcal{C}, p\rangle$ no es consistente con el resto de los argumentos de interferencia de la línea argumentativa a la que pertenecería en $\mathcal{T}^*_{\langle\mathcal{B}, q\rangle}$. Sin embargo, esto no es posible ya que dicho conjunto se corresponde con el conjunto de argumentos de apoyo de la línea argumentativa a la que pertenece $\langle\mathcal{C}, p\rangle$ en $\mathcal{T}^*_{\langle\mathcal{A}, h\rangle}$. Por lo tanto, si $\langle\mathcal{C}, p\rangle$ es consistente con el conjunto de argumentos de apoyo de la línea argumentativa a la que pertenece en $\mathcal{T}^*_{\langle\mathcal{A}, h\rangle}$, también será consistente con el conjunto de argumentos de interferencia de la línea argumentativa a la que pertenece en $\mathcal{T}^*_{\langle\mathcal{B}, q\rangle}$.*

- *El argumento $\langle\mathcal{C}, p\rangle$ es un sub-argumento de alguno de los argumentos que aparecen con anterioridad en la línea argumentativa a la que pertenecería en $\mathcal{T}^*_{\langle\mathcal{B}, q\rangle}$. Esto da lugar a dos posibles situaciones:*

  - *$\langle\mathcal{C}, q\rangle$ es sub-argumento un argumento distinto de $\langle\mathcal{B}, q\rangle$ que aparece con anterioridad en la línea argumentativa en $\mathcal{T}^*_{\langle\mathcal{B}, q\rangle}$. Sin embargo, esto no es posible dado que, a excepción de $\langle\mathcal{B}, q\rangle$, todo argumento que aparece con anterioridad en la línea argumentativa a la que pertenecería $\langle\mathcal{C}, p\rangle$ en $\mathcal{T}^*_{\langle\mathcal{B}, q\rangle}$ aparece también* *(con anterioridad a $\langle \mathcal{C}, p \rangle$) en la línea argumentativa a la que pertenece $\langle \mathcal{C}, p \rangle$ en $\mathcal{T}^*_{\langle \mathcal{A}, h \rangle}$.*

- *$\langle \mathcal{C}, p \rangle$ es sub-argumento de $\langle \mathcal{B}, q \rangle$. Como $\langle \mathcal{C}, p \rangle$ derrota a un argumento $\langle \mathcal{D}, t \rangle$ en $\mathcal{T}^*_{\langle \mathcal{B}, q \rangle}$, para cualquier tipo de derrota el conjunto $\mathcal{B} \cup \mathcal{D}$ será conflictivo, en el siguiente sentido:*

  * *Si la derrota de $\langle \mathcal{C}, p \rangle$ sobre $\langle \mathcal{D}, t \rangle$ corresponde a una derrota primaria en el $\mathsf{BUAF}_{\mathcal{P}}$ asociado a $\mathcal{P}$, puede ser el caso que $\langle \mathcal{C}, p \rangle$ sea un rebut, un undermine, o un undercut para $\langle \mathcal{D}, t \rangle$. Si $\langle \mathcal{C}, p \rangle$ es un rebut o un undermine para $\langle \mathcal{D}, t \rangle$, entonces el conjunto $\Pi \cup \mathcal{C} \cup \mathcal{D}$ será contradictorio. Luego, como $\langle \mathcal{C}, p \rangle$ es un sub-argumento de $\langle \mathcal{B}, q \rangle$, por Definición 5.13, tenemos que $\mathcal{C} \subseteq \mathcal{B}$. Por lo tanto, $\Pi \cup \mathcal{B} \cup \mathcal{D}$ será un conjunto contradictorio y, por definiciones 5.26, 5.27, y 5.28, $\langle \mathcal{D}, t \rangle$ no aparecerá en $\mathcal{T}^*_{\langle \mathcal{B}, q \rangle}$. De este modo, $\langle \mathcal{A}, h \rangle$ quedará marcado como **U** en $\mathcal{T}^*_{\langle \mathcal{B}, q \rangle}$, lo cual contradice la suposición. Si, por el contrario, $\langle \mathcal{C}, p \rangle$ es un undercut para $\langle \mathcal{D}, t \rangle$, entonces el conjunto $\mathcal{C} \cup \mathcal{D}$ contiene simultáneamente una regla rebatible R y una regla de undercut para R. Nuevamente, como $\langle \mathcal{C}, p \rangle$ es un sub-argumento de $\langle \mathcal{B}, q \rangle$, por Definición 5.13, tenemos que $\mathcal{C} \subseteq \mathcal{B}$. Luego, $\mathcal{B} \cup \mathcal{D}$ contiene simultáneamente una regla rebatible R y una regla de undercut para R. Por lo tanto, por definiciones 5.26, 5.27, y 5.28, $\langle \mathcal{D}, t \rangle$ no aparecerá en $\mathcal{T}^*_{\langle \mathcal{B}, q \rangle}$. En consecuencia, $\langle \mathcal{A}, h \rangle$ quedará marcado como **U** en $\mathcal{T}^*_{\langle \mathcal{B}, q \rangle}$, lo cual contradice la suposición.*
  * *Si la derrota de $\langle \mathcal{C}, p \rangle$ sobre $\langle \mathcal{D}, t \rangle$ corresponde a una derrota implícita en el $\mathsf{BUAF}_{\mathcal{P}}$ asociado a $\mathcal{P}$, tendremos que $p = t = R$, donde R es una regla rebatible de $\mathcal{P}$ (i. e., $R \in (\Pi, \Delta)$). Por lo tanto, $\langle \mathcal{C}, p \rangle$ es un argumento de backing para R y $\langle \mathcal{D}, t \rangle$ es un argumento de undercutting para R o viceversa. En cualquiera de los dos casos, el conjunto $\mathcal{C} \cup \mathcal{D}$ contendrá simultáneamente una regla de backing y una regla de undercutting para R. Luego, como $\langle \mathcal{C}, p \rangle$ es un sub-argumento de $\langle \mathcal{B}, q \rangle$, por Definición 5.13, tenemos que $\mathcal{C} \subseteq \mathcal{B}$. De este modo, $\mathcal{B} \cup \mathcal{D}$ contendrá simultáneamente una regla de backing y una regla de undercut para R. Por lo tanto, por definiciones 5.26, 5.27, y 5.28, $\langle \mathcal{D}, t \rangle$ no aparecerá en $\mathcal{T}^*_{\langle \mathcal{B}, q \rangle}$. En consecuencia, $\langle \mathcal{A}, h \rangle$ quedará marcado como **U** en $\mathcal{T}^*_{\langle \mathcal{B}, q \rangle}$, lo cual contradice la suposición.*
  * *Si la derrota de $\langle \mathcal{C}, p \rangle$ sobre $\langle \mathcal{D}, t \rangle$ corresponde a una derrota indirec-*

ta en el $\mathsf{BUAF}_{\mathcal{P}}$ asociado a $\mathcal{P}$, tendremos que $\langle \mathcal{C}, p\rangle$ derrota (primariamente, implícitamente o indirectamente) a un argumento $\langle \mathcal{E}, s\rangle$ que es un backing o un sub-argumento de $\langle \mathcal{D}, t\rangle$. En particular, por definiciones 5.7, 5.10, 5.11, 5.12 y 5.19, los argumentos de backing en el $\mathsf{BUAF}_{\mathcal{P}}$ serán, en particular, sub-argumentos de los argumentos que soportan. Para el caso de derrotas indirectas anidadas, esto se reducirá a una derrota primaria o implícita de $\langle \mathcal{C}, p\rangle$ sobre un sub-argumento $\langle \mathcal{F}, g\rangle$ de $\langle \mathcal{D}, t\rangle$. En consecuencia, al igual que en los casos anteriores, el conjunto $\mathcal{C} \cup \mathcal{F}$ violará la segunda condición de la Definición 5.26. Luego, como por Definición 5.13 vale que $\mathcal{F} \subseteq \mathcal{D}$ y $\mathcal{C} \subseteq \mathcal{B}$, el conjunto $\mathcal{B} \cup \mathcal{D}$ violará la segunda condición de la Definición 5.26. Por lo tanto, por definiciones 5.27 y 5.28, $\langle \mathcal{D}, t\rangle$ no aparecerá en $\mathcal{T}^*_{\langle \mathcal{B}, q\rangle}$. En consecuencia, $\langle \mathcal{A}, h\rangle$ quedará marcado como U en $\mathcal{T}^*_{\langle \mathcal{B}, q\rangle}$, lo cual contradice la suposición.

- *El argumento $\langle \mathcal{C}, p\rangle$ conduce a dos derrotas por bloqueo consecutivas en la línea argumentativa a la que pertenecería en $\mathcal{T}^*_{\langle \mathcal{B}, q\rangle}$. Sin embargo, esto no es posible ya que, de ser así, también hubiera ocurrido un doble bloqueo en $\mathcal{T}^*_{\langle \mathcal{A}, h\rangle}$.*

Por lo tanto, en cualquiera de estos casos, $\langle \mathcal{A}, h\rangle$ estará marcado como U en $\mathcal{T}^*_{\langle \mathcal{B}, q\rangle}$. □

Proposición 5.10 [pág. 212] *Sea $\mathcal{P}$ un programa lógico rebatible extendido. Si $\langle \mathcal{A}, h\rangle$ es un argumento aceptado de $\mathcal{P}$, entonces para todo argumento $\langle \mathcal{B}, q\rangle$ tal que $\langle \mathcal{A}, h\rangle$ derrota a $\langle \mathcal{B}, q\rangle$ en $\mathcal{P}$ o $\langle \mathcal{B}, q\rangle$ derrota a $\langle \mathcal{A}, h\rangle$ en $\mathcal{P}$, se verifica que $\langle \mathcal{B}, q\rangle$ es un argumento rechazado de $\mathcal{P}$.*

Demostración: Para efectuar la demostración consideraremos dos casos en los cuales, respectivamente, $\langle \mathcal{B}, q\rangle$ derrota a $\langle \mathcal{A}, h\rangle$ en $\mathcal{P}$, y $\langle \mathcal{A}, h\rangle$ derrota a $\langle \mathcal{B}, q\rangle$ en $\mathcal{P}$:

- *Supongamos que existe un argumento $\langle \mathcal{B}, q\rangle$ tal que es un argumento aceptado de $\mathcal{P}$ y $\langle \mathcal{B}, q\rangle$ derrota a $\langle \mathcal{A}, h\rangle$ en $\mathcal{P}$. Entonces, por Proposición 5.9, $\langle \mathcal{B}, q\rangle$ está marcado como U en el árbol de dialéctica marcado $\mathcal{T}^*_{\langle \mathcal{A}, h\rangle}$. Por lo tanto, por Definición 5.29, $\langle \mathcal{A}, h\rangle$ será un argumento rechazado de $\mathcal{P}$, lo cual contradice la hipótesis.*

- *Supongamos que existe un argumento $\langle \mathcal{B}, q\rangle$ tal que es un argumento aceptado de $\mathcal{P}$ y $\langle \mathcal{A}, h\rangle$ derrota a $\langle \mathcal{B}, q\rangle$ en $\mathcal{P}$. Entonces, por Proposición 5.9 $\langle \mathcal{A}, h\rangle$ estará marcado como U en el árbol de dialéctica marcado $\mathcal{T}^*_{\langle \mathcal{B}, q\rangle}$. Por lo tanto, por Definición 5.29, $\langle \mathcal{B}, q\rangle$ será un argumento rechazado de $\mathcal{P}$, lo cual contradice la suposición.*

□

Teorema 5.2 [pág. 212] *Sea $\mathcal{P}$ un programa lógico rebatible extendido. Si L es un literal garantizado a partir de $\mathcal{P}$, entonces para todo literal L' tal que L y L' están en desacuerdo se verifica que L' no es un literal garantizado a partir de $\mathcal{P}$.*

Demostración: Si L es un literal garantizado a partir de $\mathcal{P}$, entonces, por Definición 5.30, existe un argumento $\langle \mathcal{A}, L\rangle$ tal que $\langle \mathcal{A}, L\rangle$ es un argumento aceptado de $\mathcal{P}$. Por otra parte, para que L' sea un literal garantizado a partir de $\mathcal{P}$ debe existir un argumento $\langle \mathcal{B}, L'\rangle$ tal que $\langle \mathcal{B}, L'\rangle$ es un argumento aceptado de $\mathcal{P}$. Supongamos que existe un argumento $\langle \mathcal{B}, L'\rangle$ tal que los literales L y L' están en desacuerdo. Entonces, por definiciones 5.15 y 5.17, tendremos que $\langle \mathcal{A}, L\rangle$ es un rebut o un undermine para $\langle \mathcal{B}, L'\rangle$, y $\langle \mathcal{B}, L'\rangle$ es un rebut o un undermine para $\langle \mathcal{A}, L\rangle$. Luego, por Definición 5.19, en el $\mathsf{BUAF}_{\mathcal{P}}$ asociado a $\mathcal{P}$ existirá un ataque rebutting o undermining de $\langle \mathcal{A}, L\rangle$ hacia $\langle \mathcal{B}, L'\rangle$, y existirá un ataque rebutting o undermining de $\langle \mathcal{B}, L'\rangle$ hacia $\langle \mathcal{A}, L\rangle$. Por Definición 4.2, esto dará lugar a las siguientes alternativas: $\langle \mathcal{A}, L\rangle$ derrota a $\langle \mathcal{B}, L'\rangle$ en $\mathsf{BUAF}_{\mathcal{P}}$, $\langle \mathcal{B}, L'\rangle$ derrota a $\langle \mathcal{A}, L\rangle$ en $\mathsf{BUAF}_{\mathcal{P}}$, o $\langle \mathcal{A}, L\rangle$ y $\langle \mathcal{B}, L'\rangle$ se derrotan mutuamente en $\mathsf{BUAF}_{\mathcal{P}}$. Es decir, existirá en $\mathsf{BUAF}_{\mathcal{P}}$ al menos una derrota involucrando a los argumentos $\langle \mathcal{A}, L\rangle$ y $\langle \mathcal{B}, L'\rangle$. Por otra parte, por Definición 5.20, tenemos que las derrotas en $\mathcal{P}$ coinciden con las derrotas en $\mathsf{BUAF}_{\mathcal{P}}$. En consecuencia, existirá en $\mathcal{P}$ al menos una derrota involucrando a los argumentos $\langle \mathcal{A}, L\rangle$ y $\langle \mathcal{B}, L'\rangle$. Finalmente, como $\langle \mathcal{A}, L\rangle$ es un argumento aceptado de $\mathcal{P}$, por Proposición 5.10 tenemos que $\langle \mathcal{B}, L'\rangle$ es un argumento rechazado de $\mathcal{P}$. Por lo tanto, por Definición 5.30, L' no es un literal garantizado a partir de $\mathcal{P}$.□

Apéndice B

Tabla de Acrónimos

La siguiente tabla contiene los acrónimos utilizados en esta tesis, listados en orden alfabético.

Acrónimo	**Expansión en Español (*en Inglés*)**	**Ubicación**
ADF	Marco Dialéctico Abstracto (*Abstract Dialectical Framework*)	Sección 3.5.2
AF	Marco Argumentativo Abstracto (*Abstract Argumentation Framework*)	Definición 2.1
$\mathsf{AF}_{\mathsf{BAF}}$	Marco Argumentativo Abstracto asociado al BAF (*Abstract Argumentation Framework associated with BAF*)	Sección 3.1.1
AFN	Marco Argumentativo con Soporte por Necesidad (*Argumentation Framework with Necessities*)	Definición 3.8
AFS	Marco Argumentativo Abstracto con Subargumentos (*Abstract Argumentation Framework with Subarguments*)	Definición 3.16
BADF	Marco Dialéctico Abstracto Bipolar (*Bipolar Abstract Dialectical Framework*)	Sección 3.5.2
BAF	Marco Argumentativo Bipolar (*Bipolar Argumentation Framework*)	Definición 3.1

BUAF	Marco Argumentativo con Backing y Undercutting (*Backing-Undercutting Argumentation Framework*)	Definición 4.1
CAF	Marco Argumentativo de Coaliciones asociado al BAF (*Coalition Framework associated with BAF*)	Sección 3.1.1
DeLP	Programación en Lógica Rebatible (*Defeasible Logic Programming*)	Sección 2.2.2
d-BAF	Marco Argumentativo Bipolar con Soporte Deductivo (*Bipolar Argumentation Framework with Deductive Support*)	Sección 3.2
d-EBAF	Marco Argumentativo Bipolar Extendido con Soporte Deductivo (*Extended Bipolar Argumentation Framework with Deductive Support*)	Sección 3.2.1
EAF	Marco Meta-argumentativo asociado a d-BAF (*Meta-argumentation Framework Associated with d-BAF*)	Definition 3.6
EAF$^+$	Marco Meta-argumentativo asociado a d-EBAF (*Meta-argumentation Framework associated with d-EBAF*)	Definición 3.7
EAS	Sistema Argumentativo Evidencial (*Evidential Argumentation System*)	Definición 3.12
E-DeLP	Programación en Lógica Rebatible Extendida (*Extended Defeasible Logic Programming*)	Sección 5.1
SABR	Sistema Argumentativo Basado en Reglas	Sección 2.2

Bibliografía

[ABP05] AMGOUD, L., BONNEFON, J.-F., AND PRADE, H. An argumentation-based approach to multiple criteria decision. In *8th European Conference on Symbolic and Quantitative Approaches to Reasoning with Uncertainty* (2005), pp. 269–280.

[AC02a] AMGOUD, L., AND CAYROL, C. Inferring from inconsistency in preference-based argumentation frameworks. *Journal of Automated Reasoning 29*, 2 (2002), 125–169.

[AC02b] AMGOUD, L., AND CAYROL, C. A reasoning model based on the production of acceptable arguments. *Annals of Mathematics and Artificial Intelligence 34*, 1-3 (2002), 197–215.

[ACLS04] AMGOUD, L., CAYROL, C., AND LAGASQUIE-SCHIEX, M.-C. On the bipolarity in argumentation frameworks. In *10th International Workshop on Non-monotonic Reasoning* (2004), pp. 1–9.

[ACLSL08] AMGOUD, L., CAYROL, C., LAGASQUIE-SCHIEX, M.-C., AND LIVET, P. On bipolarity in argumentation frameworks. *International Journal of Intelligent Systems 23*, 10 (2008), 1062–1093.

[AK07] AMGOUD, L., AND KACI, S. An argumentation framework for merging conflicting knowledge bases. *Int. Journal of Approximate Reasoning 45*, 2 (2007), 321–340.

[AMP02] AMGOUD, L., MAUDET, N., AND PARSONS, S. An argumentation-based semantics for agent communication languages. In *15th European Conference on Artificial Intelligence* (2002), pp. 38–42.

[AP93a] Alferes, J. J., and Pereira, L. M. Contradiction: When avoidance equals removal - part i. In *ELP* (1993), pp. 11–23.

[AP93b] Alferes, J. J., and Pereira, L. M. Contradiction: When avoidance equals removal - part ii. In *ELP* (1993), pp. 268–281.

[APM00] Amgoud, L., Parsons, S., and Maudet, N. Arguments, dialogue, and negotiation. In *14th European Conference on Artificial Intelligence* (2000), pp. 338–342.

[AV11a] Amgoud, L., and Vesic, S. A new approach for preference-based argumentation frameworks. *Annals of Mathematics and Artificial Intelligence 63*, 2 (2011), 149–183.

[AV11b] Amgoud, L., and Vesic, S. Two roles of preferences in argumentation frameworks. In *11th European Conference on Symbolic and Quantitative Approaches to Reasoning with Uncertainty (ECSQARU 2011)* (2011), W. Liu, Ed., vol. 6717 of *Lecture Notes in Computer Science*, Springer, pp. 86–97.

[AV12] Amgoud, L., and Vesic, S. A formal analysis of the role of argumentation in negotiation dialogues. *J. Log. Comput. 22*, 5 (2012), 957–978.

[AyC14] *Argument & Computation - Special Issue: Tutorials on Structured Argumentation*, vol. 5 (1). Taylor & Francis Online, 2014.

[BCD07] Bench-Capon, T. J. M., and Dunne, P. E. Argumentation in artificial intelligence. *Artificial Intelligence 171*, 10-15 (2007), 619–641.

[BCG11] Baroni, P., Caminada, M., and Giacomin, M. An introduction to argumentation semantics. *Knowledge Eng. Review 26*, 4 (2011), 365–410.

[BCGG09] Baroni, P., Cerutti, F., Giacomin, M., and Guida, G. Encompassing attacks to attacks in abstract argumentation frameworks. In *European Conf. on Symbolic and Quantitative Approaches to Reasoning with Uncertainty (ECSQARU)* (2009), pp. 83–94.

[BCGG11] Baroni, P., Cerutti, F., Giacomin, M., and Guida, G. AFRA: Argumentation Framework with Recursive Attacks. *International Journal of Approximate Reasoning 52*, 1 (2011), 19–37.

[BDKP02] Benferhat, S., Dubois, D., Kaci, S., and Prade, H. Bipolar representation and fusion of preferences on the possibilistic logic framework. In *8th International Conference on Principles of Knowledge Representation and Reasoning* (2002), pp. 421–448.

[Ber01] Berge, C. *Graphs and Hypergraphs*. Dover Publications, September 2001.

[BGG05] Baroni, P., Giacomin, M., and Guida, G. SCC-recursiveness: a general schema for argumentation semantics. *Artif. Intell. 168*, 1-2 (2005), 162–210.

[BGvdTV09] Boella, G., Gabbay, D. M., van der Torre, L. W., and Villata, S. Meta-argumentation modelling i: Methodology and techniques. *Studia Logica 93*, 2-3 (2009), 297–355.

[BGvdTV10] Boella, G., Gabbay, D. M., van der Torre, L. W. N., and Villata, S. Support in abstract argumentation. In *3rd International Conference on Computational Models of Argument* (2010), P. Baroni, F. Cerutti, M. Giacomin, and G. R. Simari, Eds., vol. 216 of *Frontiers in Artificial Intelligence and Applications*, IOS Press, pp. 111–122.

[BH01] Besnard, P., and Hunter, A. A logic-based theory of deductive arguments. *Artificial Intelligence 128*, 1-2 (2001), 203–235.

[BH08] Besnard, P., and Hunter, A. *Elements of Argumentation*. MIT Press, 2008.

[BH09a] Besnard, P., and Hunter, A. Argumentation based on classical logic. In *Argumentation in Artificial Intelligence*, G. Simari and I. Rahwan, Eds. Springer US, 2009, pp. 133–152.

[BH09b] Black, E., and Hunter, A. An inquiry dialogue system. *Autonomous Agents and Multi-Agent Systems 19*, 2 (2009), 173–209.

[BNR12] Boudhar, I., Nouioua, F., and Risch, V. Handling preferences in argumentation frameworks with necessities. In *4th International Conference on Agents and Artificial Intelligence* (2012), pp. 340–345.

[BNT08] Brewka, G., Niemelä, I., and Truszczynski, M. Preferences and nonmonotonic reasoning. *AI Magazine 29*, 4 (2008), 69–78.

[BW10] Brewka, G., and Woltran, S. Abstract dialectical frameworks. In *12th International Conference on the Principles of Knowledge Representation and Reasoning* (2010), pp. 102–111.

[Cam06a] Caminada, M. On the issue of reinstatement in argumentation. In *JELIA* (2006), pp. 111–123.

[Cam06b] Caminada, M. Semi-stable semantics. In *COMMA* (2006), pp. 121–130.

[Cam07] Caminada, M. An algorithm for computing semi-stable semantics. In *ECSQARU* (2007), pp. 222–234.

[CCS05] Capobianco, M., Chesñevar, C. I., and Simari, G. R. Argumentation and the dynamics of warranted beliefs in changing environments. *Autonomous Agents and Multi-Agent Systems 11*, 2 (2005), 127–151.

[CGFE01] Coulson, A. S., Glasspool, D. W., Fox, J., and Emery, J. RAGs: A novel approach to computerized genetic risk assessment and decision support from pedigrees. *Methods of Information in Medicine* (2001).

[CGGS13a] Cohen, A., Gottifredi, S., García, A. J., and Simari, G. R. Recursive attack and support in abstract argumentation frameworks. In *Working Papers of the IJCAI-2013 Workshop on Weighted Logics for Artificial Intelligence (WL4AI 2013)* (Beijing, China, August 2013), pp. 42–49.

[CGGS13b] Cohen, A., Gottifredi, S., García, A. J., and Simari, G. R. A survey of different approaches to support in argumentation systems. *The Knowledge Engineering Review FirstView* (9 2013), 1–38.

[CGS10] Cohen, A., García, A. J., and Simari, G. R. Extending DeLP with attack and support for defeasible rules. In *12th Ibero-American Conference on Artificial Intelligence (IBERAMIA 2010)* (2010), A. F. Kuri-Morales and G. R. Simari, Eds., vol. 6433 of *Lecture Notes in Computer Science*, Springer, pp. 90–99.

[CGS11a] Cohen, A., García, A. J., and Simari, G. R. An argumentation framework with backing and undercutting. In *XVII Congreso Argentino de Ciencias de la Computación (CACIC 2011)* (La Plata, Argentina, October 2011), pp. 21–30.

[CGS11b] Cohen, A., García, A. J., and Simari, G. R. Backing and undercutting in defeasible logic programming. In *11th European Conference on Symbolic and Quantitative Approaches to Reasoning with Uncertainty* (2011), pp. 50–61.

[CGS12a] Cohen, A., García, A. J., and Simari, G. R. An argumentation framework with backing and undercutting. In *Computer Science & Technology Series: XVII Argentine Congress of Computer Science Selected Papers*. Editorial de la Universidad Nacional de La Plata (Edulp), 2012, pp. 31–41.

[CGS12b] Cohen, A., García, A. J., and Simari, G. R. Backing and undercutting in abstract argumentation frameworks. In *7th International Symposium on Foundations of Information and Knowledge Systems* (2012), pp. 107–123.

[CLS05] Cayrol, C., and Lagasquie-Schiex, M.-C. On the acceptability of arguments in bipolar argumentation frameworks. In *8th European Conference on Symbolic and Quantitative Approaches to Reasoning with Uncertainty* (2005), pp. 378–389.

[CLS07] Cayrol, C., and Lagasquie-Schiex, M.-C. Coalitions of arguments in bipolar argumentation frameworks. In *7th International Workshop on Computational Models of Natural Argument* (2007), pp. 14–20.

[CLS09] Cayrol, C., and Lagasquie-Schiex, M.-C. Bipolar abstract argumentation systems. In *Argumentation in Artificial Intelligence*, G. R. Simari and I. Rahwan, Eds. Springer US, 2009, pp. 65–84.

[CLS10] Cayrol, C., and Lagasquie-Schiex, M.-C. Coalitions of arguments: A tool for handling bipolar argumentation frameworks. *International Journal of Intelligent Systems 25*, 1 (2010), 83–109.

[CLS11] Cayrol, C., and Lagasquie-Schiex, M.-C. Bipolarity in argumentation graphs: Towards a better understanding. In *5th International Conference on Scalable Uncertainty Management* (2011), pp. 137–148.

[CLS13] Cayrol, C., and Lagasquie-Schiex, M.-C. Bipolarity in argumentation graphs: Towards a better understanding. *Int. J. Approx. Reasoning 54*, 7 (2013), 876–899.

[CMDM05] Coste-Marquis, S., Devred, C., and Marquis, P. Prudent semantics for argumentation frameworks. In *ICTAI* (2005), pp. 568–572.

[CML00] Chesñevar, C. I., Maguitman, A. G., and Loui, R. P. Logical models of argument. *ACM Computing Surveys 32*, 4 (2000), 337–383.

[DFB08] Dubois, D., Fargier, H., and Bonnefon, J.-F. On the qualitative comparison of decisions having positive and negative features. *Journal of Artificial Intellicence Research 32* (2008), 385–417.

[DKT06] Dung, P. M., Kowalski, R. A., and Toni, F. Dialectic proof procedures for assumption-based, admissible argumentation. *Artificial Intelligence 170*, 2 (2006), 114–159.

[DMT06] Dung, P. M., Mancarella, P., and Toni, F. A dialectic procedure for sceptical, assumption-based argumentation. In *COMMA* (2006), pp. 145–156.

[DP05] Dubois, D., and Prade, H. A bipolar possibilistic representation of knowledge and preferences and its applications. In *6th International Workshop on Fuzzy Logic and Applications* (2005), pp. 1–10.

[Dun95] 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 (1995), 321–358.

[FEGS07] Ferretti, E., Errecalde, M., García, A. J., and Simari, G. R. An application of defeasible logic programming to decision making in a robotic environment. In *LPNMR* (2007), pp. 297–302.

[Gar00] García, A. J. *Programación en Lógica Rebatible: Definición, Semántica Operacional y Paralelismo*. PhD thesis, Universidad Nacional del Sur, Bahía Blanca, Argentina, 2000.

[GCS08] Gómez, S. A., Chesñevar, C. I., and Simari, G. R. Defeasible reasoning in web-based forms through argumentation. *International Journal of Information Technology and Decision Making 7*, 1 (2008), 71–101.

[GGS10] Gottifredi, S., García, A. J., and Simari, G. R. Query-based argumentation in agent programming. In *IBERAMIA* (2010), pp. 284–295.

[GL88] Gelfond, M., and Lifschitz, V. The stable model semantics for logic programming. In *ICLP/SLP* (1988), pp. 1070–1080.

[GRTS07] García, A. J., Rotstein, N. D., Tucat, M., and Simari, G. R. An argumentative reasoning service for deliberative agents. In *KSEM* (2007), pp. 128–139.

[GS04] García, A. J., and Simari, G. R. Defeasible logic programming: An argumentative approach. *Theory and Practice of Logic Programming 4*, 1-2 (2004), 95–138.

[Kon88] Konolige, K. Defeasible argumentation in reasoning about events. In *ISMIS* (1988), pp. 380–390.

[Lif96] Lifschitz, V. Foundations of logic programs. In *Principles of Knowledge Representation*, G. Brewka, Ed. CSLI Pub., 1996, pp. 69–128.

[MBC11] Modgil, S., and Bench-Capon, T. J. M. Metalevel argumentation. *J. Log. Comput. 21*, 6 (2011), 959–1003.

[MGS06] Martínez, D. C., García, A. J., and Simari, G. R. On acceptability in abstract argumentation frameworks with an extended defeat relation. In *1st International Conference on Computational Models of Argument* (2006), pp. 273–278.

[MGS07] Martínez, D. C., García, A. J., and Simari, G. R. Modelling well-structured argumentation lines. In *20th International Joint Conference on Artificial Intelligence* (2007), pp. 465–470.

[MGS08a] Martínez, D. C., García, A. J., and Simari, G. R. An abstract argumentation framework with varied-strength attacks. In *11th International Conference on Principles of Knowledge Representation and Reasoning* (2008), pp. 135–144.

[MGS08b] Martínez, D. C., García, A. J., and Simari, G. R. Strong and weak forms of abstract argument defense. In *2nd International Conference on Computational Models of Argument* (2008), pp. 216–227.

[Mod09] Modgil, S. Reasoning about preferences in argumentation frameworks. *Artificial Intelligence 173*, 9-10 (2009), 901–934.

[NP06] Nielsen, S. H., and Parsons, S. A generalization of Dung's abstract framework for argumentation: Arguing with sets of attacking arguments. In *Third International Workshop on Argumentation in Multi-Agent Systems* (2006), pp. 54–73.

[NR10] Nouioua, F., and Risch, V. Bipolar argumentation frameworks with specialized supports. In *22th IEEE International Conference on Tools with Artificial Intelligence* (2010), pp. 215–218.

[NR11] Nouioua, F., and Risch, V. Argumentation frameworks with necessities. In *5th International Conference on Scalable Uncertainty Management* (2011), pp. 163–176.

[Nut88] Nute, D. Defeasible reasoning: a philosophical analysis in prolog. In *Aspects of Artificial Intelligence*, J. H. Fetzer, Ed. Kluwer Academic Pub., 1988, pp. 251–288.

[Nut94] Nute, D. Defeasible logic. In *Handbook of Logic in Artificial Intelligence and Logic Programming, Vol 3*, D. Gabbay, C. Hogger, and J.A.Robinson, Eds. Oxford University Press, 1994, pp. 355–395.

[ON08] Oren, N., and Norman, T. J. Semantics for evidence-based argumentation. In *2nd International Conference on Computational Models of Argument* (2008), pp. 276–284.

[Pol87] Pollock, J. L. Defeasible reasoning. *Cognitive Science 11*, 4 (1987), 481–518.

[Pol94] POLLOCK, J. L. Justification and defeat. *Artificial Intelligence 67*, 2 (1994), 377–407.

[Pol95] POLLOCK, J. L. *Cognitive Carpentry: A Blueprint for How to Build a Person.* MIT Press, Cambridge, MA, USA, 1995.

[Pol01] POLLOCK, J. L. Defeasible reasoning with variable degrees of justification. *Artificial Intelligence 133*, 1-2 (2001), 233–282.

[Poo85] POOLE, D. On the comparison of theories: Preferring the most specific explanation. In *IJCAI* (1985), pp. 144–147.

[Pra09] PRAKKEN, H. An abstract framework for argumentation with structured arguments. *Journal of Argument and Computation 1* (2009), 93–124.

[PS96] PRAKKEN, H., AND SARTOR, G. A dialectical model of assessing conflicting arguments in legal reasoning. *Artificial Intelligence and Law 4*, 3-4 (1996), 331–368.

[PS97] PRAKKEN, H., AND SARTOR, G. Argument-based extended logic programming with defeasible priorities. *Journal of Applied Non-Classical Logics 7*, 1 (1997).

[PS02] PRAKKEN, H., AND SARTOR, G. The role of logic in computational models of legal argument: A critical survey. In *Computational Logic: Logic Programming and Beyond* (2002), pp. 342–381.

[PSJ98] PARSONS, S., SIERRA, C., AND JENNINGS, N. R. Agents that reason and negotiate by arguing. *Journal of Logic and Computation 8*, 3 (1998), 261–292.

[PSM11] PARSONS, S., SKLAR, E., AND MCBURNEY, P. Using argumentation to reason with and about trust. In *Proceedings of the 8th International Workshop on Argumentation in Multiagent Systems* (Taipei, Taiwan, 2011).

[PV02] PRAKKEN, H., AND VREESWIJK, G. Logics for defeasible argumentation. In *Handbook of Philosophical Logic*, D. Gabbay and F. Guenthner, Eds., vol. 4. Kluwer Academic Pub., 2002, pp. 218–319.

[RA06] Rahwan, I., and Amgoud, L. An argumentation based approach for practical reasoning. In *AAMAS* (2006), pp. 347–354.

[RGS07] Rotstein, N. D., Garcáa, A. J., and Simari, G. R. Reasoning from desires to intentions: A dialectical framework. In *AAAI Conference on Artificial Intelligence* (2007), pp. 136–141.

[RMF+08] Rotstein, N. D., Moguillansky, M. O., Falappa, M. A., García, A. J., and Simari, G. R. Argument theory change: Revision upon warrant. In *2nd International Conference on Computational Models of Argument* (2008), pp. 336–347.

[RMGS10] Rotstein, N. D., Moguillansky, M. O., García, A. J., and Simari, G. R. A Dynamic Argumentation Framework. In *3rd International Conference on Computational Models of Argument* (2010), pp. 427–438.

[RRJ+03] Rahwan, I., Ramchurn, S. D., Jennings, N. R., McBurney, P., Parsons, S., and Sonenberg, L. Argumentation-based negotiation. *Knowledge Eng. Review 18*, 4 (2003), 343–375.

[RS09] Rahwan, I., and Simari, G. R. *Argumentation in Artificial Intelligence.* Springer, 2009.

[SL92] Simari, G. R., and Loui, R. P. A mathematical treatment of defeasible reasoning and its implementation. *Artificial Intelligence 53*, 2-3 (1992), 125–157.

[TCM+12] Tang, Y., Cai, K., McBurney, P., Sklar, E., and Parsons, S. Using argumentation to reason about trust and belief. *J. Log. Comput. 22*, 5 (2012), 979–1018.

[TGG+12] Tamargo, L. H., Gottifredi, S., García, A. J., Falappa, M. A., and Simari, G. R. Deliberative DeLP agents with multiple informants. *Inteligencia Artificial, Revista Iberoamericana de Inteligencia Artificial 15*, 49 (2012), 13–30.

[Tou58] Toulmin, S. E. *The Uses of Argument.* Cambridge University Press, 1958.

[Ver02] Verheij, B. On the existence and multiplicity of extensions in dialectical argumentation. In *9th International Workshop on Non-monotonic Reasoning* (2002), pp. 416–425.

[Ver03a] Verheij, B. DefLog: On the logical interpretation of prima facie justified assumptions. *Journal of Logic and Computation 13*, 3 (2003), 319–346.

[Ver03b] Verheij, B. Dialectical argumentation with argumentation schemes: An approach to legal logic. *Artificial Intelligence and Law 11*, 2-3 (2003), 167–195.

[Ver05] Verheij, B. Evaluating arguments based on Toulmin's scheme. *Argumentation 19*, 3 (2005), 347–371.

[Ver07] Verheij, B. A labeling approach to the computation of credulous acceptance in argumentation. In *IJCAI* (2007), pp. 623–628.

[Ver09] Verheij, B. The Toulmin argument model in artificial intelligence. or: How semi-formal, defeasible argumentation schemes creep into logic. In *Argumentation in Artificial Intelligence*, I. Rahwan and G. R. Simari, Eds. Springer, 2009.

[Vre91] Vreeswijk, G. The feasibility of defeat in defeasible reasoning. In *In Proceedings of the 2nd International Conference on Knowledge Representation and Reasoning* (1991), Morgan Kaufmann, pp. 526–534.

[Vre97] Vreeswijk, G. Abstract argumentation systems. *Artificial Intelligence 90*, 1-2 (1997), 225–279.

[Vre06] Vreeswijk, G. An algorithm to compute minimally grounded and admissible defence sets in argument systems. In *COMMA* (2006), pp. 109–120.

[WGM+97] Walton, R., Gierl, C., Mistry, H., Vessey, M. P., and Fox, J. Evaluation of computer support for prescribing (CAPSULE) using simulated cases. *British Medical Journal 315* (1997), 791–795.