



Sobre la definición de los modelos estructurados incompletos de argumentación

On the definition of incomplete structured argumentation frameworks

Antonio Yuste-Ginel

Departamento de Lógica y Filosofía Teórica
Universidad Complutense de Madrid
antoyust@ucm.es

RESUMEN

Este artículo versa sobre cómo instanciar los modelos abstractos incompletos de argumentación en sistemas estructurados de argumentación, en concreto en ASPIC+. El objetivo fundamental del trabajo se motiva partiendo de argumentos generales en contra de la proliferación indiscriminada de los modelos abstractos en argumentación formal y la necesidad de estudiar sus instancias para evitar el modelado *ad-hoc* de ejemplos y descubrir posibles asunciones implícitas. El énfasis del trabajo es conceptual, mostrando cómo algunas asunciones que subyacen a los modelos abstractos incompletos no son del todo ciertas cuando se estudian los detalles de sus instancias.

PALABRAS CLAVE: argumentación formal, Dung, instanciaciones, modelos abstractos, modelos estructurados, modelos incompletos, modelos con incertidumbre, sistemas estructurados.

ABSTRACT

This paper studies different options for the instantiation of incomplete abstract argumentation frameworks into ASPIC+. The main motivation of the work is rooted in general arguments that point out the danger of unlimited abstraction in argumentation models research and the need to study some instantiations in order to avoid *ad-hoc* modelling of examples and implicit assumptions behind abstract models. The work has a conceptual focus, showing how some assumptions underlying incomplete abstract argumentation frameworks are not exactly true when some of its instances are considered in detail.

KEYWORDS: Dung, formal argumentation, instantiations, abstract frameworks, structured frameworks, incomplete argumentation frameworks, frameworks with uncertainty, structured systems.

1. INTRODUCCIÓN

Contexto. La argumentación formal es un campo de investigación relativamente joven donde confluyen ideas y personas provenientes de muy distintas disciplinas: las matemáticas, las ciencias de la computación, la ingeniería informática, la filosofía y la lingüística, entre otras. A pesar de su juventud, cuenta ya con dos volúmenes de un manual monográfico (Baroni et al., 2018; Gabbay et al., 2021), con diversos congresos especializados,¹ y con una competición internacional de software,² encontrando numerosas aplicaciones en el mundo de la inteligencia artificial (Bench-Capon y Dunne, 2007). La variedad de temas dentro del campo es amplia pero el denominador común es el uso de herramientas formales (matemáticas) para el estudio de fenómenos y nociones argumentativas. En este sentido, puede entenderse como una rama de la teoría de la argumentación muy ampliamente concebida. Dentro de los estudios de argumentación formal, este trabajo se centra en los *modelos de inferencia basados en argumentos* (Prakken, 2017). La idea principal detrás de estos modelos es la de construir relaciones inferenciales entre proposiciones (normalmente no monótonas) a partir de una serie de nociones argumentativas (e.g., argumento, ataque, derrota, apoyo y aceptabilidad).

Modelos abstractos vs. modelos estructurados. Una distinción que vertebra la investigación en modelos inferenciales basados en argumentos es la que separa los modelos abstractos de aquellos estructurados. Aquí el término ‘abstracto’ se refiere a la tendencia del primer tipo de modelos a obviar deliberadamente la estructura interna y origen de los argumentos, así como la estructura y naturaleza de las relaciones dialécticas entre los mismos, para enfocarse en su lugar sobre cuestiones de alto nivel. Un claro ejemplo de estos modelos es el del ubicuo enfoque de Dung: el uso de grafos dirigidos donde los nodos representan argumentos y las flechas relaciones de derrota entre estos³ y donde distintas semánticas permiten capturar formalmente la noción intuitiva de aceptabilidad de un conjunto de argumentos. Frente a este tipo de modelos, los modelos estructurados “miran dentro” de los argumentos, dotándolos de algún tipo de estructura, normalmente basadas en una lógica subyacente. Uno de los modelos estructurados de argumentación más populares es ASPIC+ (cuyo nombre proviene del

¹ Por ejemplo, las *Conferences on Computational Models of Arguments* (COMMA) es una reunión bianual que congrega a miembros de la comunidad de todo el mundo.

² *International Competition on Computational Models of Argumentation* (ICCA).

³ Originalmente, las relaciones en los grafos de Dung se entienden en términos de *ataque* [attack] pero, como se verá más adelante, es más conveniente hacerlo aquí en términos de *derrota* [defeat].

proyecto europeo *Argumentation Service Platform with Integrated Components*). Entre otros aspectos ventajosos de ASPIC+, destacan su capacidad para simular otros sistemas estructurados de argumentación y su elaborada fundamentación teórico-filosófica (Modgil y Prakken, 2013, 2014).

La introducción de la incertidumbre. A pesar de la popularidad y la utilidad innegable de los grafos de Dung, este enfoque se ve lastrado por limitaciones expresivas muy importantes. Más allá de las relaciones de derrota con otros argumentos, y además de la naturaleza interna de la que hablamos en el párrafo anterior, parece haber otros aspectos que influyen de forma crucial la aceptabilidad de los argumentos, y a los que el modelo de Dung es ciego, véase:

- las relaciones positivas (de apoyo) entre argumentos;
- las preferencias entre argumentos (por ejemplo, basadas en el grado de confiabilidad de las premisas de cada argumento);
- las relaciones dialécticas de orden superior (por ejemplo, un argumento que ataque un movimiento dialéctico, en lugar de a otro argumento);
- la incertidumbre acerca de algunos de los elementos implicados.

Cada uno de estos puntos ha originado una corriente en la bibliografía de los modelos abstractos de argumentación (véase la Parte I de Gabbay et al., 2021). En este artículo, nos centramos en la última.

La necesidad para la argumentación formal de representar incertidumbre acerca de los argumentos de una discusión y de las relaciones de derrota existentes entre ellos emerge naturalmente en contextos multi-agente: la falta de conocimiento completo de un agente acerca de la información que otro posee hace que el primero atribuya al segundo argumentos y derrotas de forma incierta, y esto influye en cómo el primero evalúa distintos argumentos de cara a usarlos en un intercambio con el segundo. Ilustremos un poco esta idea.

Ejemplo 0. Imaginemos a un sospechoso que, al ser interrogado, baraja la posibilidad de usar dos coartadas distintas. Si el sospechoso está seguro de que su interrogador posee información para contrarrestar la primera, pero duda acerca de sus posibilidades para contrarrestar la segunda, parece que lo más razonable para el sospechoso es usar la segunda.

La introducción de la incertidumbre en el modelo de Dung se ha llevado a cabo

mediante el uso de probabilidades (véase el capítulo 7 de Gabbay et al., 2021) o cualitativamente, principalmente mediante el desarrollo de los *modelos abstractos incompletos de argumentación* (Baumeister et al., 2021, Maily, 2022), donde tanto los argumentos como las derrotas entre ellos pueden ser ciertos o inciertos, dando lugar a distintas *compleciones* [*completions*], estas son, posibles configuraciones del debate que el agente modelado atribuye a otro.

El peligro de la abstracción desmesurada. Todas las extensiones del modelo de Dung que comentamos en el párrafo anterior son abstractas: los argumentos se siguen representando como puntos de un grafo que no tienen estructura interna. La proliferación de estas extensiones ha provocado que se eleven voces críticas dentro de la comunidad. Prakken y De Winter (2018) señalan que el desarrollo indiscriminado de estos modelos se ve acompañado de dos peligros:

(P1) El análisis de ejemplos usando modelos abstractos en lugar de estructurados conduce usualmente a representaciones *ad-hoc* de los ejemplos en cuestión.

(P2) Los modelos abstractos pueden hacer asunciones implícitas que no comparten ninguna de sus instancias.

Para evitar ambos peligros, los autores proponen una solución: por cada nuevo modelo abstracto de argumentación, provéase también una instancia del mismo en un sistema estructurado aceptado por la comunidad. Como ellos señalan, el propio Dung siguió este consejo en su trabajo fundacional de 1995.

Objetivo y contribución. En un proyecto de investigación en marcha, nos hemos propuesto analizar la viabilidad de las instancias de los modelos abstractos incompletos de argumentación. Los primeros resultados se obtuvieron en (Yuste-Ginel y Proietti, 2023) donde caracterizamos la instancia de una clase específica de modelos abstractos incompletos de argumentación. En este artículo, realizamos las siguientes contribuciones al conseguimiento del objetivo general: (i) hacemos una exposición elaborada del estado de la cuestión; (ii) exploramos sistemáticamente otras opciones para instanciar los modelos abstractos incompletos de argumentación en ASPIC+; (iii) proporcionamos resultados técnicos novedosos; (iv) explicitamos una asunción implícita detrás de estos modelos no compartida por las instancias propuestas (i.e., identificamos el peligro (P2) en este contexto).

Estructura del texto. El resto de este artículo está organizado como sigue. La Sección

2 introduce las herramientas formales y conceptuales necesarias para entender el resto del trabajo. En particular, se presentan y discuten el modelo de Dung (1995), su extensión con incertidumbre cualitativa conocida como *modelos abstractos incompletos de argumentación* (Baumeister et al., 2021) y el modelo estructurado de argumentación ASPIC+ (Modgil y Prakken, 2013, 2014). En la Sección 3 concentramos las contribuciones originales del artículo, analizando en detalle las dos propuestas de instanciación lanzadas por Baumeister et al. (2021), mostrando algunos resultados relacionados, y discutiendo los mismos. Además, se estudian otras posibilidades de instanciación. Concluimos con la Sección 4 donde, después de recapitular, recogemos algunos desafíos actuales para nuestro proyecto de investigación.

2. ANTECEDENTES

2.1. MODELOS ABSTRACTOS DE ARGUMENTACIÓN DE DUNG

Un **modelo abstracto de argumentación** [*Abstract Argumentation Framework*] (**AAF**)⁴ es un grafo $\langle Arg, Der \rangle$ donde $Arg \neq \emptyset$ representa un conjunto de argumentos y $Der \subseteq Arg \times Arg$ representa una relación binaria de derrota. Usamos mayúsculas latinas, comenzando por A , para denotar argumentos (elementos de Arg). $\langle A, B \rangle \in Der$ se lee “ A derrota a B ”.

Dado un modelo abstracto $AAF = \langle Arg, Der \rangle$, existen distintas maneras intuitivamente viables de resolver los conflictos expresados por Der . La contrapartida formal de esta idea es que existen distintas *semánticas* para calcular las *extensiones* de un AAF (los subconjuntos de argumentos colectivamente aceptables). El interés de estas semánticas es ortogonal a este trabajo, así que omitimos aquí su definición formal.⁵ A modo de ejemplo, ilustramos las extensiones del siguiente AAF, que colapsan para las cuatro semánticas que definió y estudió Dung en 1995.

Ejemplo 1. El siguiente AAF captura algunos elementos relevantes del Ejemplo 0, donde A es un argumento apuntado a la culpabilidad del sospechoso, B_1 y B_2 representan las dos coartadas y C_1 y C_2 los respectivos contra-argumentos a las coartadas.

$$AAF_1 = \{\{A, B_1, B_2, C_1, C_2\}, \{\langle B_1, A \rangle, \langle B_2, A \rangle, \langle C_1, B_1 \rangle, \langle C_2, B_2 \rangle\}\}$$

⁴ En este trabajo, usamos las siglas en inglés de los nombres de los formalismos como abreviaturas de los mismos.

⁵ Para el lector interesado en el tema, recomendamos el Capítulo 4 de Baroni et al., 2018.

Nótese que AAF_1 puede representarse gráficamente como:

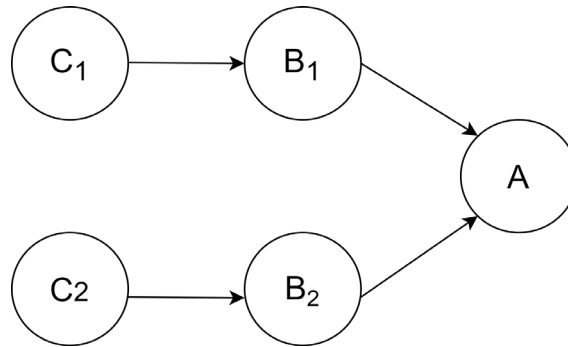


Figura 1: AAF para el Ejemplo 0

AAF_1 tiene una sola extensión para las cuatro semánticas de Dung, véase, $\{C_2, C_1, A\}$. Intuitivamente, este es el único conjunto de argumentos de AAF_1 que cumple los siguientes requisitos: (i) está libre de conflicto (no hay derrotas entre sus elementos); (ii) defiende a todos sus elementos (i.e., derrota a todos los derrotadores de sus elementos); (iii) contiene a todos los argumentos que defiende.

Usando las extensiones de un AAF pueden definirse distintas nociones de aceptabilidad de un argumento. Típicamente, se dice que un argumento es **crédulamente/escépticamente aceptable** (con respecto a AAF) si y sólo si (sysss, en adelante) pertenece a al menos una/todas las extensiones de AAF .

2.2. INTRODUCIENDO INCERTIDUMBRE CUALITATIVA EN EL MODELO DE DUNG

Un **modelo incompleto abstracto de argumentación** [*Incomplete Abstract Argumentation Framework*] (**IAAF**) es una tupla $\langle Arg^F, Arg^?, Der^F, Der^? \rangle$ donde Arg^F y $Arg^?$ son dos conjuntos disjuntos de argumentos abstractos y $Der^F, Der^? \subseteq (Arg^F \cup Arg^?) \times (Arg^F \cup Arg^?)$ son relaciones disjuntas de derrota entre estos argumentos. Arg^F (respectivamente, Der^F) representa el conjunto de argumentos (respectivamente, derrotas) fijos o conocidos mientras que $Arg^?$ (respectivamente, $Der^?$) representa el conjunto de argumentos (respectivamente, derrotas) desconocidos o inciertos.

La noción de aceptabilidad de un argumento en un IAAF está mediada por la de completión o configuración de un IAAF. Formalmente, una **completión** [*completion*] de IAAF es cualquier AAF $\langle Arg', Der' \rangle$ tal que:

- $Arg^F \subseteq Arg' \subseteq Arg^F \cup Arg^?$; y

- $Der^F \cap (Arg' \times Arg') \subseteq Der' \subseteq (Der^F \cup Der^?) \cap (Arg' \times Arg')$.

En palabras, todos los argumentos fijos son partes de la compleción, y los argumentos de la compleción son o bien fijos o bien inciertos. En cuanto a las relaciones de derrota: (i) todas las derrotas fijas forman parte de cualquier compleción (siempre que se dé la presencia del origen y del destino de esa derrota dentro de la compleción en cuestión); y (ii) las relaciones de derrota de cualquier compleción son o bien fijas o bien inciertas.

En los modelos incompletos, la noción de aceptabilidad de un argumento cobra una nueva capa de complejidad: un argumento puede ser aceptado en toda extensión (escépticamente) o en al menos una (crédulamente) de toda compleción (**necesariamente**) o en al menos una (**posiblemente**).

Interpretación epistémica de los IAAFs. Desde una perspectiva epistémica y multi-agente los IAAFs pueden interpretarse como sigue (véase Proietti y Yuste-Ginel, 2021 o Yuste-Ginel y Herzig, 2023 para una exposición más detallada). El IAAF en sí puede entenderse como el modelo que un agente (humano o artificial), llamémoslo PRO, tiene de la visión de otro agente (típicamente, su oponente, denotado mediante OPP) de cierta situación argumentativa (i.e., de cierto AAF). Los elementos de Arg^F son los argumentos de los que PRO cree que OPP es consciente.⁶ Los elementos de $Arg^?$ son los argumentos acerca de los cuales PRO duda de que OPP sea consciente. Por último, todo elemento fuera de $Arg^F \cup Arg^?$ serían argumentos de los que PRO está seguro de que OPP no es consciente.

Ejemplo 2. En el siguiente IAAF capturamos algunos aspectos del Ejemplo 0, suponiendo que el agente sospechoso duda acerca de la posibilidad de su interrogador tanto de formar el contra-argumento a la segunda coartada, es decir, C_2 , como de ver la derrota de C_2 a B_2 una vez que el primero ha sido formado.

$$IAAF_1 = \langle \{A, B_1, B_2, C_1\}, \{C_2\}, \{ \langle A, B_1 \rangle, \langle A, B_2 \rangle, \langle B_1, C_1 \rangle \}, \{ \langle C_2, B_2 \rangle \} \rangle.$$

Representamos $IAAF_1$ en la Figura 2, donde los elementos inciertos se capturan mediante líneas discontinuas. Nótese que $IAAF_1$ tiene tres compleciones y que A es posiblemente (pero no necesariamente) aceptado desde la perspectiva escéptica.

⁶ Ser consciente de un argumento [*to be aware of an argument*] es una noción que queda voluntariamente indeterminada en esta interpretación. No obstante, una posible lectura es que un agente es consciente de un argumento si es capaz de construirlo por sus propios medios y de utilizarlo en algún momento del intercambio.

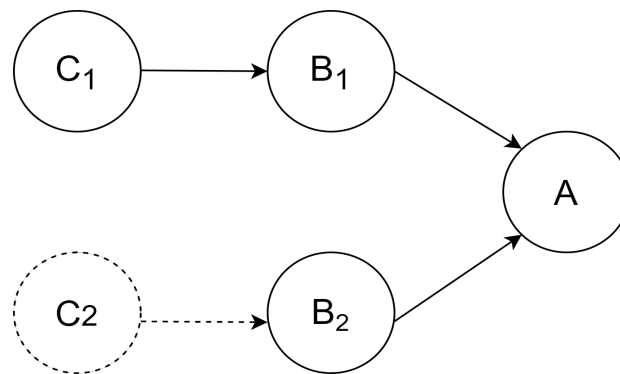


Figura 2: IAAF para el Ejemplo 0

2.3. MODELOS ESTRUCTURADOS DE ARGUMENTACIÓN

Un **modelo estructurado de argumentación** [*Structured Argumentation Framework*] (**SAF**) de ASPIC+⁷ es una tupla $\langle Len, Con, Reg, Nom, Kno, Arg, Att, Pre \rangle$ donde:

- *Len* es un **lenguaje formal** (por ejemplo, el de la lógica proposicional clásica). Usamos las meta-variables *F* y *G* (posiblemente subíndizadas) para denotar elementos arbitrarios de *Len*.
- *Con*: $Len \rightarrow \wp(Len)$ es una función que asigna a cada fórmula del lenguaje el conjunto de sus **contrarias**.
- $Reg = Reg_s \cup Reg_d$ es un conjunto de **reglas de inferencias** (secuencias finitas sobre *Len*). Asumimos que disponemos de reglas estrictas o deductivas (Reg_s) y reglas derrotables (Reg_d) y que ambos conjuntos son disjuntos.
- *Nom*: $Reg_d \rightarrow Len$ es una función que asigna “**nombres**” a las **reglas derrotables**. Podemos entender $Nom(r)$ como una proposición que asevera “la regla *r* es aplicable”.
- $Kno = Kno_a \cup Kno_p$ es una **base de conocimiento**, es decir, un conjunto de fórmulas de *Len* de donde extraeremos las premisas de los argumentos. Asumimos que disponemos de dos subconjuntos disjuntos de *Kno*, véase los axiomas o premisas inatacables (Kno_a) y las premisas corrientes (Kno_p).
- *Arg* es el conjunto de **argumentos** de SAF y se define recursivamente como sigue. Definimos además algunas funciones junto con la noción de argumento que nos permiten analizar su estructura ($Conc(A)$ denota la conclusión de *A*;

⁷ Aunque la presentación de Modgil y Prakken (2013, 2014) es ligeramente distinta a la nuestra, ambas se pueden probar equivalentes de forma sencilla.

$Prem(A)$ denota las premisas de A ; $TopRule(A)$ denota la última regla usada en la construcción de A ; y $SubArg(A)$ denota el conjunto de subargumentos de A . Arg es el menor conjunto tal que:

- $[F] \in Arg$ para cualquier $F \in Kno$; con $Prem([F]) = Conc([F]) = F$; $TopRule([F])$ queda indefinido; y $SubArg([F]) = \{[F]\}$.
- Si $A_1, \dots, A_n \in Arg$ y $\langle Conc(A_1), \dots, Conc(A_n), F \rangle \in Reg_s$ entonces $[A_1, \dots, A_n] \rightarrow^s F \in Arg$. Definimos además $Prem([A_1, \dots, A_n] \rightarrow^s F) = Prem(A_1) \cup \dots \cup Prem(A_n)$; $Conc([A_1, \dots, A_n] \rightarrow^s F) = F$; $TopRule([A_1, \dots, A_n] \rightarrow^s F) = \langle Conc(A_1), \dots, Conc(A_n), F \rangle$; y $SubArg([A_1, \dots, A_n] \rightarrow^d F) = SubArg(A_1) \cup \dots \cup SubArg(A_n)$.
- Si $A_1, \dots, A_n \in Arg$ y $\langle Conc(A_1), \dots, Conc(A_n), F \rangle \in Reg_d$ entonces $[A_1, \dots, A_n] \rightarrow^d F \in Arg$. Definimos además $Prem([A_1, \dots, A_n] \rightarrow^d F) = Prem(A_1) \cup \dots \cup Prem(A_n)$; $Conc([A_1, \dots, A_n] \rightarrow^d F) = F$; $TopRule([A_1, \dots, A_n] \rightarrow^d F) = \langle Conc(A_1), \dots, Conc(A_n), F \rangle$; y $SubArg([A_1, \dots, A_n] \rightarrow^d F) = SubArg(A_1) \cup \dots \cup SubArg(A_n)$.
- $Att \subseteq Arg \times Arg$ es una relación de **ataque** entre argumentos. ASPIC+ permite distinguir entre tres tipos de ataque. Nos desviamos de la terminología original (en inglés) y adoptamos la de Marraud (2017). Dados $A, B \in Arg$ decimos que $\langle A, B \rangle \in Att$ (A ataca a B) si y sólo si (sys, en adelante):
 - $Conc(A) \in Con(F)$ para alguna $F \in Prem(B)$ (en tal caso, decimos que A **objeta** a B (en F)); o bien
 - $Conc(A) \in Con(F)$ para algún $B' = [B_1, \dots, B_n] \rightarrow^d F \in SubArg(B)$ (en tal caso, decimos que A **refuta** a B (en B')); o bien
 - $Conc(A) \in Con(r)$ donde $r = Nom(\langle F_1, \dots, F_n \rangle)$ y $\langle F_1, \dots, F_n \rangle \in TopRule(B')$ para algún $B' \in SubArg(B)$ (en tal caso, decimos que A **recusa** a B (en B')).
- Finalmente, $Pref \subseteq Arg \times Arg$ es una relación de **preferencia** o fuerza relativa entre argumentos. Si $\langle A, B \rangle \in Pref$ entonces A es al menos tan fuerte como B . Esta relación de preferencia se puede fundar en otras más

primitivas. Por ejemplo, en una relación de preferencia entre premisas basada en la confiabilidad de las mismas. Aquí, sin embargo, nos limitamos a considerarla como primitiva.

Dados $A, B \in Arg(SAF)$ decimos que A **derrota a** B si A recusa a B o bien A objeta/refuta a B en B' y no es el caso de que $\langle B', A \rangle \in Pref$ pero $\langle A, B' \rangle \notin Pref$.

Dado SAF , denotamos mediante $Com(SAF)$ su componente Com . Por ejemplo, usamos $Arg(SAF)$ para denotar el conjunto de argumentos de SAF conforme a la anterior definición.

Dado SAF , el **modelo abstracto argumentativo asociado a SAF** (su AAF) es el par $AAF(SAF) = \langle Arg(SAF), Der(SAF) \rangle$. Esta definición establece un puente nítido entre los sistemas de argumentación estructurados y los abstractos. Además, ahora debería estar claro por qué llamamos en este trabajo derrotas a las flechas de los modelos de Dung (en lugar de ataque, como hace parte la bibliografía sobre modelos abstractos). Y es que es usando las relaciones de derrota de un SAF, y no las de ataque, como se define el AAF asociado. Una vez que tenemos el AAF asociado, podemos aplicar las semánticas de Dung para calcular el conjunto de argumentos aceptables de un SAF y, en última instancia, el conjunto de fórmulas (proposiciones) que resulta viable aceptar en un SAF dado.

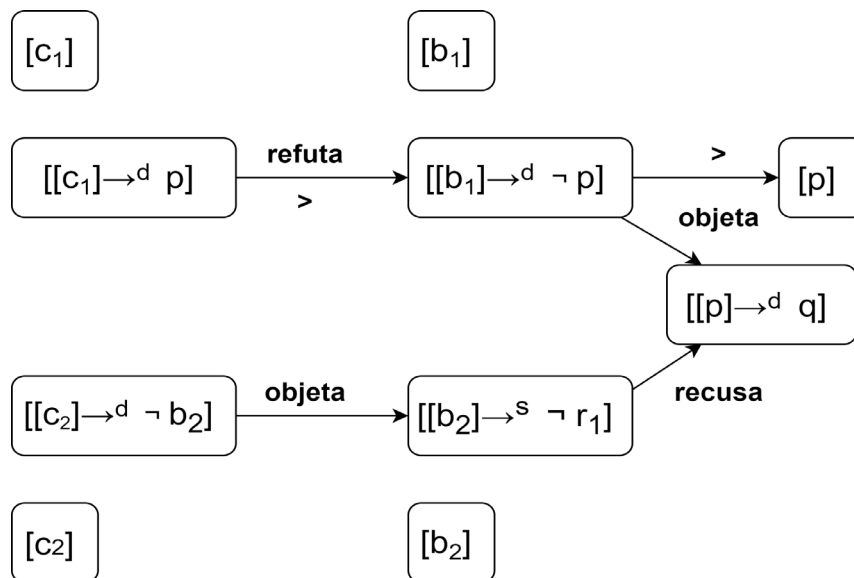


Figura 3: SAF para el Ejemplo 0

Ejemplo 3. El SAF representado en la Figura 3 captura algunos elementos del Ejemplo 0, abstrayéndonos de la incertidumbre y asumiendo cierta estructura en los argumentos modelados (omitimos detalles). Nótese que asumimos que $Nom(\langle p, q \rangle) = r_1$, que

$\langle [c_1] \rightarrow^d p, [b_1] \rightarrow^d \neg p \rangle \in Pref$ pero $\langle [b_1] \rightarrow^d \neg p, [c_1] \rightarrow^d p \rangle \notin Pref$ y que $\langle [b_1] \rightarrow^d p, [p] \rangle \in Pref$ pero $\langle [p], [b_1] \rightarrow^d p \rangle \notin Pref$.

3. MODELOS ESTRUCTURADOS INCOMPLETOS DE ARGUMENTACIÓN

En esta sección discutimos sistemáticamente cómo se pueden instanciar los modelos abstractos con incertidumbre (IAAFs) usando ASPIC+. Analizaremos primero las dos sugerencias hechas por Baumeister et al. (2021), véase:

(H1) la incertidumbre de los argumentos se puede explicar atendiendo a la incertidumbre de las reglas de inferencia; y

(H2) la incertidumbre de las derrotas se puede explicar atendiendo a la incertidumbre de las preferencias entre argumentos.

Tras esto, analizamos y discutimos algunos resultados que ponen a prueba formalmente ambas hipótesis y, en base a los mismos, revelamos una asunción implícita detrás de los IAAFs. Cerramos la sección examinando otras formas de representar incertidumbre cualitativa en los modelos estructurados de argumentación.

3.1. REGLAS INCIERTAS

Un **modelo estructurado de argumentación con reglas incompletas** [*rule-Incomplete Structured Argumentation Framework*] (**rule-ISAF**) (Yuste-Ginel y Proietti, 2023) es una tupla $rule - ISAF = \langle Len, Con, Reg, Nom, Kno, Arg, Att, Pref \rangle$ que solo se diferencia de un SAF estándar (Sección 2.3) en que el conjunto de reglas Reg se asume dividido en cuatro subconjuntos disjuntos $Reg = Reg_s^F \cup Reg_s^? \cup Reg_d^F \cup Reg_d^?$ tales que:

- Reg_s^F representa el conjunto de reglas deductivas ciertas.
- $Reg_s^?$ representa el conjunto de reglas deductivas inciertas.
- Reg_d^F representa el conjunto de reglas derrotables ciertas.
- $Reg_d^?$ representa el conjunto de reglas derrotables inciertas.

Una **compleción de reglas** de $rule - ISAF$ es cualquier conjunto de reglas Reg' tal que:

$$(Reg_s^F \cup Reg_d^F) \subseteq Reg' \subseteq (Reg_s^F \cup Reg_s^? \cup Reg_d^F \cup Reg_d^?)$$

Nótese que a cada compleción de reglas Reg' le corresponde naturalmente un SAF, véase $SAF' = \langle Len, Con, Reg', Nom', Kno, Arg', Att', Pref' \rangle$, donde cada elemento

marcado con ' es simplemente el componente original restringido al conjunto de argumentos que puede construirse con Reg' .

Dado $rule - ISAF$, una **compleción de $rule - ISAF$** es cualquier modelo abstracto de argumentación (AAF) asociado a una compleción de regla de $rule - ISAF$.

Ejemplo 4. Si en el Ejemplo 3 asumimos que la regla $\langle c_2, \neg b_2 \rangle$ es incierta (i.e., $\langle c_2, \neg b_2 \rangle \in Reg_a^?$), entonces obtenemos un rule-ISAF con dos compleciones, véase, la de la Figura 3 y la de la Figura 4.

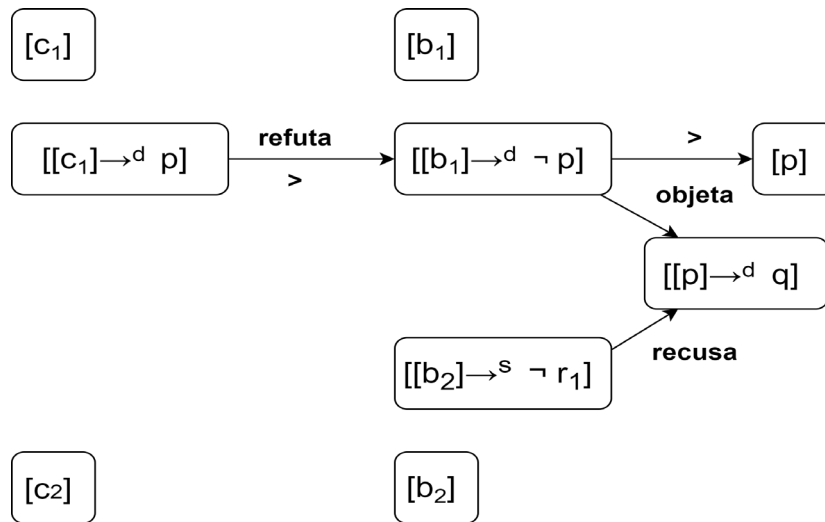


Figura 4: Segunda compleción para el Ejemplo 4

Nótese que las compleciones del ejemplo anterior se corresponden a la perfección con las compleciones de un IAAF, lo que habla en favor de (H1). Esto, sin embargo, no es el caso generalmente:

Proposición 1 (Yuste-Ginel y Proietti, 2023). Existe un rule-ISAF tal que su conjunto de compleciones no es igual al conjunto de compleciones de ningún IAAF.

3.2. PREFERENCIAS INCIERTAS

Un **modelo estructurado de argumentación con preferencias incompletas** [*preference-Incomplete Structured Argumentation Framework*] (**pref-ISAF**) es una tupla $pref - ISAF = \langle Len, Con, Reg, Nom, Kno, Arg, Att, Pref \rangle$ que solo se diferencia de un SAF estándar (Sección 2.3) en que la relación de preferencia entre argumentos $Pref$ se asume dividida en dos subrelaciones disjuntas $Pref = Pref^F \cup Pref^?$ tales que:

- $Pref^F$ representa los pares de argumentos cuya preferencia se sabe con certeza.

- $Pref^?$ representa los pares de argumento cuya preferencia es incierta.

Una **compleción de preferencia** de $pref - ISAF$ es cualquier $Pref'$ tal que:

$$Pref^F \subseteq Pref' \subseteq Pref^F \cup Pref^?$$

Nuevamente, a cada compleción de preferencia $Pref'$ le corresponde un SAF, véase $SAF' = \langle Len, Con, Reg, Nom, Kno, Arg, Att, Pref' \rangle$. Así, **una compleción** de $pref - ISAF$ es cualquier modelo abstracto (AAF) asociado a un SAF SAF' que a su vez está asociado a una compleción de preferencias $Pref'$ de $pref - ISAF$.

Ejemplo 5. Si, partiendo del Ejemplo 3, asumimos que $\langle [c_1] \rightarrow^d p, [b_1] \rightarrow^d \neg p \rangle \in Pref^?$, obtenemos dos compleciones del pref-ISAF resultante, véase, la de la Figura 3 y la de la Figura 5, donde la derrota entre $[c_1] \rightarrow^d p$ y $[b_1] \rightarrow^d \neg p$ se vuelve simétrica porque ningún argumento es estrictamente preferido al otro.

Nuevamente, las compleciones del ejemplo anterior se corresponden a la perfección con las compleciones de un IAAF, lo que habla en favor de (H2). Esto, sin embargo, no es el caso generalmente (omitimos la demostración por falta de espacio):

Proposición 2. Existe un pref-ISAF tal que su conjunto de compleciones no es igual al conjunto de compleciones de ningún IAAF.

Demostración. Considérese el pref-ISAF $pref - ISAF = \langle Len, Con, Reg, Nom, Kno, Arg, Att, Pref \rangle$ donde $Len = \{p, \neg p\}$, $Con(p) = \{\neg p\}$, $Con(\neg p) = \{p\}$, $Reg = Nom = Kno_a = Pref^F = \emptyset$, $Pref^? = \{\langle p, \neg p \rangle, \langle \neg p, p \rangle\}$. Obtenemos cuatro compleciones de reglas pero solo tres AAFs asociados, véanse, uno donde $[p]$ derrota a $[\neg p]$ (pero no viceversa), otro donde $[\neg p]$ derrota a $[p]$ (pero no viceversa) y otro donde ambos argumentos se derrotan mutuamente. Es fácil mostrar que estos tres AAFs no son igual al conjunto de compleciones de ningún IAAF. Por un lado, ninguna de las dos relaciones de derrota (de $[p]$ a $[\neg p]$ y viceversa) puede ser cierta (i.e., puede pertenecer a Der^F), porque aparecería en todas las compleciones pero sabemos que esto no ocurre. Por otro lado, tampoco puede ser el caso de que ambas relaciones sean inciertas (pertenecan a $Der^?$), porque de ser así habría una compleción donde ninguna de las dos derrotas se da.

3.3. DISCUSIÓN

Corrección de las hipótesis. Los resultados negativos de las secciones anteriores (proposiciones 1 y 2) muestran que las hipótesis (H1) y (H2), si se admite nuestra formalización de las mismas, no son del todo correctas: las reglas inciertas y las

preferencias inciertas no generan IAAFs a nivel abstracto, sino formalismos más expresivos con respecto al conjunto de compleciones que son capaces de generar. La caracterización de este formalismo para los rule-ISAFs puede encontrarse en (Yuste-Ginel y Proietti, 2023) (Teorema 1) y es un problema abierto para los pref-ISAFs.

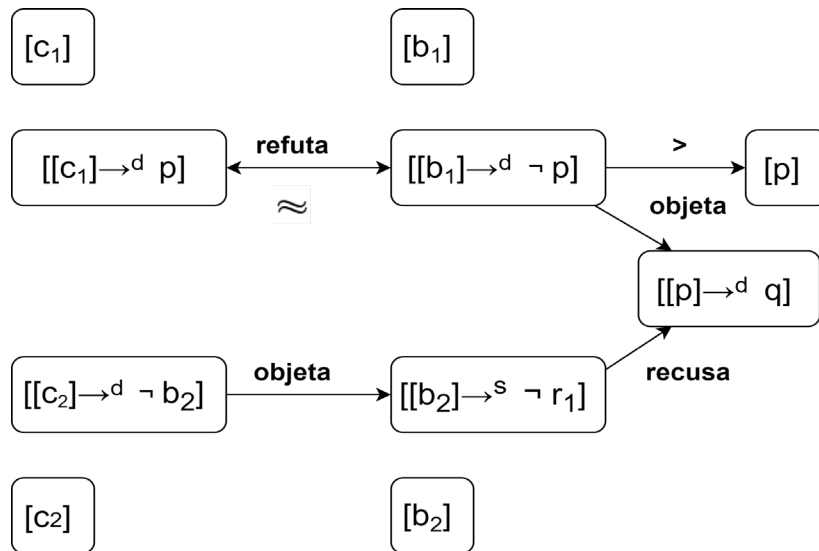


Figura 5: Segunda compleción para el pref-ISAF del Ejemplo 5

Una asunción implícita. En base a lo anterior, podemos enunciar una asunción implícita (y errónea) que subyace a los IAAFs. Para enunciarla, debemos hacer una breve distinción previa. Puede decirse que los IAAFs se valen de una manera *combinatoria* de representar incertidumbre. Más detalladamente, y centrándonos en los argumentos, lo que hace un IAAF es especificar dos conjuntos, Arg^F y $Arg^F \cup Arg^G$ y, mediante la noción de compleción, *se recorren todas las combinaciones* posibles que caigan entre ambos extremos. Esta no es, por supuesto, la única manera de representar incertidumbre cualitativamente con respecto a un AAF. Desde un punto de vista más general, basta con determinar un conjunto de AAFs directamente (Proietti y Yuste-Ginel, 2021), es decir, de determinar un conjunto de compleciones; o de ofrecer una restricción lógica que toda compleción ha de cumplir (Mailly, 2022; Yuste-Ginel y Herzig, 2023). Podemos entonces contraponer formas combinatorias de representar incertidumbre contra formas lógicas. Así, la asunción que subyace a los IAAFs y que hemos desentrañado en este trabajo puede enunciarse como:

(A1) La incertidumbre combinatoria a nivel de componentes se hereda como incertidumbre combinatoria a nivel de argumentos y derrotas.

Como hemos visto en las proposiciones 1 y 2, (A) es una asunción falsa.

3.4. MÁS OPCIONES PARA LA INSTANCIACIÓN

Pasamos, por último, a analizar otras opciones de instanciación de los IAAFs en ASPIC+, más allá de las recogidas por Baumeister et al. (2021).

Premisas inciertas. Una opción inmediata y extremadamente intuitiva es entender que los argumentos inciertos de un IAAF están enraizados en una base de conocimiento incierta. Desde una perspectiva epistémica, diríamos que un agente duda sobre si atribuir un argumento a su oponente porque duda sobre si su oponente acepta las premisas del argumento en cuestión. Curiosamente, esta opción ha aparecido recientemente en la bibliografía, en el contexto de los conocidos como problemas de relevancia y estabilidad (Odekerken et al., 2023). No se han estudiado, sin embargo, cuestiones de expresividad con respecto a compleciones como las aquí expuestas. Claramente, premisas inciertas provocan que los argumentos sean inciertos. Sin embargo, creemos que se pueden probar resultados negativos con respecto a una correspondencia estricta con los IAAFs también en este caso.

Lenguajes inciertos. Otra opción intuitivamente viable es separar el lenguaje formal de un modelo estructurado en una parte conocida Len^F y una parte incierta $Len^?$. En una lectura epistémica, Len^F sería la parte del lenguaje de la que el agente sabe que su oponente es consciente y $Len^?$ sería la parte del lenguaje tal que el agente duda de que su oponente sea consciente. Esto conectaría claramente con las conocidas como *lógicas de la conciencia* (Fagin y Halpern, 1987). Nuevamente, lenguajes inciertos provocan argumentos inciertos a nivel abstracto. Aquí también, conjeturamos que la correspondencia con los IAAFs no es estricta.

Aunque las opciones de dividir la función de contrariedad *Con* y la de nombramiento de reglas *Nom* en una parte conocida y otra incierta parecen menos atractivas desde un punto de vista intuitivo, son igualmente viables desde un punto de vista formal. El resto de componentes de un modelo estructurado (véase, *Arg* y *Att*) no son primitivos y no sirven, claramente, como opciones para instanciar los IAAFs usando SAFs.

4. CONCLUSIÓN

Este trabajo ha analizado las posibles maneras de instanciar los modelos abstractos incompletos de argumentación (IAAFs) en ASPIC+. Hemos motivado dicho estudio usando los argumentos de Prakken y De Winter (2018) sobre los peligros que conlleva

la proliferación indiscriminada de extensiones abstractas del modelo de Dung (1995). Hemos mostrado cómo las dos hipótesis de instanciación de los IAAFs en ASPIC+ propuestas en la bibliografía (Baumeister et al., 2021) no son del todo correctas, y que dichas instancias hablan a favor de modelos abstractos más expresivos en términos de incertidumbre. Por último, hemos explorado nuevas opciones para la definición de modelos estructurados incompletos de argumentación, conectándolas con otras partes de la bibliografía y conjeturando que esperamos obtener resultados negativos sobre la correspondencia de estos formalismos con los IAAFs a nivel abstracto.

Como es usual, este trabajo trae consigo más preguntas que respuestas. Aquí señalamos dos cuya resolución creemos viable y pronta. En primer lugar, una compleción natural del estudio sería ampliar los resultados de caracterización obtenidos para rule-ISAFs (Teorema 1 de Yuste-Ginel y Proietti (2023)) a pref-ISAFs y, en última instancia, a todas las posibles instancias de IAAFs introducidas en la sección 3.4 de este trabajo. En segundo lugar, creemos que existen conexiones interesantes entre las ideas aquí expuestas y los esfuerzos llevados a cabo para instanciar los modelos abstractos de argumentación bipolares en ASPIC+ (Cohen et al., 2018).

REFERENCIAS

- Baumeister, D., Järvisalo, M., Neugebauer, D., Niskanen, A., & Rothe, J. (2021). "Acceptance in incomplete argumentation frameworks". *Artificial Intelligence*, 295, 103470.
- Baroni, P., Gabbay, D., Giacomin, M., & Van der Torre, L. (2018). *Handbook of formal argumentation (volume 1)*. College Publications.
- Bench-Capon, T. J., & Dunne, P. E. (2007). "Argumentation in artificial intelligence." *Artificial intelligence*, 171(10-15), 619-641.
- Cohen, A., Parsons, S., Sklar, E. I., & McBurney, P. (2018). "A characterization of types of support between structured arguments and their relationship with support in abstract argumentation". *International Journal of Approximate Reasoning*, 94, 76-104.
- Dung, P. M. (1995). "On the acceptability of arguments and its fundamental role in nonmonotonic reasoning, logic programming and n-person games". *Artificial intelligence*, 77(2), 321-357.
- Gabbay, D., Giacomin, M., Simari, G. R. & Thimm, M. (2021). *Handbook of formal argumentation (volume 2)*. College Publications.
- Fagin, R., & Halpern, J. Y. (1987). "Belief, awareness, and limited reasoning". *Artificial intelligence*, 34(1), 39-76.
- Maily, J. G. (2022). "Yes, no, maybe, I don't know: Complexity and application of abstract argumentation with incomplete knowledge". *Argument & Computation*, 13(3), 291-324.
- Marraud, H. (2017). "De las siete maneras de contraargumentar". *Quadripartita ratio*, (4), 52-57.
- Modgil, S., & Prakken, H. (2013). "A general account of argumentation with preferences". *Artificial Intelligence*, 195, 361-397.
- Modgil, S., & Prakken, H. (2014). "The ASPIC+ framework for structured argumentation: a tutorial". *Argument & Computation*, 5(1), 31-62.
- Odekerken, D., Lehtonen, T., Borg, A., Wallner, J. P., & Järvisalo, M. (2023). "Argumentative reasoning in ASPIC+ under incomplete information". En: P. Marquis, T. Cao Son, G. Kern-Isberner (Eds.), *Proceedings of the International Conference on Principles of Knowledge Representation and Reasoning* (Vol. 19, No. 1, pp. 531-541). IJCAI Organization.
- Prakken, H. (2017). "Historical overview of formal argumentation". *IfCoLog Journal of Logics and their Applications*, 4(8), 2183-2262.

- Prakken, H., & De Winter, M. (2018). "Abstraction in Argumentation: Necessary but Dangerous". En: S. Modgil, K. Budzyska, J. Lawrence (Eds.), *Computational Models of Argument* (pp. 85-96). IOS Press.
- Yuste-Ginel, A., & Herzig, A. (2023). "Qualitative uncertainty and dynamics of argumentation through dynamic logic". *Journal of Logic and Computation*, 33(2), 370-405.
- Yuste-Ginel, A., & Proietti, C. (2023). "On the instantiation of argument-incomplete argumentation frameworks". En: G. Alfano y S. Ferilli (Eds.) *AI^3 2023 7th Workshop on Advances in Argumentation in Artificial Intelligence*. CEUR.

AGRADECIMIENTOS: He de agradecer a mi colega Carlo Proietti por su agudeza y generosidad intelectual en nuestras discusiones sobre el tema de este artículo. Mi gratitud está también con las personas asistentes al II Congreso Iberoamericano de Argumentación y a la discusión que siguió a la presentación de este trabajo.

A. YUSTE-GINEL: es profesor ayudante doctor en el Departamento de Lógica y Filosofía Teórica de la Universidad Complutense de Madrid. Antes de esto, se doctoró por la Universidad de Málaga con una tesis sobre la combinación de sistemas de lógica epistémica y argumentación formal (disponible [aquí](#)), bajo la dirección de Alfredo Burrieza. Después del doctorado, estuvo seis meses en el Institute de Recherche en Informatique de Toulouse trabajando como investigador postdoctoral. Sus intereses actuales giran en torno a la lógica modal, la argumentación formal y las aproximaciones simbólicas a la inteligencia artificial.