Research article

Argument & Computation
1–52

Article reuse guidelines:
sagepub.com/journals-permissions
DOI: 10.1177/19462174261466579
journals.sagepub.com/home/arg
Sage | IOS Press

Unveiling argumentative communities through ad-populum argumentation schemes

Paola D Budán[1,2,3], Maximiliano J Debórtoli[3], Maximiliano CD Budán[2,3], Gerardo I Simari[1] and Maria Vanina Martinez[4]

Abstract
This article introduces a method for detecting argumentative communities in semi-structured debates, which are supported by ad-populum argumentation schemes, as proposed by Walton. In these communities, opinions are characterized by their level of cohesion and the argumentative groups involved. In this context, we present a semi-structured argumentation formalism for analyzing social media debates and demonstrate how large language models can assist in instantiating the proposed formalism. We use conceptualizations, similarity and valued-based argumentation frameworks to assess argument cohesion and strength. Additionally, we model discursive interactions to analyze how stances influence one another and classify argumentative communities based on popular opinion argumentation schemes. We show that these communities can be modeled using such schemes, relying on public perception and consensus.

Keywords
argumentative communities, semi-structured argumentation, argumentation schemes, discourse analysis, large language models

Received: 4 October 2025; accepted: 23 June 2026

1 Introduction

Identifying argumentative communities in debates is a challenging yet essential task, as these communities play a crucial role in alliance formation and polarization in public opinion.[1,2] Furthermore, their detection can help uncover potential bridges between opposing groups, facilitating negotiation and dialogues. From a practical point of view, an argumentative community can be understood as a group of arguments or opinions that collectively support a particular statement, meaning that its members adopt supporting positions in a debate-like setting.[3] The internal cohesion of such a community could, consequently, be characterized by the similarity among its arguments, which reinforces shared perspectives.

A key challenge in identifying argumentative communities is understanding how the flow of information and persuasion unfolds in a debate. Through a systematic analysis of argumentative communities in structured debates, we can understand the discourse structures and how the argument interactions take place.[4] Additionally, classifying arguments into supporting, opposing, or neutral groups provides a clearer view of debate dynamics. Addressing the challenges associated with

[1]Departamento de Ciencias e Ingeniería de la Computación, Universidad Nacional del Sur (UNS) & Instituto de Ciencias e Ingeniería de la Computación (ICIC UNS-CONICET), Bahía Blanca, Argentina
[2]Consejo Nacional e Investigaciones Científicas y Técncias (CONICET), Buenos Aires, Argentina
[3]Departmento de Matemática, Universidad Nacional de Santiago del Estero, Santiago del Estero, Argentina
[4]Artificial Intelligence Research Institute (IIIA-CSIC), Bellatera, Catalonia, Spain

Corresponding author:
Maximiliano CD Budán, Consejo Nacional e Investigaciones Científicas y Técncias (CONICET), Godoy Cruz 2290 (C1425FQB), Buenos Aires, Argentina; Departmento de Matemática, Universidad Nacional de Santiago del Estero, Av. Belgrano (s) 1912 (G4200ABT), Santiago del Estero, Argentina.
Email: maximiliano699@gmail.com

argumentative communities is, therefore, fundamental to fostering more constructive debates, reducing polarization, and enhancing critical thinking in both human and artificial intelligence (AI)-driven discussions.

In this work, we introduce a method that combines both structural and content-based approaches[5] for detecting and classifying argumentative communities. The structural approach is based on labeled argumentation frameworks (LAF), which is a formalism to represent arguments and their characteristics through labels. Thus, the effects produced by the interactions between arguments, such as support, conflict, and aggregation, are reflected over these labels using operations defined in an algebraic structure. In this direction, we define two kinds of labels: one for modeling the support relation between terms in the debate's domain linked with a conceptualization, and the second for depicting relevant attributes associated with the source of the arguments. LAF can be regarded as a valuable tool in quantitative argumentation, where evidence plays a central role in supporting and explaining decision-making processes.[6] On the other hand, the content-based approach leverages Argumentation Schemes (AS) introduced by Walton,[7–9] which are semi-formal tools for structuring human opinions for computational analysis. Specifically, we focus on the *Ad-Populum AS*, or *argumentation from popular opinion*,[9,10] to address political debates and controversial topics in identifying discursive communities.

Although argumentation schemes abstract away many contextual details, they encode recurring patterns of persuasion that structure how influence propagates in a debate. In the case of ad-populum arguments, these patterns are inherently social, as they presuppose a reference group whose endorsement confers persuasive force. This makes ad-populum AS especially suitable for modeling not just isolated arguments, but networks of alignment and influence characteristic of argumentative communities. In this sense, AS act as a conceptual lens that makes explicit how individual argumentative moves contribute to the emergence of collective positions within debates.

In particular, ad-populum AS offer distinctive potential for addressing the gap between fine-grained argumentative analysis and the modeling of collective alignment in debates, as they explicitly encode references to social endorsement, collective alignment, and group-based authority. Rather than treating ad-populum as a monolithic pattern, our approach builds on Walton's typology of appeals to popular opinion, distinguishing between sub-schemes grounded in majority opinion, expert communities, social prestige, or shared identity. These sub-schemes provide principled criteria for associating arguments with specific reference groups, thereby enabling the identification of coherent argumentative communities and their internal structure. By explicitly operationalizing these sub-scheme distinctions within a formal labeled argumentation framework, we argue that ad-populum AS can serve not only as descriptive tools for individual arguments, but also as a viable modeling mechanism for capturing how group alignment and influence emerge and stabilize within debates. Particular challenges arise, however, given the informal language used to define AS, leading in some cases to ambiguity and concerns about both the fallacious nature and the adaptability of these schemes to real debate contexts. In this context, we note that the treatment of enthymemes,[11] as the implicit presence of premises or conclusions in AS, is outside the scope of this work. Importantly, our approach explicitly exploits Walton's typology of argumentum ad-populum not only to classify individual arguments, but also to characterize argumentative communities. In particular, qualitative labels associated with ad-populum sub-schemes are used to denote the type of community to which a proponent belongs, reflecting the reference group implicitly invoked in the argument. As discussed later, this perspective makes the group affiliation of proponents a central element for determining which ad-populum sub-scheme is instantiated in a debate.

Thus, our formalism integrates both quantitative and qualitative data to analyze debates according to the Ad-Populum AS. By converting the nuanced patterns of human argumentation into structured, analyzable formats, we can better understand how popular opinions shape discursive communities. Figure 1 provides an overview of our approach. This work aligns with previous studies[3,12–15] and complements them by demonstrating how a versatile formalism such as LAF, along with the extraction of argumentative elements from structured debates, aids in disambiguating AS. We clarify that the proposed framework should be understood as a normative argumentative model rather than as a descriptive or cognitive model of how humans actually determine winning or defeated communities. While the resulting classifications may align with human intuitions in certain cases, the framework defines one principled way, among many possible ones, of aggregating similarity, source quality, and group affiliation within an argumentative setting.

2 Motivational example and goals

Consider the scenario depicted in Figure 2—different opinions about dress codes generate conflicting positions, according to the structured debate from an online source (https://www.procon.org/headlines/dress-codes-top-3-pros-and-cons/). In this context, we observe the presence of two predominant positions:

- Pro stance: Highlights the benefits of dress codes, such as promoting a professional environment, improving work quality, and reinforcing safety standards in specific settings.

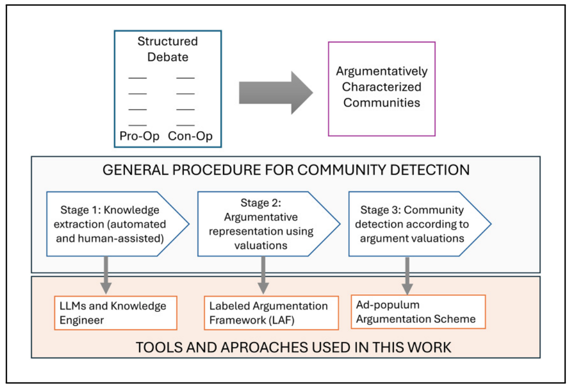

Figure 1. Proposed pipeline to detect argumentative communities.

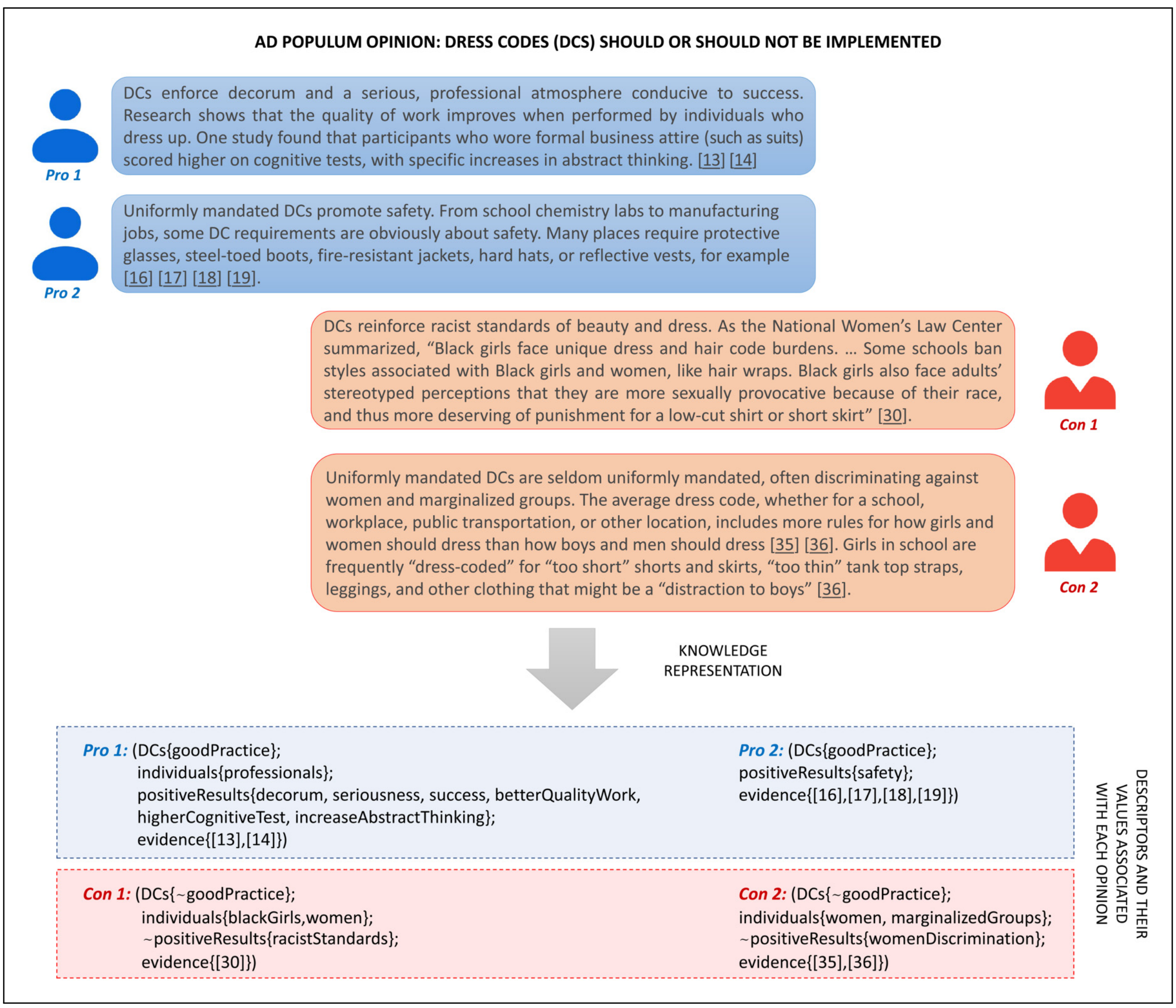

Figure 2. Opinions about the implementation of dress codes.

- Con stance: Emphasizes the negative aspects, including gender and racial discrimination that some dress code policies may impose, as well as unnecessary restrictions on individual freedom.

Determining which position is more valid is not trivial, especially when considering the credibility and reliability of the sources backing each argument. The complexity of this analysis increases due to the following aspects:

- Diverse evidence: Opinions are supported by different types of evidence, including scientific studies, institutional policies, and personal experiences.
- Multiple perspectives: A single topic (dress code implementation) can be analyzed from various angles, affecting different groups (professionals, students, women, minorities, etc.). Also, different opinions instantiate different subsets of domain concepts (for instance, the absence of individuals in Pro 2 reflects the fact that the original argument's formulation does not mention any individuals).
- Multiple claims within a single stance: A single argument may include various claims, reinforcing its validity with different types of evidence or perspectives.

To address this complexity, we use argument descriptors that formally represent the key elements of each argument. An argument descriptor is a label that identifies a relevant aspect of the discussion. For example, the descriptor "individuals" may take the value "professionals" in a pro stance (Opinion-Pro1), while in a con stance (Opinion-Con1), the same descriptor may take values such as "women" or "Black girls." This structured representation is based on the approach introduced by Budán et al.,[13] allowing us to capture the semantic complexity of debates. Despite this structuring, one key challenge in debate analysis is the identification of discourse communities, that is, groups of opinions with shared characteristics. In this case, it becomes apparent that Opinion-Pro 1, which supports the Pro Claim, assigns different values to the concept of "results" compared to Opinion-Pro 2, despite the fact that both opinions support the same claim. Finally, making a high-level analysis, we clearly distinguish two communities that hold contrasting opinions regarding dress code, with one community emphasizing the positive outcomes and the other expressing skepticism or disagreement.

Our proposal addresses this issue by modeling these communities not only based on argument similarity but also incorporating the credibility and features of the sources supporting each stance. Thus, we propose a methodological framework that enables:

- The detection and representation of discourse communities in structured debates, considering the internal cohesion of arguments within each community.
- The evaluation of controversy levels between communities with opposing stances, using metrics that combine semantic similarity and source credibility.
- The classification of arguments based on their alignment with popular opinion, facilitating the identification of recurring discourse patterns.
- The interpretation of how different discourse communities shape public perception on various issues.

These elements make our approach applicable to a wide range of contexts, from analyzing debates on social media to evaluating controversies in political and decision-making processes. In particular, we complement our theoretical contributions by developing a practical example in the context of education. Furthermore, unlike previous approaches that focus solely on argument similarity, our proposal introduces a hybrid model that incorporates both debate structure and source credibility.

This article is structured as follows: Section 3 provides a brief introduction to ad-populum AS and LAF, which contain the essential elements for our development. In Section 4, we identify and characterize opinion exchanges in a structured debate by mapping argumentative discourse onto a conceptual framework, which serves as the basis for constructing a contextualized semantic graph that models the debate's structure. The core contribution of the article is presented in Section 5, to 7, where the ad-populum LAF is introduced. Through this framework, the debate is modeled by integrating domain-specific knowledge and characteristics into a unified system that enables the detection and characterization of communities. Section 8 presents an illustrative complete application, where the instantiated formalism is applied to a specific domain. Section 9 discusses related work, and finally, Section 10 concludes the article and outlines directions for future research.

3 Conceptual background

In this section, we provide a summary of the ad-populum AS, which we will use to characterize opinions. In addition, we offer a description of the labeled argumentation framework, which serves as a key tool for formalizing the concepts required to express the structure of these schemes in a precise and computable way.

3.1 Ad-populum AS

As we mentioned previously, the *Ad-Populum AS* provided by Walton[7–9] are useful tool for analyzing opinion exchange in debates. These schemes consider an argumentative structure as a set of arguments with particular features (such as the proponent's position to know, their expertise, or moral position, etc.), recognizing provisional and defeasible arguments.[16] However, these schemes are defined in a non-computable and sometimes ambiguous manner, using colloquial language and the following general structure:

- *Argument in favor of A (accept A):*
  - Everybody (in a particular reference group g) accepts A.
  - Therefore, A is true (or you should accept A).
- *Argument against A (reject A):*
  - Everybody (in a particular reference group g) rejects A.
  - Therefore, A is false (or you should reject A).

As we can note in the AS, competing claims can coexist, but when a group of proponents opposes a claim, its validity is brought into doubt. Here, the strength of the argument relies on the collection of commitments and values embraced by the specific group g to which the proponent belongs. It can be attacked on the premises, on the weakness of an argument, and the presence of evidence that contradicts its premises.[16] Thus, each AS is accompanied by critical questions (CQs), which are specific queries designed to test the strength and validity of an argument by identifying potential weaknesses, assumptions, or gaps in reasoning within each type of argumentative structure. CQ for an ad-populum AS are:

- CQ1: Does a large majority of the cited reference group accept A as true?
- CQ2: Is there other relevant evidence available that would support the assumption that A is not true?
- CQ3: What reason is there for thinking that the view of this large majority is likely to be right?

The first CQ verifies if the claim is supported by a majority. If this is not the case, the argument weakens. The second seeks to introduce counter-evidence as a way to challenge the argument: credible evidence that contradicts the majority's belief undermines the argument's persuasiveness. The last CQ asks whether their belief is based on informed reasoning or relevant experience, rather than mere popularity. Usually, the pattern proposed by Walton, hereafter referred to as the *original* pattern, is reformulated as follows:

- If a large majority in a reference group g accepts A as true (false), then there exists a defeasible presumption in favor of (against, respectively) A.
- A large majority accepts A as true (false).
- Therefore, there exists a presumption in favor of (against, respectively) A.

AS are powerful tools because they help structure human reasoning in specific situations. In particular, AS capture content-related components of premises and conclusions, which enable the identification of recurring reasoning patterns and their organization into broader argumentative clusters. One such cluster is *argumentum ad-populum*, which encompasses a family of persuasion techniques systematically characterized through Walton's typology of ad populum sub-schemes. However, their subjective interpretation does not adequately support the use of computational tools. Therefore, we will employ technical formalisms to help translate these reasoning patterns into algorithmic structures. In this direction, the main challenge in applying this pattern lies in precisely defining the concepts of *large majority*, *accepts A*, and *a particular reference group*. Specifically, what constitutes a "large majority" can vary depending on the context, and without a clear definition, the interpretation can be subjective.[17] Similarly, determining what it means for a group to "accept A" involves understanding the level of agreement or endorsement required, which can also be ambiguous. Lastly, identifying "a particular reference group" requires specifying the criteria for group membership and relevance to the argument, which can also be open to interpretation. These ambiguities make it difficult to apply the pattern consistently and rigorously in computational models.

In this context, Walton[9] proposed a series of sub-schemes to address these issues. Specifically, these sub-schemes have the following features:

- *Position-to-Know Ad-Populum Argument (PtoK)*: The argument is provided by a group of people that holds the appropriate position to support the claims.
- *Expert Opinion Ad-Populum Argument (EX)*: The argument is provided by a group of experts in a knowledge domain.
- *Moral Justification Ad-Populum Argument (MO)*: This strategy relies on the popularity of a particular moral viewpoint to lend credibility or legitimacy to the arguer's stance.

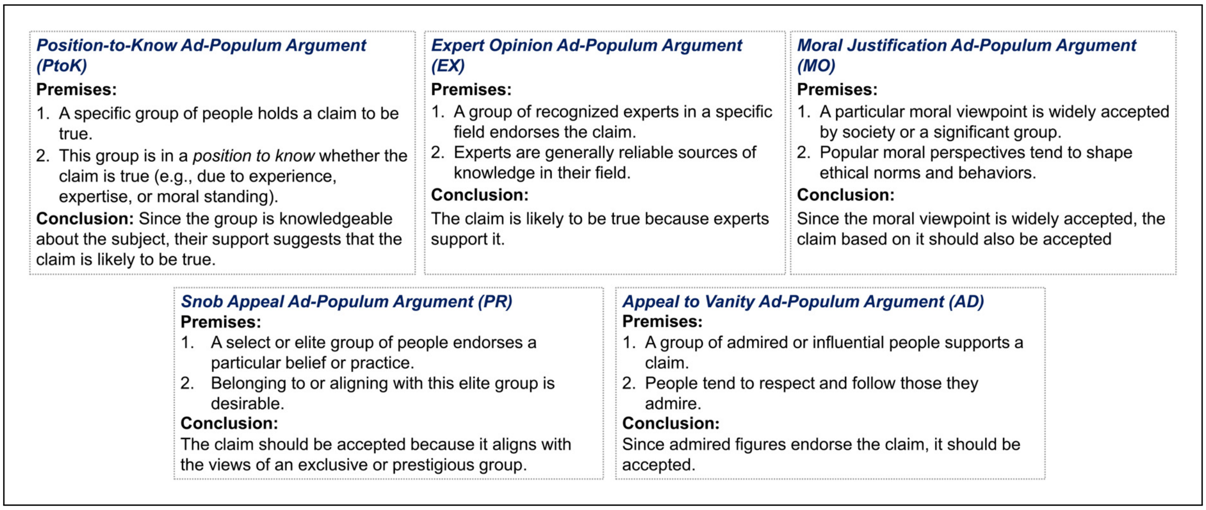

Figure 3. Ad-populum sub-scheme patterns.

- *Snob Appeal Ad-Populum Argument (PR)*: An arguer appeals to the beliefs or practices of a select or elite group of people to support their claim or position. The arguer uses the allure of exclusivity or prestige to persuade others to accept their point of view.
- *Appeal to Vanity Ad-Populum Argument (AD)*: The argument is provided by a group of admired people.

The sub-schemes are selected considering that their general structures are similar, as can be seen in Figure 3, where the premises and conclusion of each sub-scheme are considered, and the features of the proponent group highlight the distinctive support given to the argument. This similarity in structure ensures that the sub-schemes can be systematically applied across different contexts, facilitating the identification of common patterns and the assessment of argument quality.

Additionally, by maintaining a consistent structural framework, integrating these sub-schemes into computational models becomes easier, enhancing their applicability and reliability across various analytical scenarios. To achieve this, we will propose elements to reduce the ambiguity of the general pattern and enhance the precision of argument analysis. Furthermore, we will use the groups that characterize each sub-scheme to address the complexities inherent to debates.

3.2 Argument interchange format

The *Argument Interchange Format* (AIF) is an ontology created to capture, in an abstract way, the specification of argumentative information that describes a discourse domain and the relations among these pieces of information, such as inference, conflict and preference. In AIF, arguments and their mutual relations are represented through a *Typed Directed Graph*, which is an intuitive way to represent arguments in a structured and systematic way without the formal constraints of a logic.[18] The AIF core ontology is naturally defined in two parts: the *Upper Ontology* and the *Forms Ontology* (Figure 4).

The AIF is a standardized framework designed to facilitate interoperability among argumentation systems, which provides a standardized ontology for structuring and exchanging argumentation data. It consists of Information Nodes (I-nodes) and Scheme Nodes (S-nodes). I-nodes represent fundamental pieces of information such as premises, conclusions, or data sources, while S-nodes capture patterns of reasoning and are further classified into three categories: Rule of Inference Nodes (RA-nodes), which represent logical inference mechanisms; Conflict Nodes (CA-nodes), which model contradictions or attacks between arguments; and Preference Nodes (PA-nodes), which establish priority or ranking among competing arguments.

In the ontology, arguments and the relations between them are conceived of as an argument graph. The Upper Ontology defines the basic building blocks of AIF argument graphs, nodes, and edges. The Forms Ontology allows us to type the elements of the upper ontology in terms of argumentation-theoretical concepts. In AIF, there exist *information nodes* (I-nodes) and *scheme application nodes* (S-nodes), where I-nodes are used to represent propositional information contained in an argument (claim, premise, data, etc.) and S-nodes capture the application of schemes, such as inference rule schemes (RA-nodes), conflict schemes (CA-nodes) and preference schemes (PA-nodes). Figure 4 presents the full AIF framework. S-nodes can be further classified; for example, inference schemes can be deductive or defeasible, and

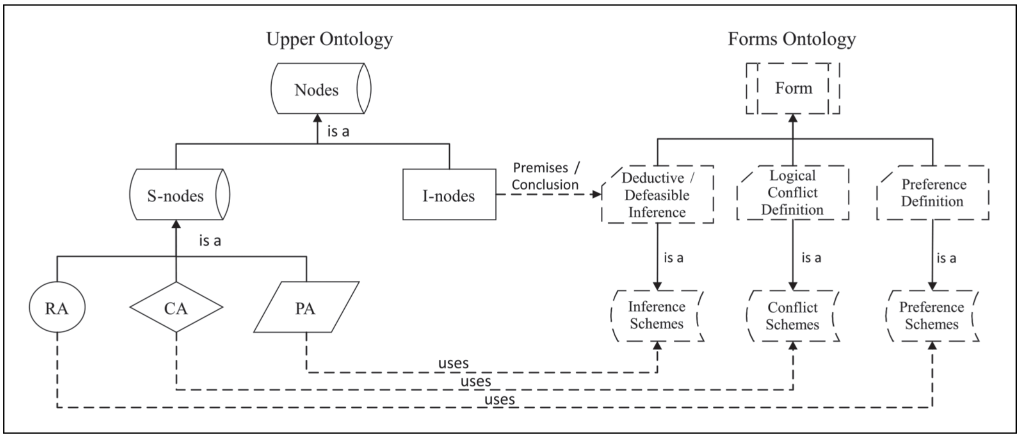

Figure 4. Argument interchange format (AIF) core ontology.

defeasible inference schemes can be subdivided into more specific AS (e.g., Expert Opinion or Witness Testimony, etc.). Thus, inference schemes, conflict schemes, and preference schemes in the Forms Ontology embody the general principles expressing how it is that q is inferable from p, p is in conflict with q, and p is preferable to q, respectively. The individual RA-nodes, CA-nodes, and PA-nodes that fulfill these schemes then capture the passage or the process of actually inferring q from p, conflicting p with q and preferring p to q, respectively.

The collection of nodes presented before can be used to build a typed directed graph, called *Argumentation Graph*, that represent the knowledge that describe a specific domain. Formally:

Definition 1 Argumentation Graph**.** An AIF argumentation graph is a digraph $G = (V, E)$, where:

- $V = I \cup RA \cup CA \cup PA$ is the set of nodes in G, where I are the I-nodes, RA are the RA-nodes, CA are the CA-nodes, and PA are the PA-nodes;
- $E \subseteq V \times V \setminus I \times I$ is the set of edges in G. Any edge $e \in E$ is assumed to have exactly one type among the following: premise, conclusion, preferred element, dispreferred element, conflicting element, and conflicted element;
- if $v \in RA$ then v has at least one direct predecessor via a premise edge and exactly one direct successor via a conclusion edge;
- if $v \in PA$ then v has exactly one direct predecessor via a preferred element edge and exactly one direct successor via a dispreferred element edge;
- if $v \in CA$ then v has exactly one direct predecessor via a conflicting element edge and exactly one direct successor via a conflicted element edge.

Technically speaking, there are two types of edges: *scheme edges* emanate from S-nodes and connected with either I-nodes or S-nodes, and *data edges* emanating from I-nodes but necessarily ending in S-nodes. Notice that edges connecting I-nodes are forbidden, because I-nodes cannot be connected without an explanation that justifies that connection. There is always a scheme, justification, inference, or rationale behind a relation between two or more I-nodes that is captured through an S-node. Moreover, only I-nodes can have zero incoming edges, as all S-nodes relate two or more components (for RA-nodes, at least one antecedent is used to support at least one conclusion; for PA-nodes, at least one alternative is preferred to at least one other; and for CA-nodes, at least one claim is in conflict with at least one other). In this work, we will concentrate only on the relationships that are relevant to our formalization, which are: I-to-RA and RA-to-I representing the premises and the conclusion of an inference scheme application, respectively. That is, we can model, for example, a legal rule; I-to-CA and CA-to-I representing the relation between conflicting or contradictory premises or claims. For example, to model conflicting interpretations of provisions or contradictory conclusions provided by different legal arguments from different legal systems. The AIF ontology also includes I-to-PA and PA-to-I relations, which represent preference relations between premises or claims. These relations are useful to establish preferred options in a decision-making process, for instance, when two non-contradictory conclusions represent alternative legal positions and one is considered more persuasive than the other. However, although PA-nodes are part of the standard AIF ontology and are mentioned here for completeness, they are not used in the present instantiation of our framework. Conflict resolution is handled exclusively through the conflict operator of the algebra (PA-nodes are included in the AIF ontology to represent preference relations. In the current framework, preference relations are not employed, and conflict resolution is performed exclusively through the conflict operator of the algebra.).

3.3 Algebra of argumentation labels

Changes in the perception of real-world situations may alter the reliability of the information that agents need to represent. As a consequence, propositions about a given domain often arise under conditions of uncertainty or partial knowledge, where both the state of the world and the available evidence are imprecise. In this setting, an argumentative formalism should be able to capture not only the logical structure of arguments, but also additional domain-dependent features that connect arguments to their contextual interpretation. In particular, the motivational example discussed in this work highlights that such features may be heterogeneous in nature. Some of them can be naturally modeled using quantitative information, such as degrees of credibility or relevance associated with real-world evidence, while others are inherently qualitative, including the origin of information sources or the values promoted by an argument. Accounting for this diversity is essential when analyzing complex debates involving multiple perspectives and sources, such as public discussions on dress code policies.

To address this issue, we introduce the use of *labels* as a mechanism for characterizing and evaluating arguments beyond their logical content. Labels encode both quantitative and qualitative, domain-dependent information associated with arguments, and they evolve as arguments interact with one another. We formalize this mechanism through the notion of an *Algebra of Argumentation Labels*, an abstract algebraic structure that defines how labels are combined, propagated, and transformed during the argumentative process. In particular, the operations of support, aggregation, and conflict capture the effects of argumentative interactions on labels. The algebra is defined over an ordered set, enabling the comparison of labels and providing a principled basis for assessing the relative strength and impact of arguments. Formally:

Definition 2 Algebra of Argumentation Labels**.** An algebra of argumentation labels A is defined as $\mathsf{A} = \langle A, \leq, \odot, \oplus, \ominus, \bot, \top \rangle$, where:

- A is a set of labels called the *domain of labels*.
- $\leq$ is a partial order over A (that is, a reflexive, antisymmetric, and transitive relation).
- $\top$ and $\bot$ are two distinguished elements of A. $\top$ is the last label with respect to $\leq$, while $\bot$ is the first.
- $\odot : A \times A \to A$ is called a *support operation* and satisfies:
  $\odot$ is a *commutative* operation: for all $\alpha, \beta \in A$, $\alpha \odot \beta = \beta \odot \alpha$.
  $\odot$ is *monotone*: for all $\alpha, \beta, \gamma \in A$, if $\alpha \leq \beta$, then $\alpha \odot \gamma \leq \beta \odot \gamma$.
  $\odot$ is *associative*: for all $\alpha, \beta, \gamma \in A$, $\alpha \odot (\beta \odot \gamma) = (\alpha \odot \beta) \odot \gamma$.
  $\top$ is the *neutral element* for $\odot$: for all $\alpha \in A$, $\alpha \odot \top = \alpha$.
- $\oplus : A \times A \to A$ is called an *aggregation operation* and satisfies:
  $\oplus$ is a *commutative* operation: for all $\alpha, \beta \in A$, $\alpha \oplus \beta = \beta \oplus \alpha$.
  $\oplus$ is *monotone*: for all $\alpha, \beta, \gamma \in A$, if $\alpha \leq \beta$, then $\alpha \oplus \gamma \leq \beta \oplus \gamma$.
  $\oplus$ is *associative*: for all $\alpha, \beta, \gamma \in A$, $\alpha \oplus (\beta \oplus \gamma) = (\alpha \oplus \beta) \oplus \gamma$.
  $\bot$ is the neutral element for $\oplus$: for all $\alpha \in A$, $\alpha \oplus \bot = \alpha$.
- $\ominus : A \times A \to A$ is called a *conflict operation* and satisfies:
  For all $\alpha, \beta \in A$, if $\beta < \alpha$, then $\alpha \ominus \beta \leq \alpha$.
  For all $\alpha, \beta \in A$, if $\beta \geq \alpha$, then $\alpha \ominus \beta = \bot$.
  $\ominus$ is *monotone in its first argument*: for all $\alpha, \beta, \gamma \in A$, if $\alpha \leq \beta$, then $\alpha \ominus \gamma \leq \beta \ominus \gamma$.
  $\ominus$ is *anti-monotone in its second argument*: for all $\alpha, \beta, \gamma \in A$, if $\beta \leq \gamma$, then $\alpha \ominus \gamma \leq \alpha \ominus \beta$.
  $\bot$ is the neutral element for $\ominus$: for all $\alpha \in A$, $\alpha \ominus \bot = \alpha$.
  For all $\alpha, \beta \in A$, if $\alpha \ominus \beta = \bot$ and $\beta \ominus \alpha = \bot$, then $\alpha = \beta$.
  For all $\alpha, \beta \in A$, if $(\alpha \oplus \beta) < \top$, then $((\alpha \oplus \beta) \ominus \beta) = \alpha$.

In this definition, we follow the standard algebraic convention of using the same symbol to denote both the algebra and its underlying carrier set. In particular, the first component of the tuple represents the domain of labels on which the algebraic operations are defined.

The support operation, denoted by $\odot$, is used to determine the valuation of a single argument based on the valuations of the arguments that directly support it. Intuitively, this operation captures how supporting reasons contribute to the strength of a conclusion. Since the order in which supporting arguments are considered should not affect the final valuation, the operation $\odot$ is required to be commutative and associative, with $\top$ as its neutral element. Moreover, if an argument is

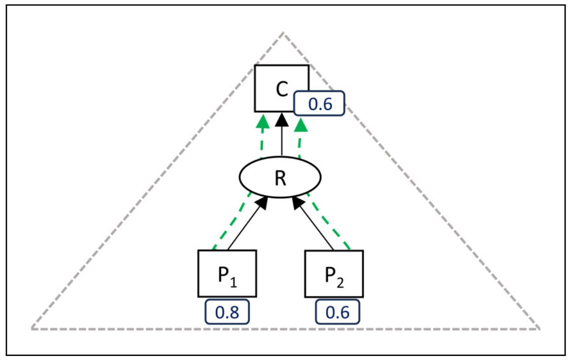

Figure 5. Support operation graphical example.

supported by stronger arguments, its resulting valuation should not decrease. This requirement is formalized through the monotonicity property, which is justified by the weakest-link principle commonly adopted in argumentation theory. These conditions allow us to characterize $\odot$ as a *t-norm*.[19,20]

Example 1 (Support Operation)**.** Consider the argument stating that enforcing a dress code improves professionalism. Assume that this argument is directly supported by two reasons: (i) the claim is semantically close to institutional regulations, with cohesion degree 0.8, and (ii) it is aligned with widely accepted workplace norms, with cohesion degree 0.6.

Using the support operation $\odot$ instantiated as the minimum t-norm, the resulting valuation of the argument is computed as follows:

$$0.8 \odot 0.6 = \min(0.8, 0.6) = 0.6.$$

This operation is illustrated in Figure 5, reflecting the weakest-link principle: the overall strength of the argument is bounded by its least convincing supporting reason.

The aggregation operation, denoted by $\oplus$, is used when several independent arguments share the same conclusion. In this case, the valuation of the accrued argument is obtained by combining the valuations of all supporting arguments. Conceptually, aggregation reflects the accumulation of independent sources of support for a claim. Accordingly, the operation $\oplus$ is required to be commutative and associative, with $\perp$ as its neutral element, ensuring that arguments with minimal valuation do not affect the accrued result. In addition, $\oplus$ satisfies monotonicity, guaranteeing that increasing the valuation of any supporting argument cannot decrease the valuation of the conclusion. These properties characterize $\oplus$ as a *t-conorm*. The use of t-conorms for aggregation is well established in several domains, including decision support and preference modeling.[19,20]

Example 2 (Aggregation Operation)**.** Assume that the conclusion "a dress code should be enforced" is supported by two independent arguments. The first argument is backed by a reliable expert, yielding a valuation of 0.7, while the second argument is supported by a respected professional community, yielding a valuation of 0.9.

Using the aggregation operation $\oplus$ instantiated as the maximum t-conorm $\max(\alpha, \beta)$, the accrued valuation of the conclusion is computed as follows:

$$0.7 \oplus 0.9 = \max(0.7, 0.9) = 0.9.$$

This operation is illustrated in Figure 6. This result reflects the accumulation of independent sources of support: the conclusion benefits from the strongest available argument.

The conflict operation, denoted by $\ominus$, determines the valuation of an argument after taking into account opposing arguments. This operation models how counterarguments weaken the strength of a claim. Its behavior is analogous to subtraction with respect to addition in numerical domains, although it is not defined as a strict inverse of aggregation. In particular, $\ominus$ satisfies anti-monotonicity with respect to the attacking argument: as the valuation of opposing arguments increases, the resulting valuation of the attacked argument decreases. Furthermore, $\perp$ acts as the neutral element for $\ominus$, indicating that counterarguments with minimal valuation do not affect the attacked claim. This design ensures a consistent and well-behaved mechanism for conflict resolution within the argumentation framework.[21,22]

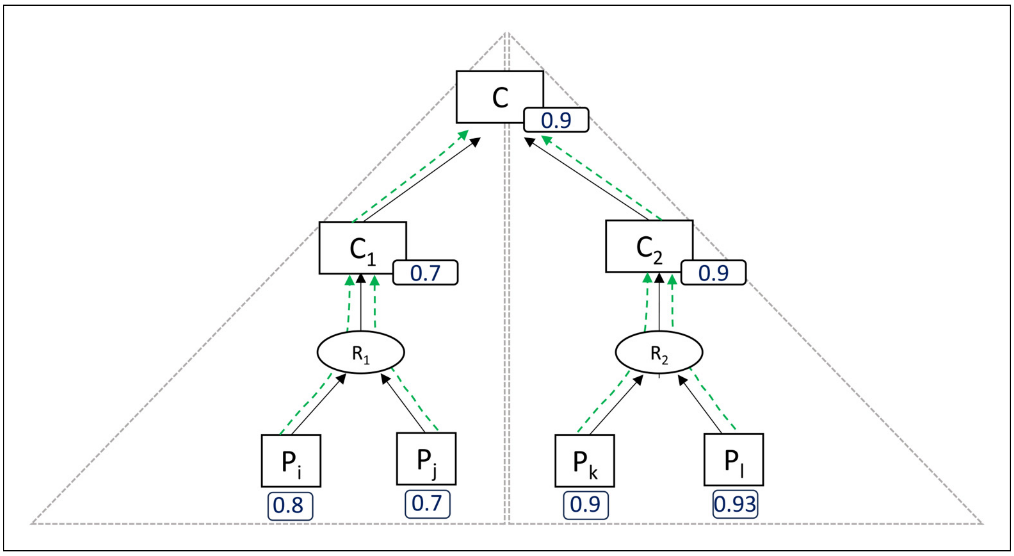

Figure 6. Aggregation operation graphical example.

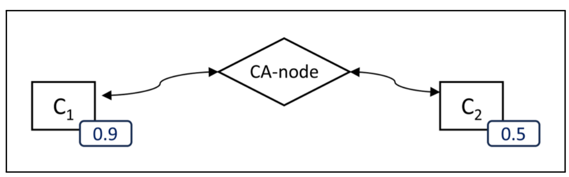

Figure 7. Example of an attack operation.

Example 3 (Conflict Operation)**.** Consider the previously accrued argument supporting the enforcement of a dress code with valuation 0.9. Suppose this argument is attacked by a counterargument claiming that dress codes limit personal expression, with valuation 0.5.

Using the conflict operation $\ominus$ instantiated as truncated subtraction, the final valuation of the attacked argument is computed as follows:

$$0.9 \ominus 0.5 = \max(0, 0.9 - 0.5) = 0.4.$$

This operation is illustrated in Figure 7. This result captures the weakening effect of counterarguments: the stronger the opposing reason, the more the original argument is reduced.

In the remainder of this work, the abstract algebra of argumentation labels is instantiated over different label domains in order to capture specific argumentative features arising in structured debates. In particular, we consider instantiations modeling semantic cohesion between arguments, the quality and credibility of the proponent, and the characteristics of the ad-populum discourse communities associated with each argument. Each instantiation preserves the algebraic properties introduced above, while providing a domain-specific interpretation of the labeling operators that enables the identification, comparison, and analysis of discourse communities.

3.4 Labeled argumentation frameworks

We now provide a concise overview of the formalism known as *Labeled Argumentation Frameworks* (*LAF*),[22,23] which enables the representation of arguments together with their internal structure, interactions, and relevant meta-level features through the use of argumentation labels. LAF extends abstract argumentation by associating each argument with one or more labels representing quantifiable attributes, such as strength, credibility, trust, or relevance, typically valued over a fuzzy or partially ordered domain. These labels propagate through an argumentative graph according to the relations between arguments—namely support, aggregation, and conflict—using the operators defined in an algebra of argumentation labels. This mechanism enables a refined and fine-grained assessment of argument acceptability beyond purely structural considerations.

In particular, LAF represents the internal structure of arguments and their interrelations by means of the *Argument Interchange Format* (AIF),[18] which provides a standardized and well-established mechanism for modeling claims, inference steps, and attack relations. The integration of AIF with an algebraic treatment of argumentation labels results in a flexible and expressive framework capable of handling complex argumentative scenarios, including cyclic argumentation graphs, through a system of equations that guarantees a consistent evaluation of arguments.

Building upon these foundations, LAF defines a labeling process that computes the final label associated with each argument by propagating and transforming labels along the argumentation graph. The resulting labels are then used to determine the acceptance status of arguments in the presence of conflicts and inconsistencies. In this work, we revisit and refine the original labeling procedure proposed by Budán et al.[22] in order to achieve improved behavior, introduce a set of postulates that the revised formalism satisfies, and analyze several properties of interest. Whenever possible, we preserve the original notation of LAF[22] to maintain consistency with prior work and to facilitate comparison with existing results.

Definition 3 (Labeled Argumentation Framework)**.** A LAF is a five-tuple of the form $\Phi = \langle \mathcal{L}, \mathcal{R}, \mathcal{K}, \mathcal{A}, \mathcal{F} \rangle$, where:

- $\mathcal{L}$ is a logical language for knowledge representation (claims) about the domain of discourse. The connectives of this language include one distinguished symbol "$\sim$" denoting strong negation.
- $\mathcal{R}$ is a set of inference rules $R_1, R_2, \ldots, R_m$ defined in terms of $\mathcal{L}$ (*i.e.*, with premises and conclusion in $\mathcal{L}$).
- $\mathcal{K}$ is the knowledge base, a set of formulas of $\mathcal{L}$ describing knowledge about the domain of discourse.
- $\mathcal{A}$ is a set of algebras of argumentation labels $\mathsf{A}_1, \ldots, \mathsf{A}_n$, one for each feature to be represented by the labels.
- $\mathcal{F}$ is a function that assigns to each element of $\mathcal{K}$ an n-tuple of elements in the algebras A_i, *with* $i = 1, \ldots, n$. That is, $\mathcal{F} : \mathcal{K} \longrightarrow \mathsf{A}_1 \times \mathsf{A}_2 \times \cdots \times \mathsf{A}_n$.

We use $\mathcal{L}$ to represent the knowledge base, where inference is governed by a set of rules that capture various patterns of reasoning. These include deductive rules (e.g., *modus ponens*), defeasible rules (e.g., *defeasible modus ponens*), and argumentation schemes (e.g., *argument from expert opinion*), among others. Every formula of the form $\sim\sim \varphi$ in $\mathcal{L}$ is considered logically equivalent to φ. Therefore, we assume that no subformula of the form "$\sim\sim \varphi$" occurs explicitly in well-formed formulas, while ensuring that the language remains closed under the application of negation. To simplify the definition of conflict between claims, the use of multiple consecutive negations (i.e., two or more $\sim$ symbols in a row) is disallowed in the syntax of $\mathcal{L}$.

Example 4 (Motivational LAF: Dress-Code Debate)**.** The debate illustrated in Figure 2 can be modeled as a LAF $\Phi = \langle \mathcal{L}, \mathcal{R}, \mathcal{K}, \mathcal{A}, \mathcal{F} \rangle$, defined as follows:

- $\mathcal{L}$ is a language defined in terms of two disjoint sets: A set of presumptions and a set of defeasible rules, where a presumption is a ground atom X or a negated ground atom $\sim X$, where $\sim$ represents strong negation; a defeasible rule is an ordered pair, denoted $C \mathrel{-\!\!\prec} P_1, \ldots, P_n$, whose first component C is a ground atom, called the conclusion and the second component $P_1, \ldots, P_n$ is a finite non-empty set of ground atoms, called the premises.
- $\mathcal{R} = \{\, dMP\}$, denoting Defeasible Modus Ponens:
  dMP: $\frac{\mathtt{P}_1, \ldots, \mathtt{P}_n \quad \mathtt{C} \mathrel{-\!\!\prec} \mathtt{P}_1, \ldots, \mathtt{P}_n}{\mathtt{C}}$.
- $\mathcal{A} = \{\mathtt{B}\}$ denotes the set of algebras of argumentation labels. In this case, the set contains the single element $\mathtt{B}$, representing the trust degree associated with the source that puts forward an argument. The domain of labels $\mathtt{B}$ is the real interval $[0, 1]$. The support, conflict and aggregation operators over labels are specified as follows:

$\alpha \odot \beta = \alpha\beta$	The support operator models the trust assigned to a conclusion as the conjunction of the trust valuations of its supporting premises.
$\alpha \oplus \beta = \alpha + \beta - \alpha\beta$	The aggregation operator defines the trust of a conclusion with multiple supporting arguments as the sum of their trust values, adjusted by a penalty term.
$\alpha \ominus \beta = \begin{cases} 1 & \text{if } \alpha = 1,\ \beta < 1 \\ \dfrac{\alpha - \beta}{1 - \beta} & \text{if } \alpha \geq \beta,\ \beta \neq 1 \\ 0 & \text{otherwise} \end{cases}$	The conflict operator reduces the trust of a conclusion according to the trust assigned to its contrary.

- $\mathcal{K}$ the knowledge base $\mathcal{K}$ contains the following elements:

$$\mathcal{K} = \begin{Bmatrix} ProfessionalEnv(goodPractice) \\ WorkQuality(goodPractice) \\ Safety(goodPractice) \\ Discrimination(\sim goodPractice) \\ FreedomRestriction(\sim goodPractice) \end{Bmatrix} \cup \begin{Bmatrix} r_1 : DressCode(X) \multimap ProfessionalEnv, WorkQuality(X) \\ r_2 : DressCode(X) \multimap Safety(X) \\ r_3 : \sim DressCode(X) \multimap Discrimination(X) \\ r_4 : \sim DressCode(X) \multimap FreedomRestriction(X) \end{Bmatrix}.$$

- $\mathcal{F}$ is the function that assigns an initial credibility value to each basic claim:

$\mathcal{F}(ProfessionalEnv(goodPractice)) = 0.8$,
$\mathcal{F}(WorkQuality(goodPractice)) = 0.7$,
$\mathcal{F}(Safety(goodPractice)) = 0.6$,
$\mathcal{F}(Discrimination(\sim goodPractice)) = 0.75$,
$\mathcal{F}(FreedomRestriction(\sim goodPractice)) = 0.65$.

All rules are assigned valuation 1.

Budan et al.[22] present the notion of an argumentation graph, which is used to represent the argumentative analysis derived from a LAF. They assume that there are no two nodes in a given graph that are named with the same sentence of $\mathcal{L}$, so each I-node in the graph refers to a sentence that is not repeated in the graph.

Definition 4 (Argumentation Graph)**.** Given a LAF $\Phi = \langle \mathcal{L}, \mathcal{R}, \mathcal{K}, \mathcal{A}, \mathcal{F} \rangle$, the associated argumentation graph is the digraph $G_\Phi = (N, E)$, where $N \neq \emptyset$ is the set of nodes and E the set of edges, constructed as follows:

(i) Each element X that is either an element of $\mathcal{K}$ or derived from $\mathcal{K}$ through $\mathcal{R}$, is represented by an I-node $X \in N$; there must not exist two distinct I-nodes $X, Y \in N$ representing the same formula φ.
(ii) For each application of an inference rule defined in Φ, there exists an RA-node $R \in N$ such that:
  - The inputs are all I-nodes $P_1, \ldots, P_m \in N$ representing the premises necessary for the application of the rule R (including the defeasible rules of $\mathcal{K}$ that may be applied).
  - The output is an I-node $Q \in N$ representing the conclusion.
(iii) If X and $\sim X$ are in N, then there exists a CA-node with edges to and from both I-nodes X and $\sim X$.
(iv) For all $X \in N$, there does not exist a path from X to X in G_Φ that does not pass through a CA-node. That is, G_Φ without the CA-nodes is acyclic.

Condition (iv) forbids cycles other than the mutual attacks between nodes. Though this may seem to be too restrictive, it must be noted that RA-node cycles are mostly generated by fallacious specifications.[24,25]

Note that although the argumentative graph is constructed automatically by the LAF, this construction does not rely on if-then rules alone. In addition to defeasible rules, the framework requires a set of contextual facts, which provide the grounding necessary to determine which rule applications are instantiated and, consequently, which arguments become active in a given situation. Facts act as enabling conditions for RA-nodes, allowing premise-to-conclusion relations to be instantiated only when the corresponding contextual conditions hold.

Moreover, the context is assumed to be dynamic: facts may cease to hold, be revised, or change their valuation over time. Such changes directly affect the activation status of arguments and may alter the resulting argumentative structure without modifying the underlying set of rules. Therefore, the automatic construction of the argumentative graph depends on the joint availability of rules and context-sensitive facts, rather than on a static collection of if-then rules in isolation.

In Figure 8, we present the argumentation graph corresponding to the motivational example. The figure allows us to identify different subgraphs representing the arguments involved in the dialectical process. In particular, we identify the set of arguments $\Phi = \{dressCode(goodPractice), dressCode(\sim goodPractice)\}$. According to Definition 5, each argument is a subgraph of G_Φ. For instance, the argument for the statement $dressCode(goodPractice)$ (represented as an I-node in G_Φ) is the subgraph containing all the reasons supporting that statement (colored in green in the figure). Both arguments have conclusions supported by multiple independent sub-arguments that converge on the same claim. Following the terminology commonly used in argumentation theory,[22,26–28] we refer to these as *aggregated arguments*. This does not mean that an argument is itself an aggregation of arguments. Rather, each argument remains a subgraph of the argumentation graph,

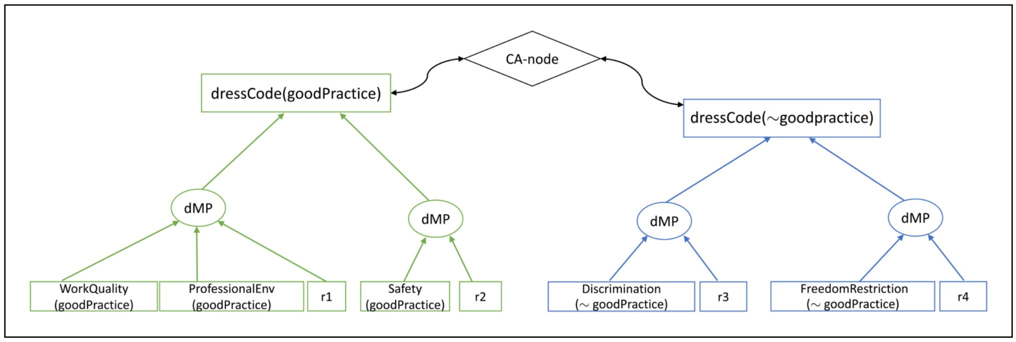

Figure 8. Argumentation graph for motivational example.

while the valuation of its conclusion is obtained by aggregating the valuations of its supporting sub-arguments. Hence, aggregation concerns the evaluation of the conclusion and not the definition of the argument itself. Formally:

Definition 5 (Argument)**.** Let $\Phi = \langle \mathcal{L}, \mathcal{R}, \mathcal{K}, \mathcal{A}, \mathcal{F} \rangle$ be a LAF, and G_Φ be the associated argumentation graph without CA-nodes for Φ. An argument for a statement represented through an I-node X is the subgraph $\mathbb{A} = (N', E')$ of G_Φ composed by all the ancestors of X, such that $N' \neq \emptyset$ and $E' \neq \emptyset$. We will denote with Arg_Φ the set of all arguments that can be identified in G_Φ.

By Definition 5, leaf I-nodes do not induce arguments. Indeed, since a leaf node has no ancestors in G_Φ, the corresponding subgraph contains no edges, and, therefore, the condition $E' \neq \emptyset$ is not satisfied.

Given the set of arguments identified in G_Φ, we now formally define the structural relations that determine how arguments interact within the LAF. In the AIF-based graph, interactions are induced by RA-nodes and CA-nodes. These give rise to support and conflict relations between arguments.

Definition 6 (Argument Interaction)**.** Let Arg_Φ be the set of all arguments that can be identified in G_Φ. Let $X, Y \in Arg_\Phi$ be two arguments. We say that:

- X supports Y iff there exists a directed path from X to Y containing at least one RA-node and no CA-node.
- X conflicts Y iff there exist a CA-node connecting X and Y in G_Φ.

In Figure 8, we identify the set of arguments $Arg_\Phi = \{dressCode(goodPractice), dressCode(\sim goodPractice)\}$, where both correspond to aggregated arguments, that is, arguments whose conclusions are supported by multiple independent sub-arguments, with their labels combined through the aggregation operator, and are connected by a CA-node representing a conflict between them. Neither the argument in favor of dressCode as a good practice nor its counterargument is atomic. Each is supported by two distinct sub-arguments, representing independent justificatory sources contributing to the respective conclusion. In the graph, aggregation is depicted by converging support relations toward both dressCode(goodPractice) and dressCode(goodPractice), respectively. This indicates that the strength of each claim results from the combination of its supporting subarguments rather than from a single premise. As such, the figure clarifies the internal structure of aggregated arguments and their role within the overall argumentation framework.

Once the argumentation graph is obtained, the author presents the process of labeling an argumentation graph, which involves creating a system of equations characterizing the formalism's knowledge. Intuitively, the labeling process distinguishes between two stages in the evaluation of a claim. The value μ_i^X represents the accrued valuation of a node X, obtained by aggregating all supporting reasons independently of conflicts. The value δ_i^X represents the weakened valuation of X after taking into account counterarguments and conflicts. This distinction, originally introduced in the LAF,[22] allows us to separate the contribution of supporting evidence from the effect of dialectical opposition.

Definition 7 (Labeling Procedure for a Graph)**.** A labeled argumentation graph is an assignment of two valuations in each of the algebras to all I-nodes of the graph, denoted with μ_i^X and δ_i^X, where μ_i^X accounts for the aggregation of the reasons supporting claim X, while δ_i^X displays the state of the claim after taking conflicts into account, such that $\mu_i^X, \delta_i^X \in A_i$. If X is an I-node, its valuations are determined as follows:

(i) If X has no inputs, then its accrued valuation is given by function $\mathcal{F}$; thus, we define $\mu_i^X = \mathcal{F}(X)$.

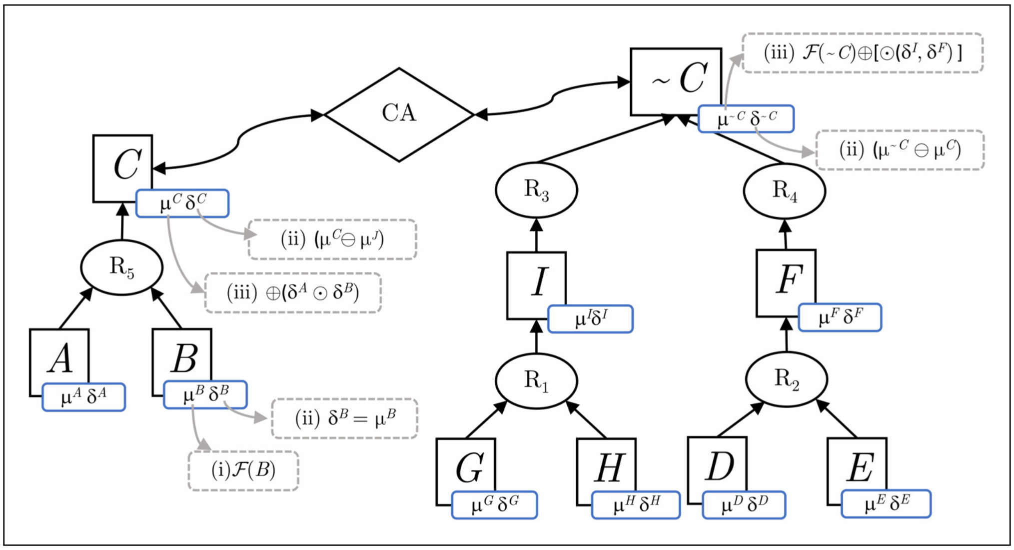

Figure 9. Example of labeling procedure for a graph.

(ii) If X has an input from a CA-node representing conflict with an I-node $\sim X$, then: $\delta_i^X = \mu_i^X \ominus \mu_i^{\sim X}$.
If there is no input from a CA-node, then: $\delta_i^X = \mu_i^X$.

(iii) If X is an element of $\mathcal{K}$ with inputs from RA-nodes $R_1, \ldots, R_k$, where each R_s has premises $X_1^{R_s}, \ldots, X_{n_s}^{R_s}$, then:

$$\mu_i^X = \mathcal{F}(X) \oplus \left[\oplus_{s=1}^{k} \left(\odot_{t=1}^{n_s} \delta_i^{X_t^{R_s}} \right) \right]$$

If X is not an element of $\mathcal{K}$ and has inputs from RA-nodes $R_1, \ldots, R_k$, where each R_s has premises $X_1^{R_s}, \ldots, X_{n_s}^{R_s}$, then:

$$\mu_i^X = \oplus_{s=1}^{k} \left(\odot_{t=1}^{n_s} \delta_i^{X_t^{R_s}} \right)$$

The label of an I-node X is then an n-tuple of pairs of valuations $\left((\mu_1^X, \delta_1^X), \ (\mu_2^X, \delta_2^X), \ldots, (\mu_i^X, \delta_i^X), \ldots, (\mu_n^X, \delta_n^X) \right)$, where each pair (μ_i^X, δ_i^X) represents the valuations of a certain feature associated with algebra $\mathsf{A}_i \in \mathcal{A}$.

Figure 9 illustrates the labeling procedure applied to an abstract argumentation graph and the dynamics induced by the interaction between support, aggregation, and conflict. Leaf I-nodes represent elementary pieces of knowledge provided by the knowledge base $\mathcal{K}$. These nodes receive their initial labels directly from the labeling function $\mathcal{F}$ and constitute the starting points of the labeling process. From these nodes, support relations propagate valuations upward through the graph via inference rules, giving rise to derived arguments represented by non-leaf I-nodes. These derived nodes may or may not correspond to elements originally contained in $\mathcal{K}$, depending on whether the corresponding claim is explicitly stated or only obtained through inference. For the purpose of illustrating the procedure shown in Figure 9, a tentative instantiation of the knowledge base can be considered, for example:

$$\mathcal{K} = \{A,\ B,\ D,\ E,\ G,\ H\}.$$

These I-nodes represent basic pieces of information initially available to the framework and, therefore, receive their initial valuations directly from $\mathcal{F}$. The remaining I-nodes in the graph (e.g., C, I, F, and $\sim C$) correspond to arguments derived through the application of inference rules. When multiple supporting paths converge on the same I-node, their associated valuations are aggregated, producing the accrued valuation μ, which reflects the cumulative strength of the supporting reasons. In parallel, the presence of conflicting arguments triggers weakening operations, resulting in the weakened valuation δ, which captures the impact of counterarguments on the overall acceptability of the node. The framework also allows for situations in which a claim belongs to the knowledge base and is simultaneously derivable through one or more inference rules. In such cases, the initial valuation assigned to the knowledge-base node is combined with the valuations obtained from derived supports, leading to a reinforced accrued valuation. Dashed boxes highlight intermediate stages of

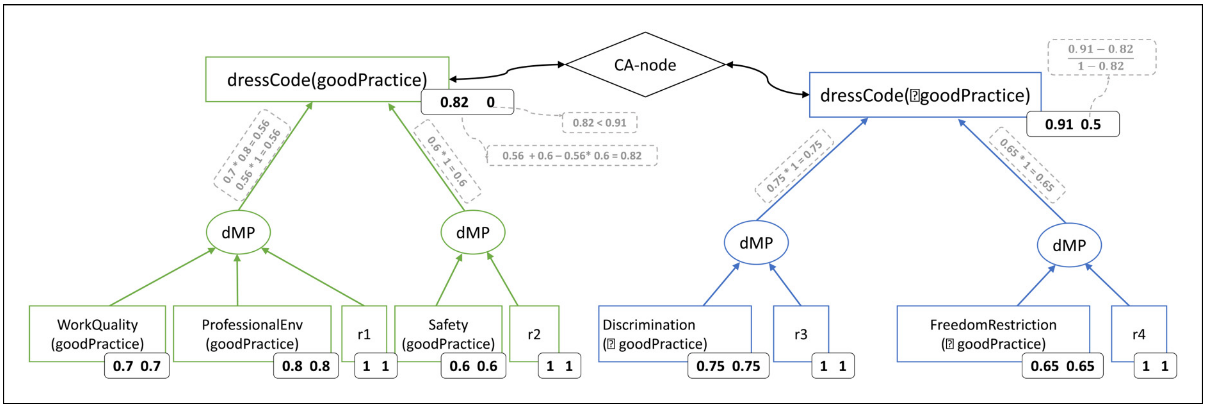

Figure 10. Labeling procedure for the argumentation graph of the motivational example.

the labeling process, making explicit how valuations are incrementally propagated, aggregated, and weakened across the graph. The numerical references shown in parentheses correspond to the specific clauses of Definition 7 that justifies each operation. Overall, the figure provides an intuitive yet precise visualization of how the labeling procedure operates over the argumentation graph, without committing to a fully specified instantiation of the underlying LAF.

Figure 10 illustrates the labeling procedure for the motivational example, according to the operators defined in Example 4. The dashed boxes represent the intermediate values resulting from the application of support, aggregation, and conflict operators. It is important to note how counterarguments weaken the strength of a claim. For example, the strength of the claim dressCode(goodPractice) is 0.82, whereas the strength of its counterargument is 0.91. Therefore, after applying the conflict operator, the resulting strength of dressCode(goodPractice) is 0.

After labeling the I-nodes of the argumentative graph, their acceptability can be assessed based on their final valuations. This additional information from argumentation labels allows us to define two processes: one that is dichotomous, and another that is gradual. The former follows the traditional binary model of acceptability, while the latter introduces a novel approach by modeling acceptability as a spectrum of values. Consequently, the gradual acceptability process refines the classical model.

Having established the conceptual foundation of ad-populum AS and LAF, we now turn our focus to how these formal structures can be effectively applied to model debates. While this section has provided the necessary theoretical background, the next step is to operationalize these concepts within a structured debate.

4 Modeling debates through conceptualizations

Our primary objective is to identify the underlying structure of a debate by modeling the relationships among its core concepts and characterizing those concepts through their associated attributes. To this end, we propose a method that maps the terms used in any argumentative discourse onto a formal conceptualization. This conceptualization organizes the terms or concepts invoked in the debate according to their semantic associations within the discussion. From this conceptual layer, we derive a contextualized semantic graph, that is, a semantic graph[29,30] that models a structured debate using exclusively the elements provided by the conceptualization.

We assume the existence of a language $\mathcal{L}'$ ($\mathcal{L}$ is the logical language for knowledge representation in LAF), which consists of a finite set of concepts C and a finite set of attributes Δ, which represents semantic values associated with descriptors of the concepts in C.

Definition 8. Let $\mathcal{L}' = \mathsf{C} \cup \Delta$ be a formal language. An argumentative debate can be represented as a conceptualization $\mathcal{O}_{\mathcal{L}'} = (\mathsf{C}, \Delta, map, \mathsf{C}^*, R_{ctx}, R_{ctxAtr})$, where:

- C is the set of domain concepts involved in the debate;
- Δ is the set of attributes that may describe those concepts;
- $map : \mathsf{C} \to 2^{\Delta}$ associates each concept with the set of attributes that characterize it;
- $\mathsf{C}^* = \{(c, a) \mid c \in \mathsf{C},\ a \in map(c)\}$ is the set of attribute-instantiated concepts.
- $R_{ctx} \subseteq \mathsf{C} \times \mathsf{C}$ is a semantic association relation between concepts;
- $R_{ctxAtr} \subseteq \mathsf{C}^* \times \mathsf{C}^*$ is an attribute-level argumentative relation between attribute-instantiated concepts.

Rather than treating concepts as purely abstract entities, the conceptualization represents the attributes under which they are invoked in the discourse. Each pair (c, a) represents an *attribute-instantiated concept*, that is, the concept $c \in \mathsf{C}$ considered under a specific attribute $a \in map(c)$. These instantiations correspond to particular interpretations or properties of the concepts as they occur in the discourse. The conceptualization, therefore, distinguishes two complementary structural levels: the relation R_{ctx} captures semantic association between domain concepts at an abstract level, representing how concepts are structurally connected within the debate, independently of the particular attributes under which they are considered. In contrast, R_{ctxAtr} operates at the level of attribute-instantiated concepts, capturing discourse relations, such as support or opposition, between particular contextualized instantiations of concepts. When clear from the context, we will use $\mathcal{O}$ instead of $\mathcal{O}_{\mathcal{L}'}$ to simplify notation.

A conceptualization can be represented through various means, such as ontologies,[31] UML diagrams,[32] or knowledge graphs.[33,34] To operationalize the conceptualization, we transform it into a contextualized semantic graph. This transformation preserves the instantiation structure defined by *map* and represents explicitly the discursive relations between attribute-instantiated concepts.

Definition 9 (Contextualized Semantic Graph)**.** Let $\mathcal{O}_{\mathcal{L}'} = (C, \Delta, map, C^*, R_{ctx}, R_{ctxAtr})$ be a conceptualization. The contextualized semantic graph derived from $\mathcal{O}_{\mathcal{L}'}$ is the directed graph $\mathcal{S}_{\mathcal{O}} = (\mathsf{V}, \mathsf{E})$, where:

(i) The set of vertices is $\mathsf{V} = \mathsf{C} \cup \mathsf{C}^*$.
(ii) The set of edges is $\mathsf{E} = \mathsf{E}_{ctx} \cup \mathsf{E}_{map} \cup \mathsf{E}_{sem} \cup \mathsf{E}_{ctxAtr}$.
The edge sets are defined as follows:

- Concept-level semantic relations $\mathsf{E}_{ctx} = R_{ctx}$.
- Instantiation edges

$$\mathsf{E}_{map} = \{(c, (c, a)) \mid c \in \mathsf{C},\ a \in map(c)\}.$$

- Inherited semantic relations

$$\mathsf{E}_{sem} = \{((c, a), d) \mid (c, d) \in R_{ctx},\ a \in map(c)\}.$$

- Contextual argumentative relations $\mathsf{E}_{ctxAtr} = R_{ctxAtr}$.

The contextualized semantic graph enables the quantification of conceptual similarities within a debate by jointly considering abstract concepts and their attribute-instantiated counterparts. Because the graph explicitly represents both semantic structure and argumentative interactions, path-based metrics[35] can capture both structural proximity and discursive closeness. In the following example, we detail the conceptualization and the contextualized semantic graph associated with the debate about dress codes.

Example 5. Figure 11 shows the conceptualization for the example presented in Section 2. The figure illustrates the set of abstract identified i the discussion and the attributes associated with each concept through the function *map*. Each attribute represents a semantic descriptor extracted from the debate and link to a particular concept. Pairs (c, a) denote attribute-instantiated concepts, that is, contextualized interpretations of abstract concepts as they appear in the discourse. The discourse further establishes favorable and opposing relations between certain attribute-instantiated concepts. These relations do not operate at the abstract conceptual level; rather, they capture argumentative interactions between specific interpretations of the concepts involved. From this conceptualization we construct the contextualized semantic graph shown in Figure 12. The graph contains:

- Edges denoting concept-level semantic relations R_{ctx}.
- Abstract concept nodes $c \in \mathsf{C}$.
- Attribute-instantiated nodes $(c, a) \in \mathsf{C}^*$.
- Instantiation edges linking each concept c to its instances (c, a).
- Contextual relation between attribute-instantiated concepts specified by R_{ctxAtr}.

This enriched representation makes explicit both the semantic organization of concepts and the argumentative relations emerging between their attribute-instantiated interpretations. Such a structure later supports similarity computations and the identification of ad-populum argumentative communities.

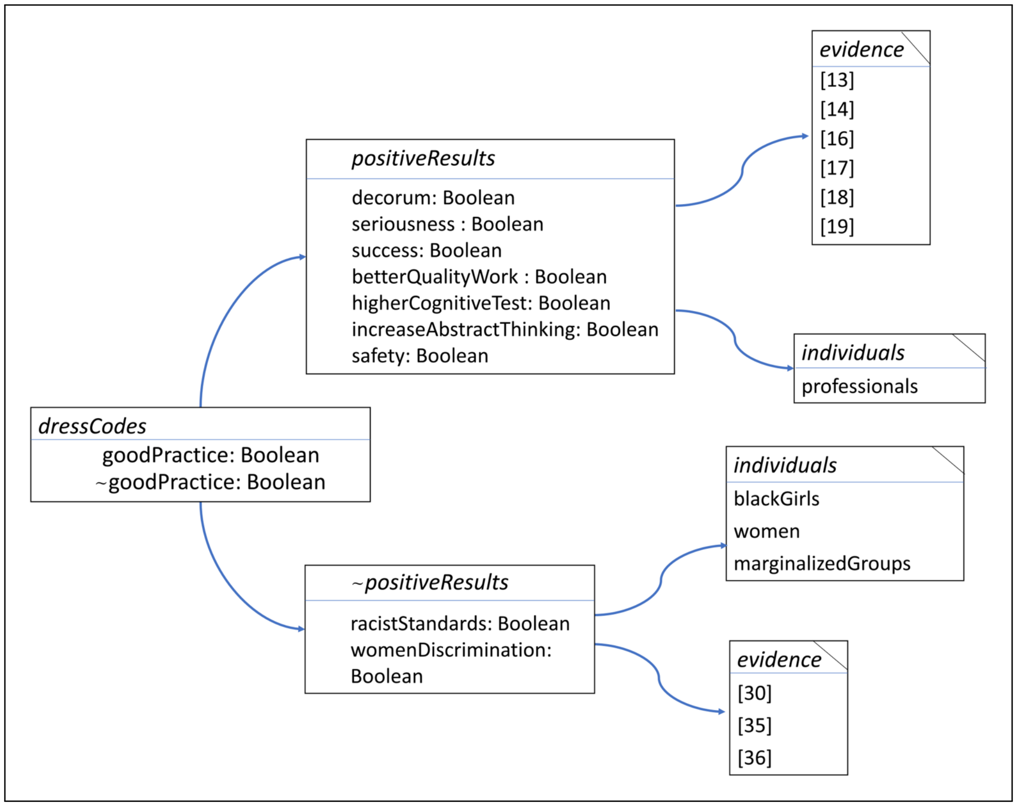

Figure 11. Proposed conceptualization for Example 2.

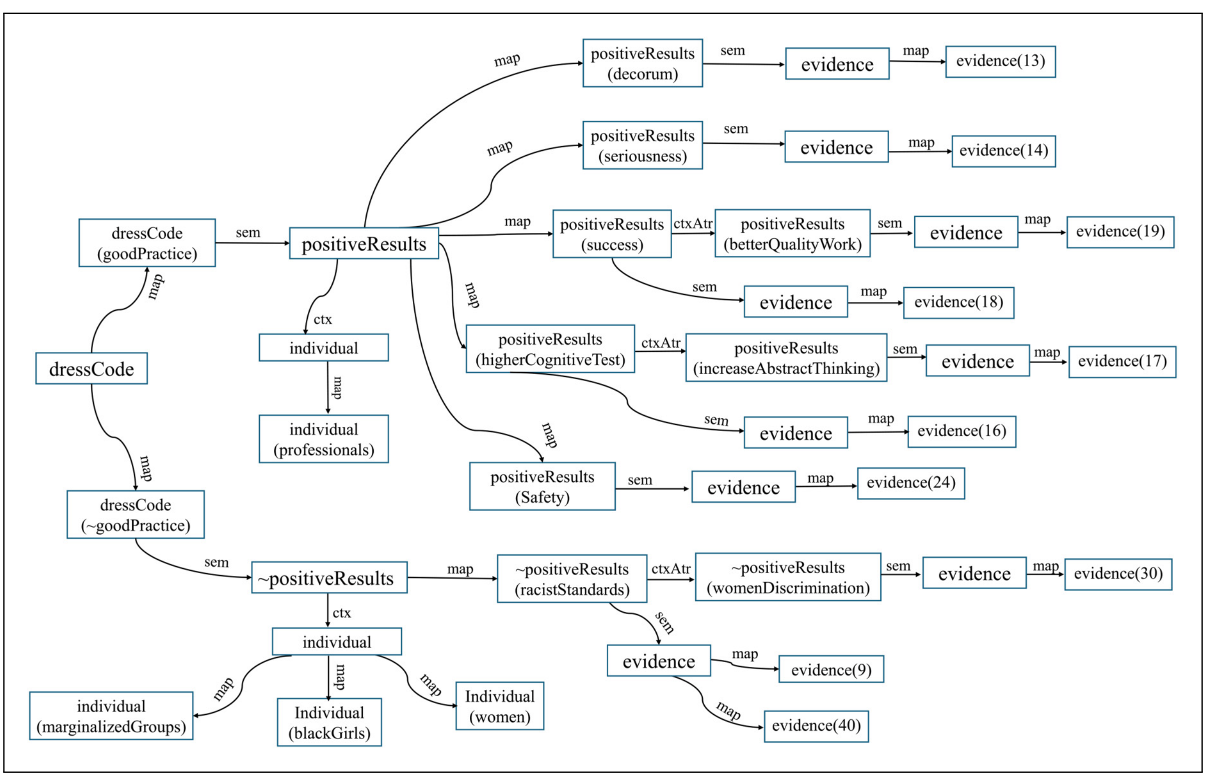

Figure 12. Contextualized semantic graph for Example 2.

This example illustrates how the contextualized semantic graph provides the structural basis for identifying ad-populum argumentation. By explicitly representing concepts, attribute-instantiated argumentative positions, and their discursive relations, the graph allows arguments to be linked to socially salient reference groups they implicitly invoke, such as institutional norms or collectively endorsed standpoints.

Starting from the structured debate, its conceptualization, and the contextualized semantic graph, we next identify the opinions expressed in the debate as coherent sets of attribute-instantiated concepts associated with a particular stance. These conceptual structures capture the semantic descriptors and contextual attributes through which each position in the debate is articulated. Based on this representation, we define a mechanism to compute the cohesion of an opinion by measuring the similarity between the conceptual structures associated with different arguments. This similarity is computed over the contextualized semantic graph using path-based metrics that capture the structural proximity between concepts and their instantiated attributes. As a result, arguments that rely on closely related conceptual descriptors and contextual attributes exhibit higher similarity and contribute to stronger internal cohesion within the corresponding opinion. From the perspective of *argumentum ad populum*, cohesion can be interpreted as an indicator of collective alignment with respect to a socially salient reference group. In other words, arguments that exhibit high conceptual similarity tend to reflect shared assumptions, values, or sources of authority invoked by participants in the debate. These shared elements correspond to the reference groups that underlie Walton's ad-populum schemes, such as expert communities, socially admired groups, or broadly accepted moral viewpoints. Under this interpretation, different ad-populum sub-schemes generate distinct patterns of alignment among arguments. Consequently, the cohesion values derived from the similarity analysis provide a principled way to detect clusters of arguments that exhibit consistent alignment with particular reference groups. These clusters correspond to argumentative communities, whose internal structure and influence patterns can then be analyzed within the proposed framework. Note that the same concept or attribute-instantiated concept may appear in arguments associated with different ad-populum labels (e.g., EX and AD) when distinct proponents or reference groups endorse the same claim. In such cases, each individual argument instantiates a single ad-populum sub-scheme according to the reference group it invokes.

In the next two sections, we instantiate the LAF to formally model the debate. In particular, we define argument labels that capture similarity-based cohesion, source-related features, and group affiliation, together with the algebraic structure required to propagate these labels through the argumentation graph. This formalization allows the framework to integrate conceptual similarity, argumentative interactions, and reference-group alignment into a unified computational representation of the debate.

5 Modeling argument features

We will now instantiate the algebra of argumentation labels introduced in Definition 2 to represent three characteristics of the knowledge used to model the debate: the primarily upheld viewpoints within a debate are identified by computing the similarity between arguments and enabling a structured interpretation of the argumentative positions; the quality of the source, considering factors such as expertise, reliability, and social acceptance; and the argumentative group to which the source is affiliated, enabling the identification of ideological or strategic alignments within the debate. It is important to highlight that these last two characteristics are closely intertwined, as will become evident in Section 6.

5.1 Modeling similarities between arguments

We propose a instantation to model the similarity between two pieces of knowledge using a label from the conceptualization. At this stage, the similarity between two concepts, c_1 and c_2, is determined using the *path metric*,[35] which is defined as follows:

$$\mathsf{sim}(c_1, c_2) = \frac{1}{1 + \mathsf{length}(c_1, c_2)}, \tag{1}$$

where $\mathsf{length}(c_1, c_2)$ represents the length of the shortest path between c_1 and c_2 in the contextualized semantic graph. When computing similarity, we rely on the structural connectivity of the graph; therefore, paths are considered as *undirected paths*, independently of the direction of the underlying relations. This measure is chosen to capture conceptual proximity rather than textual or lexical similarity; it directly reflects the length of conceptual paths between arguments, although its effectiveness depends on the quality and structure of the underlying semantic graph. Thus, the concepts connected by shorter paths are considered more similar, whereas concepts connected by longer paths exhibit lower similarity.

The following example illustrates how the similarity measure is computed for concepts in the dress codes debate using the contextualized semantic graph.

Table 1. Similarity calculation between elements of the semantic graph (Figure 12).

Elements	Similarity measure
sim(positiveResults(increaseAbstractThinking), (∼ positiveResults(racistStandards)))	$\frac{1}{1+7} = 0.125$
sim(positiveResults(success), dressCode(goodPractice))	$\frac{1}{1+2} = 0.33$
sim(positiveResults, positiveResults)	1
sim(dressCode(∼ goodPractice), ∼ positiveResults(womenDiscrimination))	$\frac{1}{1+3} = 0.25$

Example 6. Considering the conceptualized semantic graph depicted in Figure 12, Table 1 presents the similarity measures between the concepts mentioned in the debate.

We point out that similarity characterizes the relation between two pieces of knowledge, but a deeper analysis of this feature is outside the scope of this article. As mentioned before, it is possible to instantiate it using any other well-studied function, such as the one proposed by David and Mailly.[36] The impact of the similarity measure on the argumentative process will be addressed later using the LAF manipulation processes.

This similarity measure is used within the algebra of argumentation labels defined in Definition 2 to represent the knowledge behavior (For a comprehensive examination of the mathematical properties associated with each operator, please refer to the work by Budan et al. (2015) in Budán et al.[21])

Instantiation 1 (S-Algebra of Labels)**.** *The algebra of argumentation labels* $\mathsf{S} \in \mathcal{A}$ *is used to represent the similarity measure between the piece of knowledge and its ancestor in the argumentative graph. The domain of labels S is the real interval* $[0, 1]$ *representing a normalized similarity valuation, where* $\top = 1$ *is the maximum valuation (and neutral element for* $\odot$*), while* $\bot = 0$ *is the minimum valuation (and neutral element for* $\oplus$ *and* $\ominus$*). Thus, let* $\alpha_1, \alpha_2 \in S$ *be two labels, the operators of support, conflict and aggregation over labels representing the similarity valuations associated with arguments are specified as follows:*

- $\odot : S \times S \rightarrow S$ *is called a support operation and it is used to model the cohesion of an argument. Here it is defined as:* $\alpha_1 \odot \alpha_2 = \alpha_1 + \alpha_2 - \alpha_1 * \alpha_2$*, for all* $\alpha_1, \alpha_2 \in S$ *(Probabilistic Sum).*
- $\oplus : S \times S \rightarrow S$ *is an aggregation operation, which states that if a claim has more than one argument supporting it, the claim's cohesion valuation is the sum of the similarity valuations of the arguments supporting it. It can be defined as:* $\alpha_1 \oplus \alpha_2 = \dfrac{\alpha_1 + \alpha_2}{1 + \alpha_1 \alpha_2}$*, for all* $\alpha_1, \alpha_2 \in S$ *(Einstein Sum).*
- $\ominus : S \times S \rightarrow S$ *is a conflict operation or impaired similarity; it reflects that the cohesion of a conclusion is weakened by the similarity value of its contrary. It can be defined as follows:*

$$\text{For all } \alpha_1, \alpha_2 \in S,\ \alpha_1 \ominus \alpha_2 = \begin{cases} \dfrac{\alpha_1 - \alpha_2}{1 - \alpha_2} & \text{if } \alpha_1 > \alpha_2, \alpha_2 \neq 1, \\ 0 & \text{otherwise.} \end{cases}$$

Regarding the aggregation operation, note that the aggregation intuition is designed to capture the cumulative effect of multiple similarity contributions and should be interpreted as a global measure of coherence between arguments within the framework. With respect to the conflict operation, note that when α_1 takes the value 1, the conflict operation reaches value 1, that is, the cohesion of α_1 is entirely preserved. Next, we present an example to illustrate how the labels work.

Example 7. Figure 13 exemplifies the support operation for the calculus of partial cohesion for a conclusion C as $0.75 = 0.5 + 0.5 - 0.5 * 0.5$, through R_l, given the similarity measures between C and its premises P_i and P_j. A similar analysis yields the valuation for C from its premises P_k and P_l, where $0.65 = 0.5 + 0.3 - 0.5 * 0.3$. The valuation for the conclusion C is the result of operating the aggregation of the similarity valuations between the previous arguments: $\frac{0.75+0.65}{1+0.75\times0.65} = 0.94$.

Figure 14 exemplifies conflict resolution using the impaired similarity operation defined above, assuming that conclusion C with a cohesion value of 0.94 conflicts with conclusion D, which has a cohesion value of 0.82. The final weakened value for each conclusion is calculated in accordance with the conflict operation definition. The values taken into account were, first $\alpha_1 = 0.94$ and $\alpha_2 = 0.82$ yielding the result $\frac{0.94-0.82}{1-0.82} = 0.76$. And second, $\alpha_1 = 0.82$ and $\alpha_2 = 0.94$ yielded a result of 0.

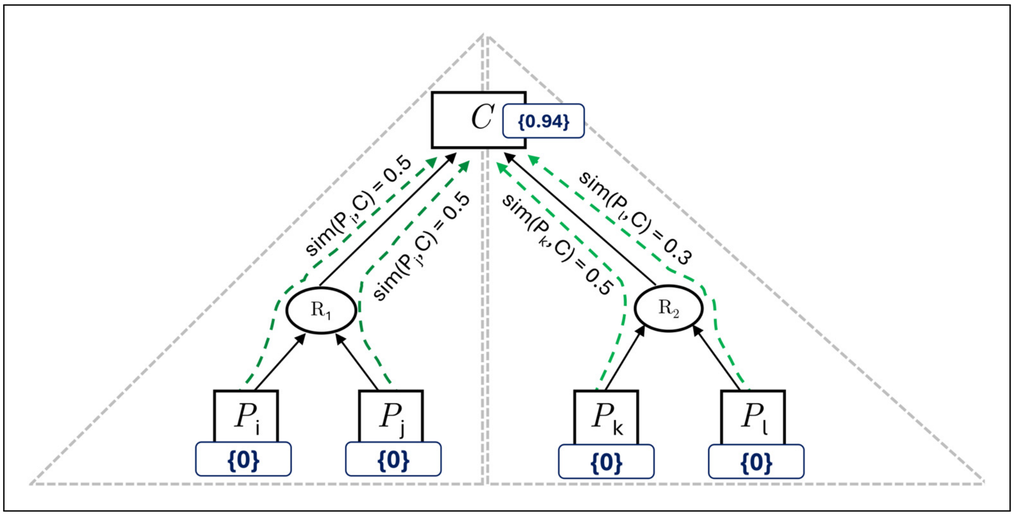

Figure 13. Support and aggregation from the similarity measure.

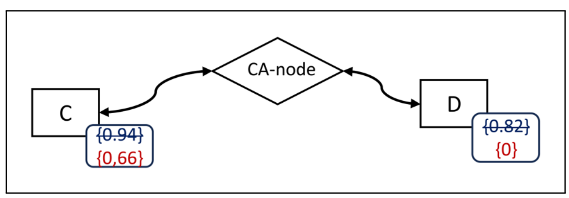

Figure 14. Conflict resolution using similarity measures.

In the approach presented by Budán et al.,[14,21,24] operations are defined on the labels of two pieces of knowledge. In contrast, the above algebra defines operations based on the similarity measure, which is determined by the conceptual distance between two concepts in the contextualized semantic graph.

In Section 6 and 7, we propose using these kind of labels to evaluate the *cohesion of an argument*. This evaluation is based on the *similarity labels* of the supporting arguments, which measure how closely related or consistent the supporting points are with each other.

5.2 Modeling the quality of the argument's proponent

We now model the proponent, which consists of two type of labels: one quantitative and one qualitative. The quantitative labels represent the *quality of the proponent*, which includes factors such as their degree of expertise, whether the source is in a special position to know, their degree of morality, and other relevant attributes.

Instantiation 2 (Q-Algebra Modeling Quality Proponents)**.** *The algebra of argumentation labels* $\mathrm{Q} \in \mathcal{A}$ *is used to represent the quality measure associated with the source that proposes a piece of knowledge. The domain of labels* Q *is the real interval* $[0, 1]$ *representing the quality of the source, where* $\top = 1$ *is the maximum valuation (and neutral element for* $\boxdot$*), while* $\bot = 0$ *is the minimum valuation (and neutral element for* $\boxplus$ *and* $\boxminus$*). Thus, let* $\alpha_1, \alpha_2 \in Q$ *be two labels, the operators of support, conflict and aggregation over labels representing the quality valuations associated with arguments are specified as follows:*

- $\boxdot : Q \times Q \rightarrow Q$ *is called a support operation; it is used to model the qualification of the sources, and satisfies the commutativity, associativity, and monotony properties. It is defined as:* $\beta_1 \boxdot \beta_1 = \beta_1 * \beta_2$ *for all* $\beta_1, \beta_2 \in Q$*.*
- $\boxplus : Q \times Q \rightarrow Q$ *is called an aggregation operation, it states that if a claim has more than one argument supporting it, the qualification corresponding to the source claim is the sum of the qualifications of the sources that present arguments supporting it. The aggregation for all* $\beta_1, \beta_2 \in Q$ *is defined as* $\beta_1 \boxplus \beta_2 = \max(\beta_1, \beta_2)$*.*

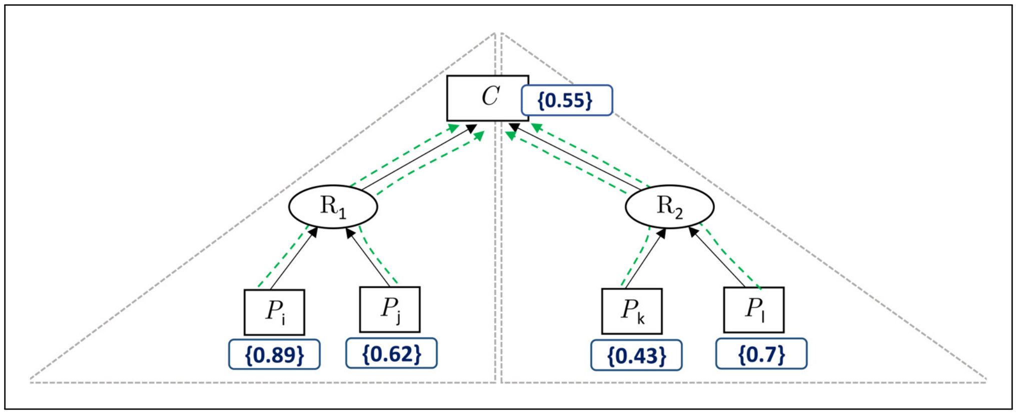

Figure 15. Support and aggregation from the quality proponents.

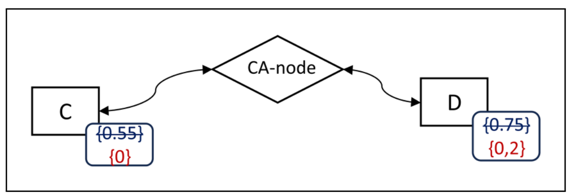

Figure 16. Conflict resolution considering quality proponents.

- $\boxminus : Q \times Q \rightarrow Q$ *is called a conflict operation and it reflects that the qualification of a conclusion proponent is weakened by the qualification value of its contrary. For* $\beta_1, \beta_2 \in Q$, *it is defined as:* $\beta_1 \boxminus \beta_2 = \max(\beta_1 - \beta_2, 0)$. *The conflict operation satisfies all the properties defined for the conflict operation in the work of Budan et al.*[21]

The values of Q labels are assumed to be provided by knowledge engineers during the debate mining and modeling process. The definition of procedures for extracting or assigning these values is outside the scope of this work.

Example 8. Figure 15 exemplifies the support operation used to calculate the qualification of the sources. The qualification for argument C, through R_l, is determined as $0.55 = 0.89 * 0.62$, given the values of premises P_i and P_j. A similar analysis yields the valuation for C from premises P_k and P_l, resulting in $0.30 = 0.43 * 0.70$. The final valuation for conclusion C is obtained by aggregating the previous arguments: $max(0.55, 0.30) = 0.55$.

Figure 16 exemplifies the conflict operation defined in the Q-Algebra, assuming that conclusion C with a proponent quality value of 0.55 conflicts with conclusion D, whose proponent has a quality value of 0.75. The final weakened value for each conclusion is calculated in accordance with the conflict operation definition. The values taken into account were, first $\beta_1 = 0.55$ and $\beta_2 = 0.75$ yielding the result $max(0.55 - 0.75, 0) = 0$, and second, $\beta_1 = 0.75$ and $\beta_2 = 0.55$ yielded a result of 0.2.

The qualitative labels denote the argumentative *type of community* to which the proponent belongs. According to Walton,[9,10] the group to which the argument's proponent belongs is a crucial factor in determining the dominant sub-scheme in a debate and, consequently, the discursive communities involved.

Instantiation 3 (AP-Algebra Modeling Ad-Populum Proponents)**.** *The algebra of argumentation labels* $\mathsf{AP} \in \mathcal{A}$ *is used to represent instantiations of the Ad-Populum AS or type of community to which the proponent belongs. Let* $AP_{base} = \{POP$ *(original ad-populum AE), PtoK (position to known proponent), EX (expert proponent), MO (morally good proponent), PR (from elite or prestige proponent), and AD (admired people or vain proponent)*$\}$ *be the set of basic ad-populum categories. The domain of labels AP is defined as* $AP = 2^{AP_{base}}$, *that is, each label is a set of ad-populum categories, allowing an argument to be associated with one or more categories simultaneously. Furthermore, the categories in* AP_{base} *are equipped with the total order* $\bot \prec POP \prec AD \prec MO \prec PtoK \prec PR \prec EX \prec \top$, *ranging from the least demanding to the most demanding in terms of the proponents' requirements. Then, let* $\gamma_1, \gamma_2 \subseteq AP_{base}$ *be two labels, the operators*

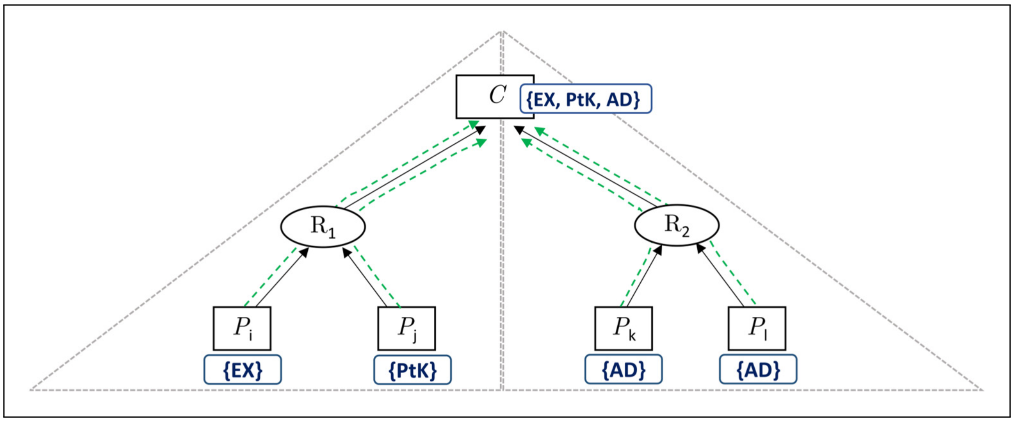

Figure 17. Conveyance operation in the line of reasoning.

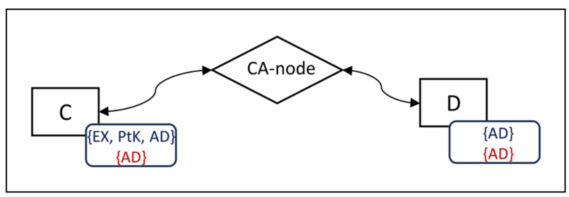

Figure 18. Conflict operation resolution between confronted groups.

of support, agregattion and conflict over labels representing the ad-populum proponents associated with arguments are specified as follows:

- $\circledcirc : 2^{AP_{base}} \times 2^{AP_{base}} \rightarrow 2^{AP_{base}}$ *is the conveyance operation representing the groups through which the line of reasoning is spreading. Thus, the operator is applicable to the support and aggregation structures. It is defined as* $\circledcirc(\gamma_1, \gamma_2) = \gamma_1 \cup \gamma_2$, *for* $\gamma_1, \gamma_2 \subseteq AP_{base}$.
- $\Cap : 2^{AP_{base}} \times 2^{AP_{base}} \rightarrow 2^{AP_{base}}$ *is called a conflict operation, and it reflects that when two lines of reasoning are confronted, the groups of proponents that are non-defeated will prevail. It can be defined as (Recall that the conflict relation is bidirectional according to Definition 4.):*

$$\Cap(\gamma_1, \gamma_2) = \begin{cases} \gamma_1 \cap \gamma_2 & \text{if } \gamma_1 \cap \gamma_2 \neq \emptyset, \\ \gamma_1 & \text{otherwise.} \end{cases}$$

As before, the values of AP labels are assumed to be provided by knowledge engineers, and the corresponding procedures are outside the scope of this work. The following example illustrates how the labels function under this algebra.

Example 9. Figure 17 illustrates the conveyance operation used to identify all the groups supporting an argument. In the case of argument C, the label obtained through R_1 is $\{EX, AD, PtK\}$, while the label obtained through R_2 is $\{AD\}$.

Figure 18 illustrates the conflict operation in the AP-Algebra. Here, a conflict arises between conclusion C, associated with the label $\{EX$ (experts), PtK (position-to-known proponent), AD (admired people)$\}$, and conclusion D, associated with the label $\{AD\}$. The conflict operator determines that such conflicts are resolved by focusing exclusively on the shared groups. Applying the conflict operator yields $\{EX, PtK, AD\} \cap \{AD\} = \{AD\}$. It is worth noting that this example only illustrates the application of the AP-Algebra conflict operator. However, the operator's full scope is discussed later when quantitative and qualitative labels form an ordered pair, where the quantitative component describes the qualitative one.

Therefore, the purpose of the labels' operations presented above is twofold: on the one hand, to determine the qualifications associated with the supporting proponents of a claim, and on the other hand, to identify the set of ad-populum groups conveying a particular stance. We now show an example illustrating how these labels work together.

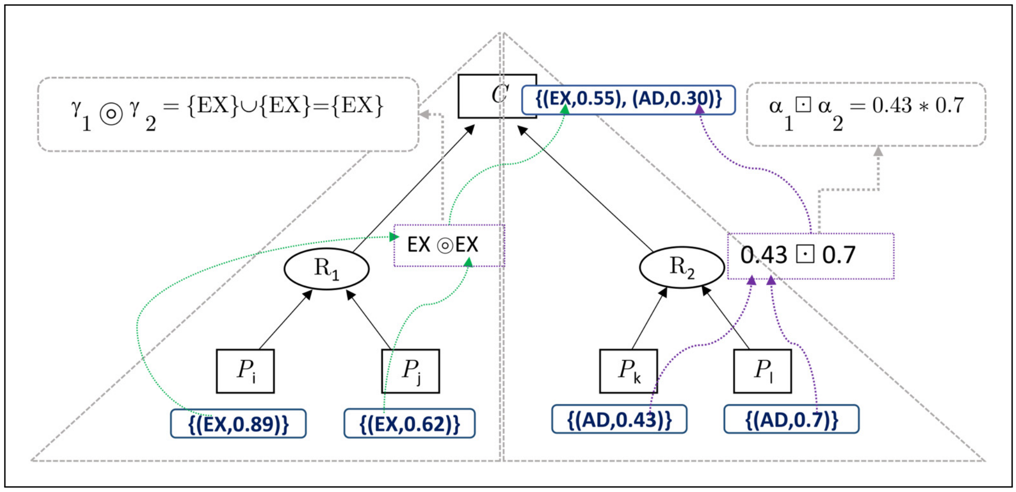

Figure 19. Support and aggregation of labels describing the argument's proponent.

Example 10. Figure 19 describes the behavior of the labels for the support, aggregation, and conveyance operations. Note that the label associated with the piece of knowledge identified as C denotes that the conclusion is conveyed by two ad-populum categories: experts (EX) with a qualification of 0.55 (where $0.55 = 0.89 \times 0.62$), and admired people (AD) with a qualification of 0.30 provided by the other argument supporting this conclusion. Consequently, the AP-component of the resulting label is $\{EX, AD\}$, while the corresponding qualifications are represented through the associated Q-labels. In this sense, the labels associated with the piece of knowledge C are obtained by applying the conveyance operation to the labels of its ancestors, yielding the union of the ad-populum categories that support the conclusion together with their corresponding qualifications.

In the next section, we build upon this algebraic framework to systematically identify all groups supporting each viewpoint, taking into account the nuances of their respective stances. By applying this approach, we aim to capture not only the alignment of arguments across different groups, but also the degrees of agreement and divergence among them, providing a more comprehensive understanding of the argumentative structure.

6 An ad-populum LAF

Towards our objective of identifying discursive communities and analyzing how they influence one another, we now instantiate a LAF that is explicitly grounded in argumentum ad populum. In this framework, labels are not limited to abstract notions of strength or preference, but encode socially grounded attributes associated with the reference group implicitly invoked by an argument, such as majority opinion, expert endorsement, or group identity. As a result, the propagation of labels throughout the argumentation graph captures how appeals to popular opinion give rise to collectively endorsed standpoints and to the emergence of coherent argumentative communities.

We refer to this instantiation as the *Ad-Populum Labeled Argumentation Framework* (AdP-LAF). It relies on the associated conceptualization to reconstruct the main lines of reasoning in a debate and employs the algebras defined in the previous section to characterize both the internal coherence of opinions and the influence relations between the communities they induce.

Instantiation 4 (Ad-Populum Framework)**.** *Given a conceptualization $\mathcal{O}$ based on $\mathcal{L}' = \mathrm{C} \cup \Delta$, an* AdP-LAF *is instantiated as follows:*

(i) *$\mathcal{L}$ is a logical language defined from $\mathcal{L}'$, such that: An atom is of the form $p(t_1, \ldots, t_n)$, with $p \in \mathrm{C}$ and $t_1, \ldots, t_n \in \Delta$; a literal is a ground atom X or $\sim X$; a presumption is a literal; and a defeasible rule is of the form $C \multimap P_1, \ldots, P_n$, where C is a ground literal (conclusion) and $P_1, \ldots, P_n$ are ground literals (premises), with $n > 0$.*

(ii) *$\mathcal{R} = \{dMP\}$ is an inference rule defined in terms of $\mathcal{L}$ (e.g., a defeasible Modus Ponens with premises and conclusion in $\mathcal{L}$).*

(iii) *$\mathcal{K}$ is the knowledge base composed by a set of presumptions and defeasible rules representing the knowledge from the debate.*

r_1: dressCode(goodPractice) $\prec$ positiveResults(decorum), individuals(professionals) : {0, (PtoK, 0.33)}
r_2: dressCode(goodPractice) $\prec$ positiveResults(seriousness), individuals(professionals) : {0, (MO, 0.54)}
r_3: dressCode(goodPractice) $\prec$ positiveResults(success), individuals(professionals : {0, (EX, 0.63)}
r_4: dressCode(goodPractice) $\prec$ positiveResults(betterQualityWork), individuals(professionals) : {0, (PtoK, 0.85)}
r_5: dressCode(goodPractice) $\prec$ positiveResults(higherCognitiveTest), individuals(professionals) : {0, (AD, 0.70)}
r_6: dressCode(goodPractice) $\prec$ positiveResults(increaseAbstractThinking), individuals(professionals) : {0, (AD, 0.42)}
r_7: dressCode(goodPractice) $\prec$ positiveResults(safety) : {0, (MO, 0.57)}
r_8: dressCode(~goodPractice) $\prec$ ~positiveResults(racistStandards) , individuals(blackGirls), individuals(women) : {0, (PR, 0.80)}
r_9: dressCode(~goodPractice)$\prec$ ~positiveResults(womenDiscrimination), individuals(women), individuals(marginalizedGroup) {0, (EX, 0.76)}

positiveResults(success) : {0, (EX, 0. 79)}
positiveResults(seriousness) : {0, (MO, 0. 62)}
positiveResults(decorum) : {0, (AD, 0.46)}
~positiveResults(racistStandards) : {0, (PR, 0.78)}
positiveResults(betterQualityWork) : {0, (MO, 0. 29)}
individuals(professionals) : {0, (PtoK, 0.53)}
dressCode(goodPractice) : {0, (PtoK, 0.83)}
dressCode(~goodPractice) : {0, (EX, 0.54)}
positiveResults(safety) : {0, (MO, 0. 34)}
positiveResults(higherCognitiveTest) : {0, (AD, 0.57)}
~positiveResults(womenDiscrimination) : {0, (PR, 0.27)}
positiveResults(increaseAbstractThinking) : {0, (AD, 0.78)}
individuals(blackGirls) : {0, (MO, 0.33)}
individuals(marginalizedGroup) : {0, (PtoK, 0.83)}
individuals(women) : {0, (EX, 0.73)}

Figure 20. The knowledge base for the motivational example.

(iv) $\mathcal{A} = \{\mathsf{S}, \mathsf{AP}, \mathsf{Q}\}$ *represents the set of algebras corresponding to similarity labels, the features of the proponent's group, and the proponent's group itself, respectively.*

(v) $\mathcal{F} : \mathcal{K} \longrightarrow S \times AP \times Q$ *is a function assigning to each element of* $\mathcal{K}$ *three labels: one corresponding to the similarity between the element and its ancestor, one identifying the proponent's group, and one representing the qualification associated with that group. Although the labels in AP and Q are interpreted jointly when representing ad-populum information, they remain distinct components of the labeling function.*

Next, we will apply the elements of our formalism to accurately represent the example that we initially presented.

Example 11. Figure 20 provides details of the KB developed from the opinions presented in Section 2, and the conceptualizations developed in Example 5 from Section 4. In addition, we initialize the similarity value of a piece of knowledge with "0" conforming to Instantiation 1. Furthermore, Figure 21 is the argumentation graph representing the debate presented in Section 2, where the discussion on dress codes is formally modeled. We identify two major arguments: one supporting the positive effects of dress codes and another emphasizing their negative consequences. It is important to highlight that both are aggregated arguments, each supported by independent chains of reasoning. From one side we can notice, for example, proponents arguing that dress codes contribute to a more structured and productive environment through seven independent lines of reasoning, promoting decorum, enhancing career success, improving work quality, and the development of abstract thinking. On the other side, opponents highlight two main negative consequences. The first is the reinforcement of racist standards, as certain dress code requirements disproportionately affect individuals from minority groups, as well as gender discrimination, particularly against women, since dress codes tend to be more restrictive for them, reinforcing gender inequalities.

The argumentative graphs introduced in this section play a central role in the ad populum interpretation of the debate. From an ad-populum standpoint, these graphs make explicit how arguments cluster around shared reference groups, with labels reflecting different forms of appeal to popular opinion. In this way, the graphical representation manifests how collective endorsement and group-based influence are structurally encoded in the framework.

Building on this interpretation, we now use an AdP-LAF to find and characterize the main opinions in a debate, and to determine how coherent each opinion is. First, we need to find the arguments and calculate their cohesion. To do that, we use the concept of argumentative graph (Definition 4) and argument (Definition 5) in the domain of AdP-LAF, using G_Γ and Arg_Γ to denote the argumentative graph and the set of arguments, respectively. Then, we will use the defined labels and the operations of the corresponding algebras to obtain the values associated with each argument that belongs to G_Γ. According to the procedure proposed by Budán et al.,[21] the labeling of each node in the graph requires the assignment of specific labels for the defined algebras to individual nodes, representing both *accrued valuations* denoted by μ^X, and

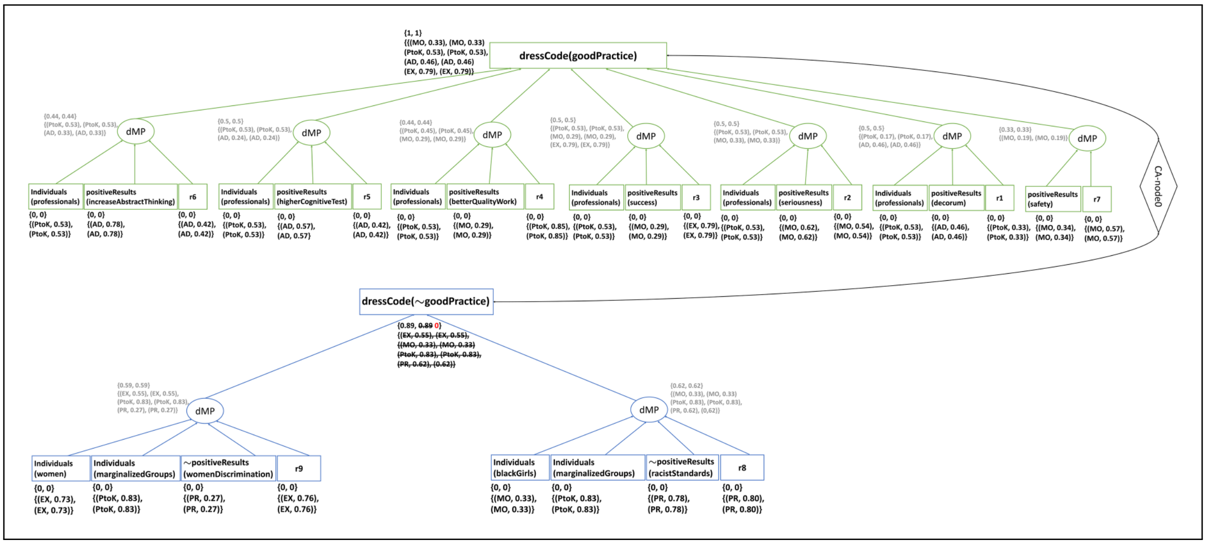

Figure 21. The labeled argumentation graph for Example 2.

weakened valuations indicated by δ^X, respectively, where X is the piece of knowledge in the graph. Thus, in the following we instantiate the labeling procedure for an argumentation graph, which derives a system of equations that characterizes the knowledge that describe the domain of the discourse.

Instantiation 5 (AdP-Labeling Procedure for a G_Γ)**.** *Let Γ be an instantiated AdP-LAF, and G_Γ be its corresponding argumentation graph. Let* S, Q, *and* AP *be the algebras representing the similarity measure, the quality of the proponent's group, and the proponent's group itself, respectively. An AdP-labeling argumentation procedure assigns:*

- *Two valuations $\mu_S^X, \delta_S^X \in S$ label all the graph's nodes, where μ_S^X represents the aggregation of partial cohesion of the reasons supporting the claim X, and δ_S^X displays the cohesion of the claim after taking conflicts into account.*
- *Two compound valuations (μ_{AP}^X, μ_Q^X) and $(\delta_{AP}^X, \delta_Q^X)$ to all the graph's nodes where $\mu_{AP}^X, \delta_{AP}^X \in AP$, and $\mu_Q^X, \delta_Q^X \in Q$, and μ_{AP}^X represents the aggregation of all the types of argumentative groups adhering to a claim X, μ_Q^X represents the aggregation of the qualities of the proponents adhering to a claim X, δ_{AP}^X selects the type of argumentative groups related to non-defeated proponents, and δ_Q^X stands for the quality of non-defeated proponents.*

The valuation of each I-node X is determined as follows:

(i) *If X has no inputs, then its accrued valuations are:*
$\mathcal{F}_1(X) = \mu_S^X = 0$ *(the node has no ancestors), and*
$\mathcal{F}_2(X) = (\mu_{AP}^X, \mu_Q^X)$*, assigning both the initial proponent quality and the group to which the proponent belongs (jointly, as the $AP \times Q$ component of $\mathcal{F}$), where $\mu_Q^X \neq 0$.*

(ii) *If X has an input from a CA-node representing a conflict with another node $\sim X$, then:*

$\delta_S^X = \mu_S^X \ominus \mu_S^{\sim X}$, *Obtaining the weakened valuation of a piece of knowledge via similarity with its contrary.*

If $\delta_S^X = 0$, then

all compound labels for X are eliminated; *Compound labels are not considered for a defeated piece of knowledge.*

else

while $\mu_{AP}^X = \mu_{AP}^{\sim X}$ do

$\delta_Q^X = \mu_Q^X \ominus \mu_Q^{\sim X}$ *If two pieces of knowledge are contrary, the weakened proponent's qualification is computed for each common argumentative group.*

end-while
If $\delta_Q^X \neq 0$, *then*

$\delta_{AP}^X = \mu_{AP}^X$; *The argumentative group is maintained while the valuation w.r.t the quality of the proponent is still considerable.*

$(\delta_{AP}^X, \delta_Q^X) = (\mu_{AP}^X, \mu_Q^X)$ *The compound label is composed by the weakened proponent's qualification and the proponent's argumentative group.*

else

eliminate (μ_{AP}^X, μ_Q^X) *When the quality of the proponent valuation is defeated, the compound label is eliminated.*

end-If
If μ_{AP}^X *is not an element in* $\mu_{AP}^{\sim X}$ *then*

$\delta_{AP}^X = \mu_{AP}^X$ *and* $\delta_Q^X = \mu_Q^X$ *If the argumentative group is only present in a piece of knowledge, then the group labels remain unchanged.*

end-If
end-If
If there is no input from a CA-node, then $\delta_S^X = \mu_S^X$ *and* $(\mu_{AP}^X, \mu_Q^X) = (\delta_{AP}^X, \delta_Q^X)$.

(iii) *Let X be an I-node with inputs from relations* $R_1, \ldots, R_k$, *where each* R_j *has premises* $X_1^{R_j}, \ldots, X_{n_j}^{R_j}$.

(a) $X \in \mathcal{K}$.
The accrued similarity label of X is given by the following equation:

$$\mu_S^X = \mathcal{F}_1(X) \oplus \left[\oplus_{j=1}^{k} \left(\odot_{t=1}^{n_j} \delta_S^{X_t^{R_j}}\right)\right].$$

If all premises in each R_j *belong to the same proponent group, that is,* $\mu_{AP}^{X_t^{R_j}} = \mu_{AP}^{X_y^{R_j}}$, *then the accrued proponent quality label of X is*

$$\mu_Q^X = \mathcal{F}_2(X) \circledcirc \left[\circledcirc_{j=1}^{k} \left(\circledcirc_{t=1}^{n_j} \delta_Q^{X_t^{R_j}}\right)\right].$$

For an I-node X in the KB that corresponds to an aggregated (accrued) conclusion, the similarity accrual label of X is obtained by aggregating the valuations resulting from each intervening support operation contributing to X. Then, the aggregation of the proponent quality is computed using the support operator defined in the P-Algebra. This computation is carried out independently for each qualitative label, that is, the quality of each group of proponents supporting X is obtained separately.

(b) $X \notin \mathcal{K}$.
The accrued similarity label of X is computed as follows:

$$\mu_S^X = \oplus_{j=1}^{k} \left(\odot_{t=1}^{n_j} \delta_S^{X_t^{R_j}}\right).$$

If all premises in each R_j *belong to the same proponent group, then*

$$\mu_Q^X = \circledcirc_{j=1}^{k} \left(\circledcirc_{t=1}^{n_j} \delta_Q^{X_t^{R_j}}\right), \quad \mu_{AP}^X = \circledcirc_{j=1}^{k} \left(\circledcirc_{t=1}^{n_j} \delta_{AP}^{X_t^{R_j}}\right).$$

For notational convenience, we denote by $\mathcal{F}_1$ *the projection of* $\mathcal{F}$ *onto the similarity algebra S, and by* $\mathcal{F}_2$ *the projection onto the pair* $AP \times Q$. *That is,* $\mathcal{F}_1 : \mathcal{K} \to S$ *and* $\mathcal{F}_2 : \mathcal{K} \to AP \times Q$, *where the latter groups the ad-populum and quality*

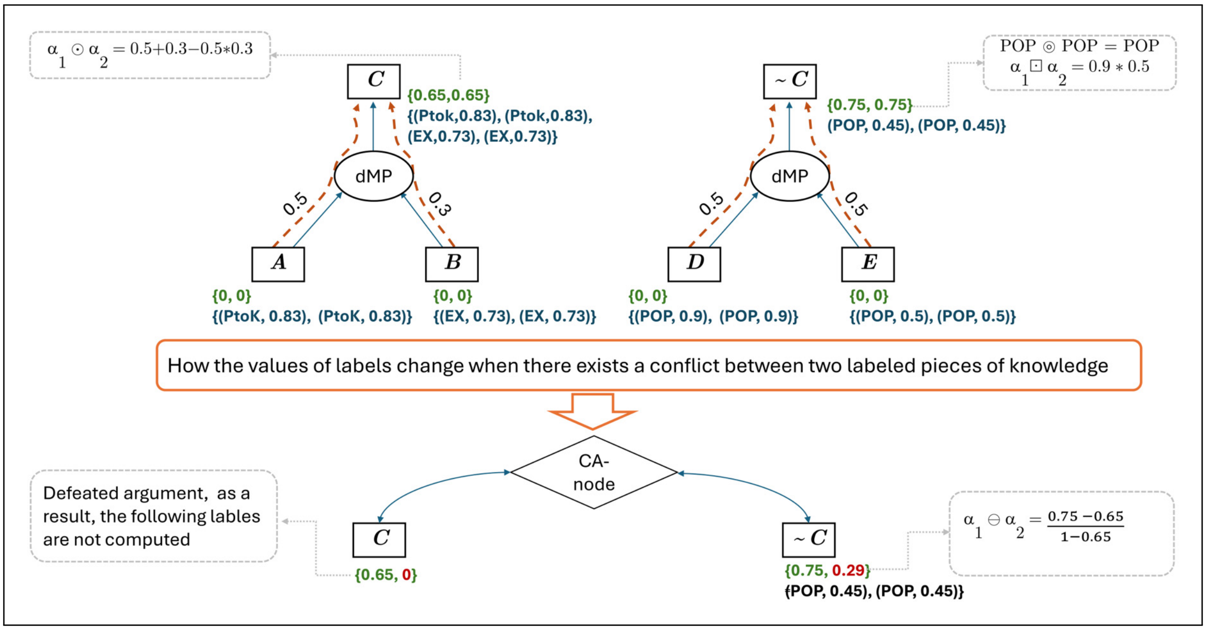

Figure 22. Graphical overview of the AdP-labeling procedure.

components in accordance with their joint interpretation introduced in Instantiation 4. No additional labeling functions are introduced.

Figure 22 provides a structured visual representation of the AdP-labeling procedure defined in Instantiation 5. The figure illustrates the main stages of the procedure, showing how similarity labels and proponent-related labels are computed, combined, and possibly weakened through support and conflict relations. In particular, it highlights the distinction between accrued and weakened valuations, as well as the different cases depending on whether a node belongs to the knowledge base or is derived through inference. Crossed-out labels indicate values that are no longer considered because the corresponding argument has been defeated during the evaluation process.

Example 12 demonstrates the labeling procedure for the conveyance operation by incorporating instantiated pieces of knowledge derived from our motivational example, providing a concrete illustration of how the process unfolds in a practical context. Moreover, Figure 21 depicts the completed labeled argumentation graph for the example provided in Section 2, where the gray labels represent the auxiliary or intermediate calculations made to reach the final claim label. According to Instantiation 5, the opinion against dress code is defeated and, as a result, its labels are weakened.

Example 12. In Figure 23, we can see that the claim about dress code as good practice is supported by three premises. The similarity between the claim and the supports is obtained from the distance between those elements in the contextualized semantic graph. Thus, the similarity between the premises and the conclusion is $0.25 + 0.33 - 0.25 \times 0.33 = 0.5$. Additionally, supports are conveyed by people in a position to know about the topic by people who appeal to vanity. Given that two supports are provided by people in a position to know, the label for the claim with respect to this group is obtained from the product of the qualities of its proponents. Finally, the values assigned to the levels remain unchanged because they are not under attack.

To complement the visual representation in Figure 21, in Table 2, we report a compact representation of pairwise similarities between I-nodes that are effectively related through inference steps. Rather than presenting a full similarity matrix, which would be largely sparse and difficult to interpret, we restrict attention to pairs of I-nodes that are connected through an RA-node in the argumentation graph. This choice reflects the intended semantics of the similarity measure: similarity values are meaningful only when two I-nodes jointly participate in a support relation, that is, when they contribute as premises or conclusions of the same rule application. Pairs of nodes connected exclusively through CA-nodes (conflict relations), as well as pairs with no direct argumentative connection, are thus excluded from the table. Leaf I-nodes are assigned a similarity value of 0, in accordance with the definition of the similarity algebra, as no cohesion can be derived in the absence of inferential support. Accordingly, the table reports only non-zero similarity values corresponding

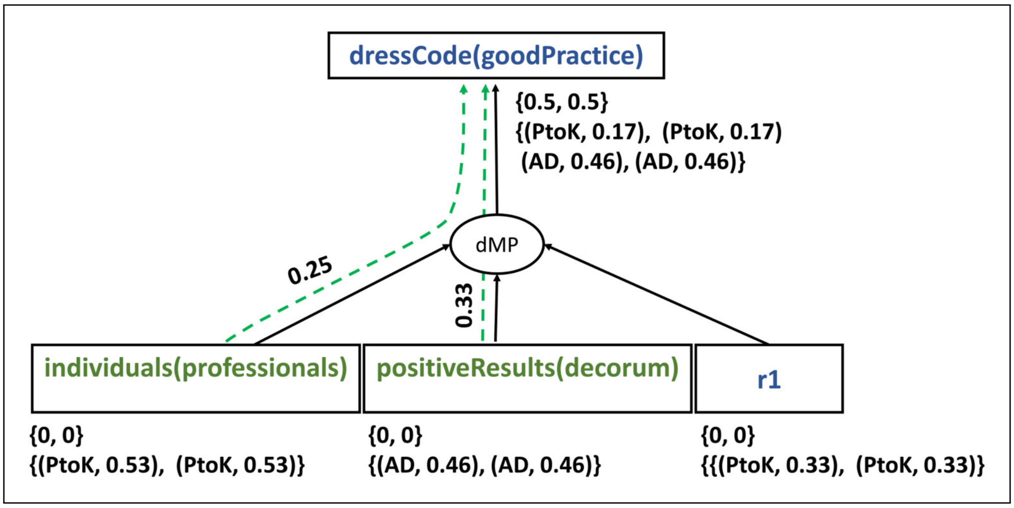

Figure 23. Example labeled opinion in favor of dress codes using the conveyance operation from Example 2.

Table 2. Similarity relations between I-nodes mediated by RA-nodes.

I-node X_1	I-node X_2	Sim (X_1, X_2)	I-node X_1	I-node X_2	Sim (X_1, X_2)
dc(gp)	dc(∼gp)	0.33	pr(d)	pr(hc)	0.33
dc(gp)	pr(s)	0.33	pr(d)	pr(ia)	0.25
dc(gp)	pr(se)	0.33	pr(bq)	pr(hc)	0.25
dc(gp)	pr(d)	0.33	pr(bq)	pr(ia)	0.20
dc(gp)	pr(bq)	0.25	pr(hc)	pr(ia)	0.50
dc(gp)	pr(hc)	0.33	pr(s)	ind(p)	0.25
dc(gp)	pr(ia)	0.24	pr(se)	ind(p)	0.25
pr(s)	pr(se)	0.33	pr(d)	ind(p)	0.25
pr(s)	pr(d)	0.33	pr(bq)	ind(p)	0.20
pr(s)	pr(bq)	0.50	pr(hc)	ind(p)	0.25
pr(s)	pr(hc)	0.33	pr(ia)	ind(p)	0.20
pr(s)	pr(ia)	0.25	¬pr(rs)	ind(w)	0.25
pr(se)	pr(d)	0.33	¬pr(rs)	ind(bg)	0.25
pr(se)	pr(bq)	0.25	¬pr(rs)	ind(mg)	0.25
pr(se)	pr(hc)	0.33	¬pr(rs)	¬pr(mg)	0.50
pr(se)	pr(ia)	0.25			

dc = *dressCode*, gp = *goodPractice*, pr = *positiveResults*, s = *success*, se = *seriousness*, d = *decorum*, bq = *betterQualityWork*, hc = *higherCognitiveTest*, ia = *increaseAbstractThinking*, rs = *racistStandards*, wd = *womenDiscrimination*, ind = *individuals*, p = *professionals*, w = *women*, bg = *blackGirls*, and mg = *marginalizedGroups*.

to RA-mediated relations, providing a clearer view of how cohesion emerges from shared inferential structure. This representation emphasizes the structural role of similarity within LAF, highlighting how local inferential connections give rise to cohesion patterns without introducing spurious similarities between unrelated or conflicting nodes.

Conducting a simple analysis, it is evident that only the initial nodes are labeled with a similarity of 0. However, as the labeling procedure progresses, it becomes impossible for any node to have a similarity measure of 0 with its ancestor, given the way the graph is constructed; that is, the cohesion of an argument is always non-zero. Therefore, we have the following proposition:

Proposition 1. *Let $\mathbb{X} \in Arg_\Gamma$ be an argument. If $\mathbb{X}$ is not a leaf node, then $\mu_S^X \neq 0$.*

As observed in the aggregation example, the operation results in a value closer to the upper bound according to the operator definition (Definition 1). This occurs because, in the conceptualization, the similarity between a concept and itself is 1. In other words, the concepts used in the claim are the same as those used in the premises, though their attributes may vary. Moreover, the claim associated with an accrued argument, that is, an argument composed of different reasoning chains, has a higher level of cohesion than its independent ancestors. It is important to note that this result is specific to the similarity measure defined in equation (1) and is not claimed to hold for other similarity measures.

Proposition 2. *Let* $\mathbb{X}, \mathbb{Y}, \mathbb{Z} \in Arg_\Gamma$ *be arguments, where* $\mathbb{Y}$ *and* $\mathbb{Z}$ *are the ancestors of* $\mathbb{X}$. *Then,* $\mu_S^X > \mu_S^Y$ *and* $\mu_S^X > \mu_S^Z$.

Note that we are considering similarity between the concepts that make up an argument, rather than between arguments as atomic entities. This notion of similarity, that is, induced by the ontological concepts underlying each argument, satisfies the similarity function properties described by Amgoud et al.[37]: *maximality*, *symmetry*, and *substitution*. In particular, maximality ensures that each concept is maximally similar to itself, symmetry guarantees that similarity is independent of the order of comparison between concepts, and substitution allows similar concepts to be interchanged without affecting the overall evaluation. These properties are naturally satisfied in our setting, as similarity is derived from ontological relations between concepts rather than from syntactic or contextual features.

According to Budan et al.,[21] this procedure generates a set of mathematical equations that outline the necessary conditions for evaluating all knowledge pieces. Each solution represents a potential set of values that ensures the stable transmission of knowledge attributes within the graph, based on the defined operations.

Next, we present the algorithms for obtaining the valuations associated with each I-node in an argumentation graph. This involves solving two systems of equations using the defined algebras.

Algorithm 1: Labeling procedure for a graph G_Γ

Input: Argumentation graph G_Γ and knowledge base $\mathcal{K}$.
Output: System of equations *EQS-S* for valuations μ_S^X and δ_S^X associated with each I-node X of G_Γ. System of equations *EQS-P* for pair of valuations (μ_{AP}^X, μ_Q^X) and $(\delta_{AP}^X, \delta_Q^X)$ associated with each I-node X of G_Γ.
Initialize *EQS-S* as an empty system of equations;
Initialize *EQS-P* as an empty system of equations;
Initialize each I-node of G_Γ with a *not visited status*;
for *each I-node* X *in* G_Γ *that is not visited* **do**
  EQS-S := *EQS-S* ∪ *LabelingFunction*($X, G_\Gamma, \mathcal{K}$, *EQS-S*);
  EQS-P := *EQS-P* ∪ *LabelingFunction*($X, G_\Gamma, \mathcal{K}$, *EQS-P*);
return (*EQS-S*; *EQS-P*).

Algorithm 1 initializes the system of equations, and Algorithm 2 analyzes each node of the graph to find and assign the valuations of its I-nodes, following the guidelines proposed by Budan et al.[21] Algorithm 2 is invoked by Algorithm 1 in a recursive manner. It incrementally constructs the systems of equations characterizing accrued and weakened valuations by extending previously generated constraints.

Algorithm 2 implements the labeling function defined within Algorithm 1. First, *LabelingFunction*($P, G_\Gamma, \mathcal{K}$, *EQS-S*) is instantiated using the similarity measure; then, *LabelingFunction*($P, G_\Gamma, \mathcal{K}$, *EQS-P*) computes the quality values of the source. Once execution begins, the system progressively generates equations in the following order: first, it computes similarity equations; then it derives the equations related to the proponents' qualities. Finally, Algorithm 2 returns control to Algorithm 1 with the full system of equations.

7 Communities from an AdP-LAF

Building on the concept of coalition as described by Amgoud[38]—understood as a set of supporting arguments—we extend this idea to the realm of argumentative communities. According to Budán et al.,[3] an argumentative community is defined as a coalition of abstract arguments, that is, a group of arguments that collectively support a common position or conclusion.

Furthermore, following Coronel et al.,[39] each argument can be seen as a unit of reasoning that focuses on specific topics, which are naturally interconnected with others to varying degrees. Consequently, there is an underlying notion of distance between arguments when considering their topical proximity. Arguments that support one another and fall within each other's reach can thus be regarded as *friendly* neighbors. When such neighborhoods overlap, the range of mutual support can be extended, allowing collections of arguments to grow beyond isolated clusters and form what we refer to as an argumentative community.

Algorithm 2: Labeling function for nodes in the graph

Input: Argumentation graph G_Γ, I-node X, knowledge base $\mathcal{K}$, and partial systems of equations *EQS-S* and *EQS-P*.

Output: Updated systems of equations *EQS-S* and *EQS-P*, extended with the constraints induced by node X.

Mark the I-node X as visited;
if *X has input from RA-nodes* **then**
  if *X is an element of $\mathcal{K}$* **then**
    $Aggregation_S := \mathcal{F}_1(X)$;
    $Aggregation_P := \mathcal{F}_2(X)$;
  else
    $Aggregation_S := \perp$;
    $Aggregation_{(AP,Q)} := (AP, Q)$, where $Q \neq 0$;
  for *each RA-node R with an incoming edge from X* **do**
    $Support_S := \top$;
    $Support_{(AP,Q)} := (\perp, \top)$;
    for *for each premise P with outgoing edge to an RA-node R* **do**
      if *P is not visited* **then**
        $LabelingFunction(P, G_\Gamma, \mathcal{K}, EQS\text{-}S)$;
        $LabelingFunction(P, G_\Gamma, \mathcal{K}, EQS\text{-}P)$;
      else
        $Support_S := Support_S \odot \delta_S^P$;
        $(Support_{AP}, Support_Q) := (Support_{AP} \circledcirc \delta_{AP}^P, Support_Q \boxdot \delta_Q^P)$;
    $Aggregation_S := Aggregation_S \oplus Support_S$;
    $(Aggregation_{AP}, Aggregation_Q) := (Aggregation_{AP} \circledcirc Support_{AP}, Aggregation_Q \boxplus Support_Q)$;
  Add "$\mu_S^X = Aggregation_S$" to *EQS-S*;
  Add "$(\mu_{AP}^X, \mu_Q^X) = (Aggregation_{AP}, Aggregation_Q)$" to *EQS-P*;
else
  Add "$\mu_S^X = \mathcal{F}_1(X)$" to *EQS-S*;
  Add "$(\mu_{AP}^X, \mu_Q^X) = \mathcal{F}_2(X)$" to *EQS-P*;
if *X has input from CA-nodes* **then**
  if *$\sim X$ is not visited* **then**
    $LabelingFunction(\sim X, G_\Gamma, \mathcal{K})$;
  else
    Add "$\delta_S^X = \mu_s^X \ominus \mu_S^{\sim X}$" and "$\delta_S^{\sim X} = \mu_S^{\sim X} \ominus \mu_S^X$" to *EQS-S*;
    Add "$(\delta_{AP}^X, \delta_Q^X) = (\mu_{AP}^X \Cap \mu_{AP}^{\sim X}, \mu_Q^X \boxminus \mu_Q^{\sim X})$" and "$(\delta_{AP}^{\sim X}, \delta_Q^{\sim X}) = (\mu_{AP}^{\sim X} \Cap \mu_{AP}^X, \mu_Q^{\sim X} \boxminus \mu_Q^X)$" to *EQS-P*;
else
  Add "$\delta_S^X = \mu_S^X$" to *EQS-S*;
  Add "$(\delta_{AP}^X, \delta_Q^X) = (\mu_{AP}^X, \mu_Q^X)$" to *EQS-P*;
return *EQS-S*; **return** *EQS-P*.

Before introducing this broader and dynamic notion of argumentative community, it is important to emphasize that our conceptualization remains grounded in the structure of *argumentum ad populum* and its subtypes. In particular, the interactions, relations of support and disagreement, and forms of influence that characterize argumentative communities in our framework are interpreted through different appeals to popular opinion, such as the majority endorsement, expert authority, social prestige, or shared identity. From this perspective, argumentative communities are not merely clusters of similar arguments, but socially grounded formations whose internal dynamics reflect the specific ad-populum sub-schemes instantiated by their proponents.

With these intuitive ideas in mind, we now introduce a more comprehensive concept of an argumentative community, which is not merely a collection of abstract arguments but also encompasses the interactions and relationships between them. This perspective emphasizes the dynamic nature of argumentation, in which arguments support, oppose, and influence one another over time. By jointly considering argumentative structure and interconnections, we aim to provide a richer account of how argumentative communities arise, evolve, and function.

Definition 10 (Community in AdP-LAF)**.** Let Γ be an instantiated AdP-LAF, and G_Γ be its corresponding argumentation graph. A community in AdP-LAF is a maximal subgraph $\mathbb{X}^*$ of G_Γ without conflict-nodes supporting a sentence X. We denote the set of all communities in Γ as Com_Γ.

As we observe the process of constructing an argument within the AdP-LAF framework, it becomes evident that each argument can be viewed as a community of supporting components. This is illustrated in Figure 21, where the arguments highlighted in blue and green exemplify this structure. In AdP-LAF, an argument is built from a set of premises, evidence, and logical connections that collectively support a specific claim or statement. This assemblage of elements forms a cohesive unit, much like a community, where each part plays a role in reinforcing the overall argument. Therefore, we can conceptualize each argument in AdP-LAF as an independent community dedicated to upholding a particular assertion. This perspective allows us to see how different arguments, or communities, operate independently to support their respective conclusions. By recognizing these argument communities, we can better understand the structure and interplay of various arguments within the AdP-LAF framework.

Proposition 3. *An object $\mathbb{X}$ is an argument in Arg_Γ if and only if it corresponds to a community $\mathbb{X}^* \in Com_\Gamma$ in* AdP-LAF.

In the context of our work, it is crucial to differentiate independent arguments, which we refer to as independent communities. This distinction is important because each independent argument or community operates autonomously, without relying on or being influenced by other arguments or communities. By identifying and analyzing these independent entities, we can better understand the structure and dynamics of the overall argumentative landscape.

Definition 11 (Independent Communities)**.** Let $\mathbb{X}^*, \mathbb{Y}^* \in Com_\Gamma$ be two communities. We say that $\mathbb{X}^*$ and $\mathbb{Y}^*$ are independent iff $\mathbb{X}^* \cap \mathbb{Y}^* = \emptyset$.

Independent communities are self-contained sets of arguments that support their conclusions without needing external justification from other groups of arguments. This independence ensures that the validity and strength of each community can be assessed on its own merits. By recognizing and distinguishing these independent communities, we can achieve a more accurate and nuanced analysis of how different argumentative strategies and perspectives coexist and interact within a broader discourse.

Also, when two communities conflict, the one that weakens the other may still receive support from groups not involved in the dispute. This suggests that even if a community's position is undermined during a conflict, it can continue to garner backing from unaffected groups.

Proposition 4. *Let $\mathbb{X}^*, \mathbb{Y}^* \in Com_\Gamma$ be two communities in Γ related by a CA-Node, where $\delta^X_{AP} \setminus \delta^Y_{AP} \neq \emptyset$, then $\delta^X_Q \neq 0$.*

As per Caminada et al.,[40] the concept of independent argumentative communities is governed by the *principle of non-interference*, assuming a set of atoms, a set of well-formed formulas, and a consequence function. To make the logical setting fully explicit and self-contained, we introduce the basic elements of the underlying propositional language and the auxiliary functions used in the following definitions.

Definition 12 (Propositional Logical Formalism)**.** A propositional logical formalism is a triple $(Atom, Form, Cn)$ such that:

- *Atom* is a finite, non-empty set of propositional symbols.
- *Form* is the smallest set satisfying:
  - Every $a \in Atom$ is in *Form*.
  - If $F \in Form$, then $\sim F \in Form$.
  - If $F, G \in Form$, then $(F \wedge G)$, $(F \vee G)$ and $(F \rightarrow G)$ belong to *Form*.
- *Cn* is a consequence operator $Cn : 2^{Form} \rightarrow 2^{Form}$ satisfying, for all $Z, K \subseteq Form$:
  - $Z \subseteq Cn(Z)$ (Extensivity);
  - If $Z \subseteq K$, then $Cn(Z) \subseteq Cn(K)$ (Monotonicity);
  - $Cn(Cn(Z)) = Cn(Z)$ (Idempotence).

Consider $Atom = \{h, s\}$, where h denotes *"dress code increases decorum"* and s denotes *"dress code causes discrimination."* According to the previous definition, h is a formula, $\sim s$ is a formula, and $h \rightarrow\sim s$ is also a well-formed formula in *Form*. Let $Z = \{h,\ h \rightarrow\sim h\}$. By the definition of the consequence operator Cn, it follows that $\sim h \in Cn(Z)$, since it is derivable from Z by Modus Ponens. Moreover, by extensivity, $Z \subseteq Cn(Z)$.

In what follows, we will refer to the set of propositional symbols occurring in a formula; for instance, the formula $h \rightarrow\sim s$ contains the atoms h and s. This intuition is formally captured by the function $atoms(\cdot)$ defined below.

Definition 13 (Atoms Function)**.** We define the function

$$atoms : Form \rightarrow 2^{Atom}$$

that returns the set of atoms occurring in a formula. It is inductively defined as follows:

- If $F = a$ and $a \in Atom$, then $atoms(F) = \{a\}$.
- If $F = \neg G$, then $atoms(F) = atoms(G)$.
- If $F = (G \circ H)$, where $\circ \in \{\wedge, \vee, \rightarrow\}$, then

$$atoms(F) = atoms(G) \cup atoms(H).$$

Consider the formula $F = h \rightarrow (s \wedge \neg h)$. The propositional symbols occurring in F are precisely h and s, and, therefore, $atoms(F) = \{h, s\}$. This function allows us to identify the syntactic components involved in a set of formulas and to restrict the consequences of a theory to the atoms that actually appear in it. In particular, the notion of syntactic disjointness used in the following definition relies on the idea that two sets of formulas do not share propositional symbols, which can be formally checked through the function $atoms(\cdot)$.

Definition 14 (Non-Interference[40])**.** A propositional logical formalism $(Atom, Form, Cn)$ satisfies the principle of non-interference iff for every $G_1, G_2 \subseteq Form$ such that G_1 and G_2 are syntactically disjoint it holds that $Cn(G_1)_{|atoms(G_1)} = Cn(G_1 \cup G_2)_{|atoms(G_1)}$ and $Cn(G_2)_{|atoms(G_2)} = Cn(G_1 \cup G_2)_{|atoms(G_2)}$.

The principle of non-interference expresses a strong form of independence between syntactically disjoint sets of formulas. It requires that, when two such sets G_1 and G_2 are combined, each set preserves exactly its own logical consequences once reasoning is restricted to the atoms occurring in that set. In other words, the inferential behavior of one set is not affected by the presence of the other, thereby ensuring the modularity of the underlying logical formalism. The above definition uses the notation $G_{|A}$ to denote the subset of formulas in G whose atoms belong to A. Accordingly, non-interference guarantees that, for syntactically disjoint sets of formulas, reasoning within each set remains independent when attention is restricted to its corresponding vocabulary.

Proposition 5. *If two communities $\mathbb{X}^*, \mathbb{Y}^*$ in Γ are independent, then they are non-interfered communities.*

While Definition 14 establishes the non-interference principle in an abstract manner, our framework operationalizes it by structuring argumentative communities within the LAF as disjoint subgraphs. Each subgraph represents a coherent set of arguments that interact according to well-defined inferential rules, without being influenced by external, unrelated argument structures. This transition from an abstract principle to a structured formalism is achieved through the labeling process, where argumentative relations—such as support, conflict, and aggregation—are explicitly encoded within the LAF. As a result, we ensure that conclusions within a given argumentative community emerge solely from its internal logical dependencies, preserving both coherence and independence across communities. Through this instantiation, the non-interference principle becomes a concrete methodological foundation for analyzing structured debates. This ensures the integrity and autonomy of each community's reasoning process.

Proposition 6. AdP-LAF *satisfies the non-interference principle.*

Additionally, according to Caminada et al.,[40] two independent argumentative communities must satisfy the *principle of non-contamination*. This principle is designed to maintain a clear delineation between relevant and irrelevant pieces of knowledge during the process of reaching a conclusion.

Definition 15 (Contamination)**.** Let $(Atom, Formulas, Cn)$ be a logical formalism. A set $G_1 \subseteq Formulas$, with $atoms(G_1) \subsetneqq Atoms$, is called contaminating iff for every $G_2 \subseteq Formulas$ such that G_1 and G_2 are syntactically disjoint it holds that $Cn(G_1) = Cn(G_1 \cup G_2)$.

Intuitively, contamination captures the idea that a set of formulas G_1 is self-sufficient with respect to logical consequence. Adding formulas that are syntactically disjoint from G_1 does not yield any new consequences, meaning that the inferential behavior of G_1 remains unaffected by unrelated information. In this sense, a contaminating set behaves as if it were logically isolated from the remainder of the language.

Conversely, the absence of contamination ensures that only information that is logically relevant to a given set contributes to its conclusions, while syntactically unrelated data is systematically disregarded. By adhering to this principle, independent argumentative communities are able to focus exclusively on the information that is pertinent to the formation of their conclusions. This selective treatment of information enhances both the accuracy and the robustness of the reasoning process, preventing the distortion of outcomes by extraneous factors.

Proposition 7. *$\mathbb{X}^*$ and $\mathbb{Y}^*$ satisfy the non-contamination principle.*

As before, Definition 15 establishes the concept of contamination in an abstract manner and our framework operationalizes it by structuring argumentative communities within the LAF to prevent the uncontrolled propagation of irrelevant information. This is achieved by defining formal constraints on how arguments within a community interact and ensuring that only relevant information influences the reasoning process. Through the labeling mechanism of AdP-LAF, we explicitly encode dependencies and argumentative structures to filter out information that does not contribute meaningfully to the derivation of conclusions. By moving to a structured formalism, we maintain the integrity of each community, ensuring that their conclusions are not distorted by unrelated or redundant data. By enforcing non-contamination, our framework enhances the precision and robustness of argument evaluation within structured debates.

Now, once the community is defined, it is of interest to determine its level of cohesion, that is, how each argument contributes to the community's general opinion. If most sub-arguments refer in a similar way to the topics in a conversation, the community will be considered to be cohesive. In this direction, we know that for a given argument $\mathbb{X}$ that supports a statement X, its δ_s^X value represents the aggregation of partial coherence of the reasons supporting the claim X overcoming the conflict (if a conflict between such argument and its contradiction exists), so it is natural to think that the cohesion of an argument can be determined through this label:

Definition 16 (Community Cohesion)**.** Let $\mathbb{X}^* \in Com_\Gamma$ be a community of Γ supporting a statement X that is labeled with $(\mu_S^X, (\mu_{AP}^X, \mu_Q^X))$ and $(\delta_S^X, (\delta_{AP}^X, \delta_Q^X))$. The cohesion associated with $\mathbb{X}^*$ is defined as $\mathsf{Coh}(\mathbb{X}^*) = \delta_S^X$.

We will now use the elements provided to evaluate the acceptability of communities. As mentioned earlier, ad-populum AS Walton[7–9] (based on popular opinion) are useful for analyzing trends and identifying communities in complex debates. These schemes look at arguments and recognize their provisional nature. We will adapt these schemes into a more computable format to help classify communities. First, we will outline the general scheme pattern to define a defeated community based on the presented concepts.

Considering both Definition 7, which states that when an argument is rejected, the state of the claim after taking conflicts into account is $\bot$, and Definition 16, we have the necessary elements to define when a community is rejected.

Definition 17 (Defeated Community)**.** Let $\mathbb{X}^*, \mathbb{Y}^* \in Com_\Gamma$ be two communities with contradictory claims X and Y, respectively, where (μ_S^X, δ_S^X) and (μ_S^Y, δ_S^Y) are the similarity labels associated with X and Y, and $((\mu_{AP}^X, \mu_Q^X), (\delta_{AP}^X, \delta_Q^X))$ and $((\mu_{AP}^Y, \mu_Q^Y), (\delta_{AP}^Y, \delta_Q^Y))$ are the proponent's group and quality label associated with X and Y, respectively. Then:

Original Pattern	Formalized Pattern
Everybody in a reference group g accepts that X is true, therefore you should accept X.	*If* $\bot < \mu_S^X \leq \top$, *and* $\bot < \mu_Q^X < \top$, then X should be accepted.
Everybody in a reference group g rejects X.	*If* $\mu_S^X < \mu_S^Y$ *or* $(\mu_S^X = \mu_S^Y$, *with* $\mu_Q^X < \mu_Q^Y$ *when* $\mu_{AP}^X = \mu_{AP}^Y)$, then X should be rejected.
X is false, you should reject X.	*Therefore,* X *is a rejected community with* $\mathsf{Coh}(\mathbb{X}^*) = \bot$.

It is important to clarify that in the formalized pattern presented above the expressions "*everybody*" and "*majority*" should be not interpreted as referring to explicit cardinality constraints over a population. Instead, the proposed pattern operates over aggregated acceptance degrees computed within an AdP-LAF. In this context, the similarity label μ associated with a community captures the overall strength of alignments among the arguments supporting a claim, rather than the number of individuals endorsing it. Thus, the value $\top$ is understood as a contextual parameter that encodes when aggregated support is considered sufficient under the adopted scheme.

Furthermore, it is possible to express the critical questions[9] about an argument X more precisely as follows:

Original CQ	Reformulated CQ
CQ1: Does the majority of the reference group accept the argument as true?	*Is* $\mu_s^X > \perp$*?*
CQ2: Is there other relevant evidence available that would support the assumption that the argument is not true?	*Does there exist Y such that* $\mathrm{Coh}(\mathbb{X}^*) < \mathrm{Coh}(\mathbb{Y}^*)$ *and* $\perp < \delta_q^Y$*, and does there exist a CA-node between X and Y?*
CQ3: What reason is there for thinking that the view of this large majority is likely to be right?	Is $\mathrm{Coh}(\mathbb{Y}^*) > \mathrm{Coh}(\mathbb{X}^*)$?

It is important to note that μ_s^X represents the cohesion value of argument X, and therefore it naturally helps us model CQ1. Furthermore, for CQ2 we consider the Coh value since the question implies an existing conflict between two communities. With the aim of clarifying CQ2, we point out that the role of the opposing community Y is not limited to establishing relevance. Within our formalism, defeat arises from a comparative evaluation between two communities X and Y performed under the same similarity and coherence computation. The rejection of X follows when Y achieves a higher aggregate acceptance or coherence degree according to the ordering induced by the algebra. Finally, with respect to CQ3, the relationship between providing reasons and the coherence value is given by the labeling mechanism of AdP-LAF. Reasons contribute to the structure and labeling of arguments within each community, and the final cohesion value is obtained through the aggregation of these labels.

Furthermore, our formalism allows us to express the variant of the general *ad-populum AS* from which we obtain *presumptive arguments* or, equivalently according to Proposition 3, *believable communities*:

Definition 18 (Believable Community)**.** Let $\mathbb{X}^* \in Com_\Gamma$ be a community, where (μ_S^X, δ_S^X) is the similarity label associated with X, and $((\mu_{AP}^X, \mu_Q^X), (\delta_{AP}^X, \delta_Q^X))$ is the proponent's group and quality label associated with X, respectively. Then:

Original Pattern	Formalized Pattern
There exists a presumption in favor of X.	There exist $\mathrm{Coh}(\mathbb{X}^*)$ and μ_S^X such that $\perp < \mathrm{Coh}(\mathbb{X}^*) < \mu_S^X$.
A large majority in a group g accepts X as true.	There exists at least one pair $(\delta_{AP}^X, \delta_Q^X)$ s.t. $\perp < \delta_{AP}^X$ and $\perp < \delta_Q^X$
Therefore, there exists a presumption in favor of X.	Therefore, $\mathbb{X}^*$ represents a believable community

A believable community can be understood as a feasible argument, *i.e.*, *if a viewpoint gathers widespread support, it may appear to hold some truth*. However, this scheme can be fallacious if it is used as the only basis for argument, since the popularity of a belief does not guarantee its veracity. Walton[7,9,10] points out that for an ad-populum argument to be valid, certain requirements must be met, such as that the majority of the population has access to accurate information and that popular beliefs are based on a solid and well-founded understanding of the facts. It is because of this that the *ad-populum AS* has several particular instantiations where the peculiarity lies in the argument source; for example, people in a special *position to know* about the main topic of the debate. Under this approach based on AdP-LAF, if the graph G_Γ is labeled with a single type of label from the AP-Algebra, for example, the label "PtoK," the *feasible popular opinion* would be obtained from the Position-to-Know Ad-Populum AS. In a real debate, different types of proponents talk about a specific matter. Thus, the AdP-LAF is useful to characterize the whole group of proponents that supports a certain position, in a precise way, according to their qualifications:

Definition 19 (Groups Advocating From Opinions)**.** Let $\mathbb{X} \in Arg_\Gamma$ be an argument labeled with (μ_S^X, δ_S^X) as the similarity label associated with X, and $((\mu_{AP}^X, \mu_Q^X), (\delta_{AP}^X, \delta_Q^X))$ as the compound label representing the proponents' group associated with X and their quality label, respectively. If $\perp < \delta_s^X < \mu_s^X$ and there exists $(\delta_{AP}^X, \delta_Q^X)$ s.t. $\perp < \delta_{AP}^X$ and $\perp < \delta_Q^X$, then X is a believable community supported by each group δ_{AP}^X such that $\perp < \delta_Q^X$.

We have defined *believable* and *defeated* communities as outcomes resulting from the application of the variant of ad-populum AS that provides these kinds of arguments. From the original patterns selected from this work, it is only possible to derive an accepted argument from the Expert Opinion variant. However, that is an unrealistic approach, since an accepted argument can be obtained from non-expert argumentative groups, although it may be subject to scrutiny. So, next we introduce a framework to derive an accepted argument, that is, an accepted community:

Definition 20 (Accepted Community)**.** Let $\mathbb{X}^* \in Com_\Gamma$ be a community. Then, $\mathbb{X}^*$ is an accepted community iff $\delta_S^X = \mu_S^X$ and there exists at least one group supporting $\mathbb{X}^*$, *i.e.*, there exists $(\delta_{AP}^X, \delta_Q^X)$ s.t. $\perp < \delta_{AP}^X$ and $\perp < \delta_Q^X$.

An accepted community is one whose force remains unaltered after the conflict, and it is supported by at least one group.

In the proposed system, the concept of argumentative communities plays a crucial role in structuring and evaluating reasoning processes. The following propositions refine and enhance this framework by ensuring coherence, stability, and logical independence among accepted communities. First, two accepted communities represent two different lines of reasoning. This result is natural due to the mechanism followed to construct the argumentation graph and is further supported by Caminada's principle of non-contamination[40]:

Proposition 8. *If $\mathbb{X}^*$ and $\mathbb{Y}^*$ are two accepted communities, each understood as a set of arguments sharing the same ad-populum reference group, then $\mathbb{X}^* \cup \mathbb{Y}^*$ satisfies the non-contamination principle.*

In this sense, Proposition 8 strengthens the non-contamination principle by asserting that the union of two accepted communities remains free from conflict. For example, consider two argumentative communities discussing dress codes. One supports dress codes in schools, arguing that they promote discipline and a professional learning environment. The other defends dress codes in hospitals, stating that they contribute to hygiene and safety in medical settings. Since each of these arguments is independently valid, their union (dress codes both promote discipline in education and ensure hygiene in healthcare) does not introduce contradictions, ensuring a logically coherent combination of arguments.

As a consequence of the previous statement, the information presented by two accepted arguments can be amalgamated while retaining their acceptability status.

Proposition 9. *If $\mathbb{X}^*$ and $\mathbb{Y}^*$ are two accepted communities, each understood as a set of arguments sharing the same ad-populum reference group, then $\mathbb{X}^* \cup \mathbb{Y}^*$ is an accepted community.*

Proposition 9 extends the previous proposition by demonstrating that the union of two accepted communities also forms an accepted community. That is, by continuing the analysis of dress codes in schools and hospitals, we see that beyond merely avoiding contradictions, the union of these two communities constituted an accepted argumentative community. In order words, the broader claim that "dress codes contribute to order, discipline, and safety in structured environments" emerges as a logically sound and accepted conclusion, encompassing the perspectives of both communities.

Finally, when two communities are linked by a CA-node, it cannot happen that both are accepted simultaneously, which is a natural result derived from Definition 5.

Proposition 10. *If $\mathbb{X}^*$ and $\mathbb{Y}^*$ are two accepted communities, then they cannot be linked by a CA-node.*

Under Proposition 10, we introduce a crucial restriction by establishing that two simultaneously accepted communities cannot be linked by a CA-node (a node representing conflict). To complete the analysis of our example, imagine two argumentative communities discussing the role of dress codes. One community argues that dress codes in schools are essential for discipline and professionalism. Another, however, claims that dress codes impose unnecessary restrictions and limit individual expression. If both were accepted but directly contradicted each other, a CA-node would highlight the conflict, preventing both from being simultaneously valid. This proposition ensures that only logically compatible communities remain accepted.

Together, these properties enforce the logical consistency of the system, ensuring that argumentation remains structured, scalable, and resistant to contamination or interference. In the following section, we detail the procedure used to identify argumentative communities in a debate and characterize them using the elements provided in our formalism.

8 Illustrative application: Homework debate

Beyond illustrating the construction of contextualized semantic graphs and knowledge bases, this use case highlights the role of *argumentum ad populum* as a key mechanism for modeling the dynamics of argumentative communities. In particular, the debate on homework and large language models makes visible how arguments progressively align around different reference groups, giving rise to evolving patterns of agreement, disagreement, and polarization that are common in ad populum reasoning.

We begin by analyzing a debate on homework (https://www.procon.org/headlines/homework-pros-cons-procon-org/), in which three opinions are presented, both in favor of and against (The complete application, together with usage instructions, is publicly available at http://167.99.145.74/ and https://goo.su/r4aC3). To streamline the process and clarify the purpose of each step, we leverage four large language models (LLMs) as automated tools to assist in tasks that would be otherwise time consuming. Specifically, we use LLMs to (semi-)automatically build the corresponding argumentative graph, for example, building part of the conceptualization and extracting arguments from the debate text. In particular, we selected the following LLMs through the Poe Platform (https://poe.com): ChatGPT 4.0; Claude 3.7 Sonnet; Gemini 2.0 Flash; and Meta AI LLaMA 2-70B model, always in their free versions. All LLMs used in the argument mining experiments were employed in their off-the-shelf configurations, without task-specific fine-tuning. These choices were made to ensure comparability across models and to focus the evaluation on their suitability for structured argument extraction, rather than on performance gains induced by model-specific optimizations. We acknowledge, however, that the exact quantization scheme applied by the Poe platform to the LLaMA 2-70B model was not exposed to end users, and, therefore, we cannot report it with certainty: the model was accessed through its hosted configuration, without any user-side quantization or decoding control beyond the platform defaults. We further note that LLaMA 2-70B has since been retired from the Poe platform, which limits the exact reproducibility of this particular run. This does not affect the validity of our contribution, since the LLMs are employed only as auxiliary tools for concept and argument extraction, and the proposed framework is model-agnostic. To support reproducibility, the complete set of extracted concepts, if-then rules, and arguments produced in our experiments is reported in Appendix B, so that the argumentation-theoretic results can be reproduced independently of the specific model or platform used for extraction. For independent re-execution, the same extraction prompts can also be applied to an openly available deployment of the model (e.g., `llama2:70b` served through Ollama, whose default quantization is `q4_0`), allowing the extraction stage to be reproduced under a fully specified and controllable configuration. In what follows, we detail the development of each stage.

8.1 Stage 1: From debate to conceptualizations

In this first stage, we need to find the concepts that are present in the debate and their corresponding attributes to construct the conceptualization and the contextualized semantic graph, according to Definition 8. This implies the following tasks:

1. Select the specific structured debate that will be input to the chosen LLMs for analysis. We noticed that it is important to select debates where the pro and con opinions are clearly differentiated, and where the proponent's group affiliation can be identified. In this case study, the selection has been done manually.
2. Use the LLMs to extract the concepts present in each opinion, as well as their attributes. To do it, we used the prompt "*Could you find keywords or the main concepts and their attributes in the following text?*" This prompt specifies, in a very simple manner, the desired outcome and the context in which the search is conducted. For a few-shot approach, we provided the LLM with examples of how the output should be structured. In this case, we gave each LLM two opinions related to the motivational example (Section 2). We obtained the results provided in Table 3.
3. Explore the responses generated by each LLM and compare them with the desired answers, which are provided by the expert knowledge engineer. The number of knowledge engineers responsible for providing the key concepts and their attributes from a structured debate vary with debate complexity. In this case, we relied on three experts. In this initial stage, we prioritize obtaining the necessary elements to construct the conceptualization using the simplest possible method. With this in mind, the outputs from the LLMs are evaluated based on precision, recall, and F1-score.[41,42]

Table 3. Metrics characterizing LLM performance in concept and attribute extraction, with the best obtained values are highlighted in bold text.

Input: Pro-Op 1	ChatGPT	Claude	Gemini	Meta	KE
# of keywords found	16	11	14	18	11
# of coincidences LLM and KE	10	6	9	10	–
Precision	0.625	0.545	**0.643**	0.556	–
Recall	**0.909**	0.545	0.818	**0.909**	–
F1-Score	**0.740**	0.545	0.719	0.689	–
Input: Pro-Op 2	ChatGPT	Claude	Gemini	Meta	KE
# of keywords found	25	23	36	15	15
# of coincidences LLM and KE	10	10	11	8	–
Precision	0.4	0.435	0.306	**0.533**	–
Recall	0.667	0.667	**0.733**	0.533	–
F1-Score	0.5	0.528	0.429	**0.533**	–
Input: Pro-Op 3	ChatGPT	Claude	Gemini	Meta	KE
# of keywords found	15	19	17	14	8
# of coincidences LLM and KE	4	7	8	4	–
Precision	0.267	0.368	**0.471**	0.286	–
Recall	0.5	0.875	**1**	0.5	–
F1-Score	0.347	0.519	**0.641**	0.364	–
Input: Con-Op 1	ChatGPT	Claude	Gemini	Meta	KE
# of keywords found	16	16	13	18	12
# of coincidences LLM and KE	10	12	12	12	–
Precision	0.625	0.75	**0.923**	0.667	–
Recall	0.833	**1**	**1**	**1**	–
F1-Score	0.714	0.857	**0.960**	0.8	–
Input: Con-Op 2	ChatGPT	Claude	Gemini	Meta	KE
# of keywords found	14	14	15	12	11
# of coincidences LLM and KE	7	6	8	7	–
Precision	0.50	0.429	0.533	**0.5833**	–
Recall	0.636	0.545	0.727	**0.777**	–
F1-Score	0.56	0.48	0.616	**0.666**	–
Input: Con-Op 3	ChatGPT	Claude	Gemini	Meta	KE
# of keywords found	9	9	9	8	11
# of coincidences LLM and KE	9	9	8	7	–
Precision	**1**	**1**	0.89	0.875	–
Recall	**0.83**	**0.83**	0.73	0.636	–
F1-Score	**0.91**	**0.91**	0.80	0.738	–

LLM = large language model; KE = knowledge engineer.

4. If the answers are useful according to the expert knowledge engineers' assessment, that is, the LLM provides precisely the set of concepts and their attributes mentioned in the debate, we create the conceptualization using them. If not, refine the prompt and repeat the directives 2 to 4 until appropriate results are obtained. In our experiment, the model did not use additional prompts. Nevertheless, we were able to obtain the necessary elements for manually constructing the conceptualization in significantly less time than if we had performed the entire process without assistance. As shown in Table 3, LLMs are relatively consistent in identifying keywords, though they exhibit variations in precision, recall, and F1-score. This suggests that while LLMs can play a significant role in structured debate mining by extracting concepts and attributes, they cannot fully replace human knowledge engineers due to limitations in precision and contextual understanding. The most effective approach appears to be a hybrid model, where LLMs provide an initial extraction that is subsequently verified and refined by engineers.

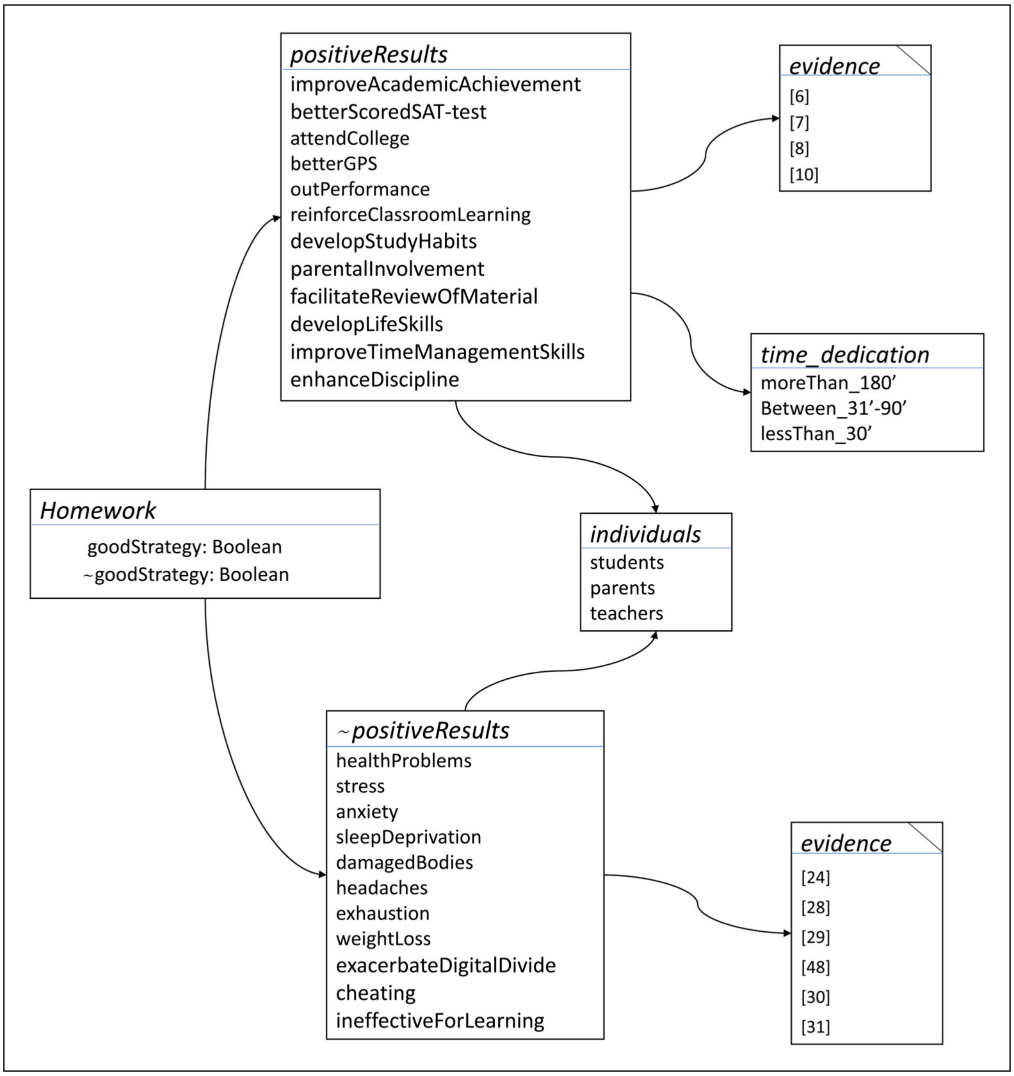

Figure 24. Conceptualization for the debate about homework.

5. Construct the graphs (the general conceptualization and the contextualized semantic graph derived from the first) using the concepts and attributes obtained as the result of applying the previous steps. Given that the case study is simple, the graphs were constructing manually. However, automatic tools may also be used for this purpose.

The analysis conducted in Step 4 shows that none of the tools explored perfectly meet our representational requirements. However, by following the aforementioned hybrid approach, we can create the conceptualization depicted in Figure 24, saving time while ensuring consistent results with the manual procedure. From this conceptualization, we obtain the contextualized semantic graph shown in Figure 25.

8.2 Stage 2: Mining arguments and representing them as if-then rules

Mining arguments may be a complex task if performed manually, requiring considerable time to read, interpret, and extract arguments from texts originating from diverse or potentially complex domains. A study of the state-of-the-art in argument mining is outside of the scope of this article; however, several tools exist that can (semi)-automatically be used to perform this task (see Section 9.5 for related work in this area). For our use case, we focus on obtaining arguments expressed as *if-then rules* that can easily be translated into an AdP-LAF. To do so, we decided to explore the possibility of using LLMs. We used the same LLMs from the previous stage and we employed the following prompt: Suppose you are a knowledge engineer that is mining a debate. *Could you give me the arguments in [text] expressed as if-then rules? And*

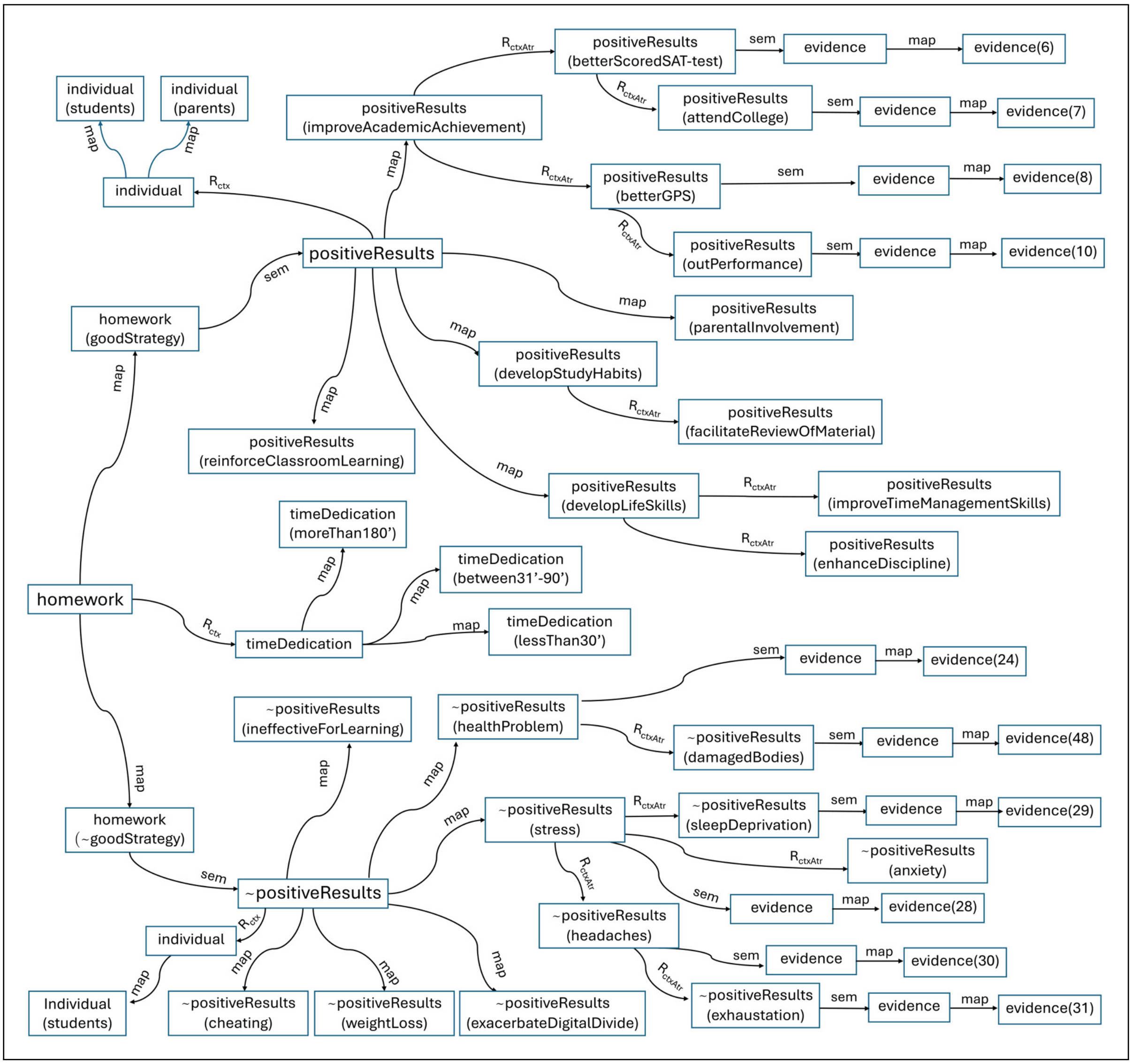

Figure 25. Contextualized semantic graph for the homework debate.

then, Could you find the premises and conclusion for each argument? For example, considering the Pro-Opinion 1 from the Homework debate, Figure 26 depicts the responses provided by ChatGPT. The complete set of mined arguments is available in Supplemental Appendix B.

We selected *ChatGPT* as the primary engine for argument mining based on both qualitative inspection and comparative experimental analysis. In particular, ChatGPT-generated outputs consistently identify premises as explicit reasons supporting a conclusion, making the aggregation process transparent and closely aligned with the representations produced by a human knowledge engineer. This property is essential for the subsequent construction of the argumentative graph within the LAF, where the explicit availability of premises directly affects argument activation and support relations.

In contrast, Gemini 2.0 Flash tends to generate concise responses that collapse multiple reasons into a single statement, often mirroring the corresponding if-then rule without exposing the internal argumentative structure. As a result, although Gemini exhibits high precision, it frequently under-represents the full set of arguments involved. Meta AI, on the other hand, fails to consistently identify all relevant arguments, while Claude 3.7 Sonnet detects a larger number of candidate rules but does not maintain a one-to-one correspondence between rules and arguments, leading to structural inconsistencies in the extracted knowledge base.

These qualitative observations are supported by the comparative evaluation reported by Debórtli.[43] Across the tested datasets, ChatGPT (GPT-4) exhibits a more balanced trade-off between semantic coverage and structural formalization,

Argument 1 (Main Claim): *Homework improves student achievement.*
If-Then Rule:
IF students do homework, **THEN** their academic achievement improves.
Premises and Conclusion:
Premise 1: Studies have shown that homework improves student achievement in terms of grades, test results, and the likelihood of attending college.
Premise 2: Research published in the *High School Journal* found that students who spent between 31 and 90 minutes per day on homework scored about 40 points higher on the SAT-Mathematics subtest than peers who did no homework.
Premise 3: Students in classes that were assigned homework outperformed 69% of students without homework on both standardized tests and grades.
Premise 4: Meta-studies have shown that take-home assignments were effective at improving academic achievement in 64% and 72% of cases, respectively.
Premise 5: Research by the Institute for the Study of Labor (IZA) concluded that increased homework led to better GPAs and a higher probability of college attendance for high school boys.
Premise 6: Boys who attended college did more than three additional hours of homework per week in high school.
Conclusion: Homework improves student achievement.
Argument 2 (Supporting Claim 1): ***Homework improves standardized test scores.***
If-Then Rule:
IF students spend time on homework, **THEN** they perform better on standardized tests.
Premises and Conclusion:
Premise 1: Research from the *High School Journal* found that students who did 31–90 minutes of homework daily scored about 40 points higher on the SAT-Mathematics subtest than those who did none.
Premise 2: A meta-study found that students in classes with homework outperformed 69% of students without homework on standardized tests.
Conclusion: Homework improves standardized test scores.
Argument 3 (Supporting Claim 2): ***Homework increases the likelihood of attending college.***
If-Then Rule:
IF students complete more homework, **THEN** they are more likely to attend college.
Premises and Conclusion:
Premise 1: Research by the Institute for the Study of Labor (IZA) found that increased homework led to higher GPAs.
Premise 2: Higher GPAs are associated with a higher probability of college attendance.
Premise 3: Boys who attended college did more than three additional hours of homework per week in high school.
Conclusion: Homework increases the likelihood of attending college.

Figure 26. Example of if-then rules and their derived arguments provided by ChatGPT 4.

achieving higher stability in terms of precision, recall, and F1-score. While Gemini favors precision at the expense of recall, and Claude favors recall at the expense of structural consistency, ChatGPT maintains a stable extraction behavior that preserves the intended meaning of colloquial arguments while conforming to the formal requirements of if-then rule generation. Importantly, the proposed framework remains model-agnostic: ChatGPT is used as a concrete and reproducible instantiation of the argument mining stage, rather than as a dependency of the LAF itself.

8.3 Stage 3: Constructing the argumentation graph and analyzing the results

Although we are working on developing a tool capable of automatically constructing the argumentation graph from a set of if-then rules, in this work the graph was manually constructed based on the following process:

1. Construct the corresponding knowledge base from the obtained if-then rules.
2. Consider two (arbitrary) opposing opinions about the matter in question.
3. Develop the subgraph for a "pro" opinion and label it according to the procedure established in Definition 5.
4. Build the graph for a "con" opinion and label it.
5. Repeat Steps 2 to 4 until all distinct and relevant opinions expressed in the debate have been identified and modeled.

Each iteration of Steps 2 to 4 is conditioned on the identification of a distinct argumentative stance within the debate. While the prompt structure used for argument mining remains fixed across iterations, the contextual focus is deliberately shifted to reflect a different position (e.g., pro or con) with respect to the issue under discussion.

Importantly, deterministic inference settings are employed throughout this process, and no stochastic sampling is used to induce variability in the extracted arguments. Consequently, previously identified keywords or argumentative elements are not explicitly accumulated or injected into subsequent prompts. Instead, argumentative diversity emerges from the plurality of stances present in the debate itself, rather than from random sampling effects. This design choice ensures that the resulting argumentation graph reflects structural differences between positions, as opposed to variations introduced by model randomness.

The number of iterations is therefore not fixed a priori, but depends on the structure of the debate under analysis. Each iteration corresponds to the identification and modeling of a distinct argumentative stance with respect to the topic. In debates characterized by low diversity or weak polarization, only a limited number of iterations may be required, as arguments tend to converge around a small set of positions. Conversely, in highly polarized or heterogeneous debates, additional iterations are necessary to capture the variety of opposing viewpoints, reference groups, and argumentative strategies involved. The iteration process continues until all relevant stances present in the debate have been represented in the argumentation graph.

The knowledge base for the debate, depicted in Figure 27 was manually developed by a single expert knowledge engineer following a well-defined and formalized procedure. Although we explored the use of LLMs to automatically generate this knowledge base, the resulting representations proved inadequate for the requirements of the proposed formalism. To facilitate the construction of the labeled argumentation graph shown in Figure 28, each fact in the knowledge base is identified as f_{number} and each rule as r_{number}. The argumentation graph was manually constructed, as any fully automatic tool capable of producing it would require inputs derived from the contextualized semantic graph, and to the best of our knowledge, no such tools are currently available. Since the annotation and modeling process was carried out by a single expert knowledge engineer, inter-annotator agreement statistics are not applicable in this setting. The focus of this work is, therefore, placed on the formal argumentation framework and its valuation mechanisms, rather than on the empirical assessment of annotation consistency.

From an *ad-populum* perspective, the full argumentative map shown in Figure 28 makes explicit how community-level polarization emerges from the aggregation of appeals to popular opinion. Arguments cluster around distinct reference groups—such as pedagogical authorities, student-centered perspectives, or technologically optimistic positions—corresponding to different *ad-populum* subtypes. This structure illustrates how group-based endorsement and source credibility jointly shape the dynamics of polarization within the debate.

The numerical valuations reported in Figure 28 provide a quantitative characterization of this phenomenon. These values originate from the knowledge extraction and source characterization stage of the proposed pipeline, before the application of the argumentation algebras. At this stage, each piece of knowledge is assigned an initial source-quality valuation together with an *ad-populum* label, based on observable characteristics of its proponent, such as expertise, institutional role, or social relevance within the debate context.

Initial valuations are obtained through a human-assisted annotation process, supported by structured debate metadata and, when available, LLMs used to identify the role, standing, and credibility of the proponent. Importantly, the assignment of these initial values is external to the argumentation formalism: the ad-populum LAF does not assume a specific mechanism for generating them, but instead provides a principled algebraic machinery for propagating and contrasting such valuations.

Once assigned, initial values are propagated through the argumentation graph using the support, aggregation, and conflict operators defined in the S- and Q-algebras. Accordingly, the numerical values shown in Figure 28 correspond to the final δ-valuations obtained after all argumentative interactions have been taken into account, while their epistemic origin lies in the prior source characterization stage.

Following the example, we can see two clear positions about homework: one highlighting its positive results, and the other indicating exactly the opposite. The two mentioned postures or arguments represent two discursive communities under our approach, that is, each maximal set of arguments that can characterize the stances in the debate by providing reasons in favor (or against) the benefits of doing homework, according to the sources that support their points of view. Now, considering the labels obtained through the labeling procedure, the argument indicating a negative result of doing homework represents a defeated community in this debate, whereas the opposite is an accepted community. This distinction arises because, although both arguments exhibit the highest level of coherence, the values associated with the shared groups supporting each opinion enable us to assess which argument is stronger. Suppose, however, that in another debate we are faced with an identical situation regarding coherence, but the groups supporting each opinion are entirely disjoint. In such a case, it would not be possible to determine which argument is a stronger, and both would be regarded as defeated.

Overall, this use case illustrates how the proposed formalism not only captures static argumentative structures, but also supports the analysis of dynamic phenomena, such as group alignment and polarization, through the lens of argumentum ad populum. This reinforces the suitability of ad populum as a conceptual bridge between natural argumentation patterns and formal graph-based models of argumentative communities.

r_1: homework(goodStrategy) ≺ positiveResults(improveAcademicAchievement), timeDedication(moreThan_180’) individuals(students) : {0, (PtoK, 0.68)}
r_2: homework(goodStrategy) ≺ positiveResults(attendCollege), positiveResults(betterScoredSAT-test), positiveResults(betterGPS), individuals(students) : {0, (AD, 0. 5)}
r_3: homework(goodStrategy) ≺ positiveResults(reinforceClassroomLearning), positiveResults(facilitateReviewOfMaterial), individuals(students) : {0, (EX, 0.45)}
r_4: homework(goodStrategy) ≺ positiveResults(developStudyHabits), individuals(students) : {0, (AD, 0.84), (PtoK, 0.2)}
r_5: homework(goodStrategy) ≺ positiveResults(developLifeSkills), positiveResults(improveTimeManagementSkills), individuals(students) : {0, (AD, 0.66)}
r_6: homework(goodStrategy) ≺ positiveResults(parentalInvolvement), individuals(parents), individuals(teachers) : {0, (EX, 0.52)}
r_7: homework(goodStrategy) ≺ positiveResults(betterGPAs), individuals(students) : {0, (AD, 0.73)}
r_8: homework(~goodStrategy) ≺ ~positiveResults(stress), ~positiveResults(anxiety), ~positiveResults(damagedBodies) individuals(students) : {0, (PR, 0.58)}
r_9: homework(~goodStrategy) ≺ ~positiveResults(healthProblems), ~positiveResults(headaches) , ~positiveResults(weightLoss), ~positiveResults(exhaustation), individuals(students) : {0, {(AD, 0.73), (PR, 0.24)}}
r_{10}: homework(~goodStrategy) ≺ ~positiveResults(sleepDeprivation), individuals(students) : {0, (PtoK, 0.65), (AD,0.15)}
r_{11}: homework(~goodStrategy) ≺ ~positiveResults(exacerbateDigitalDivide), individuals(students) : {0, (EX, 0.52)}
r_{12}: ~positiveResults(ineffectiveForLearning) ≺ timeDedication(Between_31’-90) , individuals(students) : {0, (AD, 0.32)}
r_{13}: positiveResults(attendCollege)≺ timeDedication(moreThan_180'), individuals(students) : {0, {(PtoK, 0. 08), (EX, 0.4)}
r_{14}: homework(~goodStrategy) ≺ ~ positiveResults(cheating), individuals(students) : {0, (PtoK, 0.47)}
r_{15}: homework(~goodStrategy) ≺ ~positiveResults(ineffectiveForLearning), individuals(students) : {0, (PR, 0.8)}
r_{16}: homework(~goodStrategy) ≺ positiveResults(ineffectiveForLearning), individuals(students) : {0, (PR, 0.8)}

f_1: homework(goodStrategy) : {0, (PR, 0.13)}
f_2: homework(~goodStrategy) : {0, (PR, 0.44)}
f_3: individuals(students) : {0, (EX, 0.87)}
f_4: individuals(parents) : {0, (AD, 0. 28)}
f_5: individuals(teachers) : {0, (PtoK, 0.27)}
f_6: timeDedication(moreThan_180’) : {0, (AD, 0.18)}
f_7: timeDedication(between 31’-90’) : {0, (AD, 0. 47)}
f_8: ~positiveResults(cheating) : {0, (PtoK, 0. 31)}
f_9: positiveResults(improveAcademicAchievement) : {0, (PtoK, 0. 88)}
f_{10}: positiveResults(betterScoredSAT-test) : {0, (PtoK, 0. 25)}
f_{11}: positiveResults(attendCollege) : {0, (AD, 0. 68)}
f_{12}: positiveResults(betterGPS) : {0, (PtoK, 0. 8)}
f_{13}: positiveResults(outPerformance) : {0, (PtoK, 0. 5)}
f_{14} : positiveResults(enhanceDiscipline) : {0, (EX, 0. 3)}
f_{15}: positiveResults(reinforceClasroomLearning) : {0, (PtoK, 0. 69)}
f_{16}: positiveResults(developStudyHabits) : {0, (EX, 0. 47)}
f_{17}: positiveResults(developLifeSkills) : {0, (AD, 0.44)}
f_{18}: positiveResults(parentalInvolment) : {0, (EX, 0. 52)}
f_{19}: positiveResults(betterGPS) : {0, (AD, 0. 73)}
f_{20}: positiveResults(facilitateReviewOfMaterial) : {0, (AD, 0. 29)}
f_{21}: ~positiveResults(ineffectiveForLearning) : {0, (PR, 0. 62)}
f_{23}: ~positiveResults(stress) : {0, (PR, 0. 47)}
f_{24}:~positiveResults(healthProblems) : {0, (AD, 0. 22)}
f_{25}: ~positiveResults(anxiety) : {0, (PR, 0. 54)}
f_{26}:~positiveResults(sleepDeprivation) : {0, (PtoK, 0. 65)}
f_{27}: ~positiveResults(damagedBodies) : {0, (AD, 0. 09)}
f_{28}:~positiveResults(headaches) : {0, (AD, 0. 28)}
f_{29}: ~positiveResults(exhaustation) : {0, (PtoK, 0. 57)}
f_{30}:~positiveResults(weightLoss) : {0, (PtoK, 0. 2)}
f_{31}: ~positiveResults(exacerbateDigitalDivide) : {0, (EX, 0. 71)}
f_{32}: positiveResults(improveTimeManagementSkills) : {0, (EX, 0.15)}

Figure 27. Knowledge base for homework debate.

9 Related work

As this article addresses various interconnected topics, we now discuss the latest advancements in such areas in relation to our work.

9.1 Community detection from graph structures

Among the relevant related work, we will mention Zhao et al.’s research[44] in which the authors proposed an algorithm to find communities in large-scale social networks by compressing the graph and using density and quality measures to

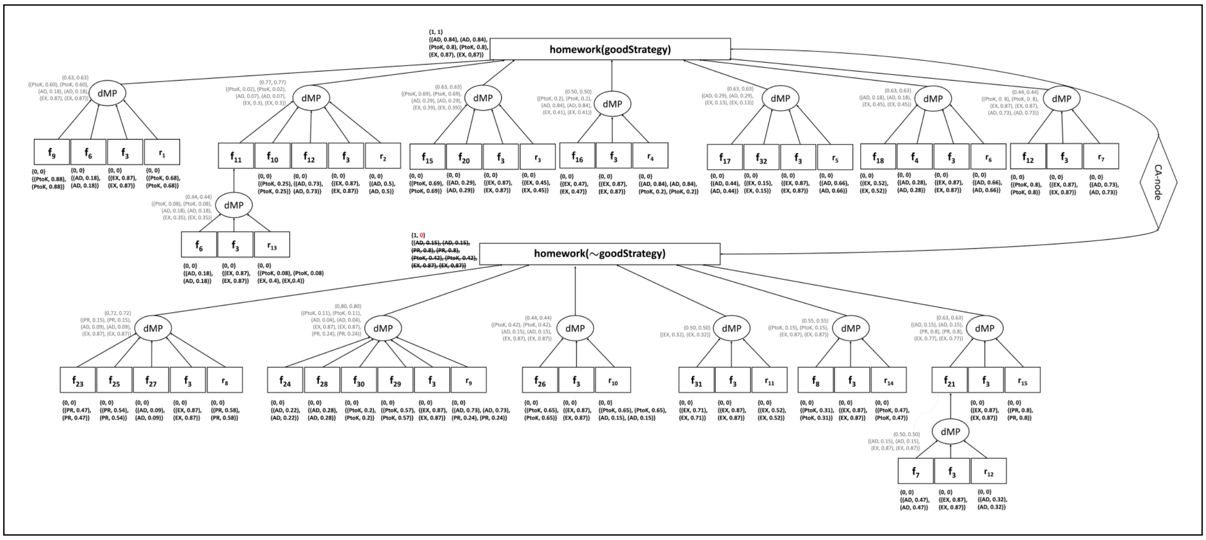

Figure 28. Labeled argumentation graph for the debate about homework.

determine initial community seeds. The nodes represent users, while the edges are relations between them. The algorithm is based on initial community seeds, which are expanded to obtain the complete community structure. The approach has proved to be efficient and could be applied to general graphs. In our work, we focus on argumentative graphs in which the opinions are represented (instead of specific users) and center our analysis on both the argumentative structure and the content of the nodes. Another recent work that combines graphs with *similarity measures* in community detection is that of Paul et al.,[45] in which the authors present an algorithm called local group assimilation (LGA) for detecting communities in complex networks. The approach combines local and global information to identify *similar groups* of nodes and merge them into larger communities. In this case, several similarity measures are used to calculate common neighbors to a node, link probabilities, and resource transmission through common neighbors. As in the previously mentioned work, it is a general algorithm based solely on the graph structure, while we propose a mechanism that combines the structure with a variety of information about each node. Also aligned with our work, Zhang et al.[46] address the problem of detecting communities in attribute networks, where existing algorithms often fail to consider both the topological and attribute similarities between nodes. The authors propose an evolutionary multi-objective approach to improve the accuracy of community detection by integrating a similarity fusion strategy with central nodes. The authors use diffusion kernel similarity to compute the topological similarity and design an attribute similarity matrix that accounts for the relationships between nodes based on their attributes. A fusion similarity matrix is constructed by weighting the topological and attribute similarities. The nodes with the highest fusion similarity values are chosen as central nodes to initialize the community detection process. The use of evolutionary multi-objective optimization and a modularity-based correction strategy leads to more accurate and coherent community detection, especially when compared with existing algorithms. However, the high computational cost is noted as a limitation, and future research will focus on optimizing the process further. Clearly, our work does not refer to topological features and the use of the similarity measures has a well-differentiated semantic notion.

Coronel et al.[39] present an extension of multi-topic abstract argumentation semantics, introducing the concepts of neighborhoods and communities of legitimate defenses. Each argument is associated with a zone of relevance determined by the topics it addresses, leading to the formation of proximity-based neighborhoods. Admissibility is defined by the presence of defenders within these neighborhoods, while arguments outside this range are considered less relevant. The work also introduces the idea of communities, where arguments that are not direct neighbors can still influence acceptability through shared connections with other arguments. This approach expands the scope of defense beyond individual neighborhoods, creating a network of interconnected argument communities. By incorporating mathematical tools to establish argument-specific proximity thresholds based on local and global topological features, the proposed framework provides a more nuanced understanding of argument relations. This proximity semantics not only improves the precision of accepted arguments but also offers deeper insights into the structure of argumentative networks. In the present work, community detection is not based on proximity but rather on the similarity between pieces of knowledge, while also incorporating features of the quality of the proponent. Budan et al.[3] present a method for detecting and analyzing communities on social media through an abstract argumentation-based approach. In this work, communities are viewed as groups of opinions that may be in conflict or supporting each other according to bipolar argumentation frameworks.[38,47] Under this approach, arguments with similar stances form coalitions that have associated a measure of cohesion which is obtain from a similarity-based evaluation. By assessing similarities between arguments, the authors classify communities and examine their internal cohesion and controversy. Although the similarity measure is applied in both works, the first one considers abstract argumentation, while the present work focuses on semi-structured argumentation. Furthermore, Budán et al.[3], do not consider the groups that put forward an argument or their qualities in the model. Malik et al.[48] address the problem of detecting hate speech in Nastaliq Urdu Facebook posts, a low-resource language. Using transfer learning models, the aim is to automatically classify content into two levels: hate speech detection and target community identification, including political, religious, or gender-based communities. A dataset was created from Pakistani Facebook posts, and fine-tuning techniques were applied to enhance model performance. Their results achieve 84.1% accuracy for target community detection. Our model considers any debate where opinions are identified as pro or con, not just hate speech. Furthermore, the technical models employed in both proposals are different.

9.2 Argumentation schemes

Regarding annotation schemes for identifying argumentation patterns, there exist several related approaches in the literature.[49–52] Visser et al.[51] state that the main problem with the AS proposed by Walton lies in deciding when a discourse fits the pattern, which is a very subjective task since it depends on who makes the annotations. For that reason, the annotators use a heuristic implemented via a decision tree classifier as a guide to make the annotations. Here, we rather focus on how to propagate the qualities of the source that puts forward and argument, rather than on who carries out the annotation process. According to experimental results,[52] the arguments from popular opinion are difficult to identify, even more so than the identification of their variants. An interesting work is presented by Jo et al.,[53] where Jo et al. present a method for classifying argumentative relations between statements based on logical mechanisms and AS. The mechanisms considered by the authors are: factual consistency, sentiment coherence, causal relation, and normative relation, demonstrating that these mechanisms can effectively classify argumentative relations without directly training on labeled data. For each of those mechanisms, they present the related AS. In our work, we do not directly identify the arguments from popular opinions, instead, we formalized the patterns proposed by Walton[10] according to the different sources that support an argument. Laurence et al.[49] develop a novel annotation tool designed to help with the empirical study of argument schemes. The authors highlight the importance of understanding inferential principles in argumentative discourse and describe how existing annotated corpora are limited in size and scope. The article introduces the argument scheme key as a method for annotating argument schemes based on Walton's typology. Clearly, our work not only considers the usefulness of AS, but the proposed mechanism also aims to identify argumentative communities within a debate. In relation to argumentative discourse representation, Benini et al.[54] introduced a framework for expressing these types of discourses using *Adpositional Argumentation*. This framework was designed to ensure a representation that is free from ambiguity. The authors based their research on adpositional trees, which serve to provide structure for linguistic material. A basic adpositional tree comprises a *governor* or ruler, a *dependent* or ruled element, and a connecting element known as an *adposition*. The adposition can be instantiated in various ways depending on the abstraction level, such as a proposition or an argument type. While the authors assert that adpositional argumentation addresses the limitations of Walton's AS specification, they do not provide an illustrative example of how they applied their formalism to AS. From our perspective, we work with an argumentative graph, decorated with quantitative and qualitative labels that describe an argument, indicating its cohesion and the groups that adhere to it. We are not only focused on the structure of the discourse but also on describing relevant features within it.

9.3 Valued based formalisms

According to Cayrol and Lagasquie-Schiex,[55] Amgoud and Ben-Naim, and[56] Potyka,[57] argumentation is modeled as a two-step process: (i) assessing the relative strength of arguments, and (ii) selecting the most acceptable ones. The authors introduce a gradual valuation of arguments based on their interactions and propose a graded concept of acceptability. Their framework defines aggregation and reduction operators for argument valuations but does not account for argument structure, relying solely on interactions for evaluation. In contrast, our approach incorporates the internal structure of arguments to identify and characterize communities that support specific statements assigning multiple valuations based on the attributes being modeled. Once these communities are established, we analyze their interactions within the debate and propagate these interactions through argumentation labels using operations defined within the algebra of argumentation labels. In contrast to this framework, enables users to define operations over argument relations, thereby allowing explicit adaptation to specific problem domains. Finally, we use the attributes assigned to the communities to determine the acceptability status and characterize their properties.

Dung's abstract argumentation framework is well-known for handling inconsistencies in knowledge bases by identifying arguments, establishing defeat relations, and determining acceptability through a collective semantic process. However, due to its high level of abstraction, it lacks the ability to assess arguments individually. To address this, Amgoud and Cayrol introduced the preference-based argumentation framework,[58] incorporating a preference relation to account for user-defined priorities. While the defeat relation captures logical conflicts (e.g., rebuttals or undercuts), the preference relation represents external meta-knowledge. An argument is defended if it is preferred over its attackers or supported by a set of stronger arguments. In contrast, our formalism explicitly details argument features to define preferences that satisfy specific contextual constraints. Unlike in preference-based frameworks,[58,59] our approach does not use preferences for conflict resolution; instead, this role is fulfilled by the conflict operator in the algebra of argumentation labels, which determines how argument valuations are affected in conflicts.

Kaci and van der Torre[60] extend Bench-Capon's value-based argumentation theory by allowing arguments to promote multiple values and defining preferences in various ways. Each value may be linked to multiple arguments and vice versa. Once values are assigned to arguments, conflicts are analyzed to determine successful attacks. In Bench-Capon's value-based framework, the attack of an argument $\mathbb{A}$ over an argument $\mathbb{B}$ is successful if and only if $\mathbb{A}$ attacks $\mathbb{B}$ and the value promoted by $\mathbb{B}$ is not preferred over that promoted by $\mathbb{A}$. The extended framework complicates this process, as arguments can support multiple values. To address this, the authors propose ordering values based on minimal/maximal specificity principles, enabling a total order that determines successful attacks and acceptable arguments. This process integrates non-monotonic reasoning with abstract argumentation techniques. Our approach shares similarities with this idea: argument valuations help establish argument strength, and multiple valuations capture attributes beyond logical soundness. However, we argue that each argument feature requires a distinct interpretation, ordering, and treatment. To this end, we introduce a set of algebras of argumentation labels to represent and compute different argument characteristics, modeling the dynamics of argumentation process. Furthermore, we use the argumentation structure in LAF to identify argumentation communities within a debate and characterize them based on the features of the knowledge elements they encompass.

Motivated by the idea of developing and encouraging argumentation about a given issue in social networks, Leite and Martins[61] proposed extending Dung's setting with voting capabilities on arguments, and a semantics where the arguments are assigned values that are drawn from a given set of possible values, and represent the argument strength (with regard to the graph topology and the stated social opinion in terms of votes). This proposal shares some aspects with the one presented in this article, both of which acquire additional information about the quality of the arguments; LAF, however, can be considered an extension of the social argumentative framework that makes it feasible to represent the multiple features associated with the arguments, such as user preferences and the truthfulness of the information carried by the arguments, among others. In addition, a number of similarities exist between the operations used to balance the social strength of arguments and the operations defined in the algebra of argumentation labels, but conflict resolution is modeled differently. Leite and Martins[61] formalize a weakening of the social values of the disputing arguments without ascertaining a strong defeat, while our LAF implementation shows that it neutralizes a rational and conventional solution in which the properties of the disputing arguments assert a strong defeat, formalizing a weakening of the winning argument and neutralizing the losing argument.

Rago and Toni[62] integrate argumentation into opinion polling to make the reasoning behind opinions more transparent. By using Quantitative Argumentation Debate for Voting (QuAD-V) frameworks, voters can add statements that justify their opinions, ensuring individual rationality. The method then aggregates these results into a collectively rational argumentation framework, enabling a structured evaluation of opinions in dialectical exchanges. In contrast, our approach focuses on enhancing Walton's ad-populum AS by introducing the concept of argumentative communities. While their method allows voters to add statements justifying their opinions and aggregates results into a collectively rational framework, our method goes further by using a labeling process to measure the similarity between popular opinions and proposed claims, reflecting the argumentative group to which the proponent belongs. Additionally, our mechanism accounts for the variability of popular opinions and the quality of the proponents based on context, which is not explicitly addressed in the QuAD-V framework. Finally, we demonstrate the implementability of our mechanism using existing tools like LLMs, and work on developing a specialized tool for generating argumentation graphs from if-then rules, aiming for further improvement.

Dupuis de Tarle et al.[63] investigate how debates evolve under gradual argumentation semantics in multi-agent environments, primarily focusing on the dynamic exchange of arguments and how agents' opinions change through iterative interactions. This study provides insights into the convergence of viewpoints and the impact of learning mechanisms in argument-based decision-making. Alternatively, Baroni et al.[64] take a more theoretical stance by systematizing the principles that govern gradual argumentation semantics, unifying various fine-grained properties into broader axiomatic frameworks, facilitating a more structured evaluation of argument strength and acceptability. Unlike these works, which emphasize individual argument strength and multi-agent opinion dynamics, our research integrates conceptual and structural analysis of arguments within social interactions. By leveraging AS introduced by Walton and employing computational models to quantify argumentative cohesion, our approach allows for the identification of discourse patterns, opinion clusters, and ideological alignments. Mossakowski and Neuhaus,[65] in turn, proposed a modular semantic model for bipolar weighted argumentation graph, separating the aggregation of supporting and attacking arguments from their influence on argument acceptability. Their framework provides an analysis of desirable semantic properties for weighted argumentation systems. While their work focuses on acceptability semantics, our proposal aims to characterize argumentative communities through ad populum schemes and source-related features using a labeled perspective.

9.4 Fuzzy and probabilistic frameworks

There has been significant progress in extending argumentation frameworks—both structured and abstract—with mechanisms to handle fuzzy and probabilistic reasoning. Approaches leveraging fuzzy logic generally fall into two categories: those that refine Dung's semantics using fuzzy sets and relations,[66],[67] and those that incorporate fuzzy logic to assign weights within the reasoning process underlying argument construction,[68–71] which aligns with our approach. Probabilistic approaches to argumentation have been extensively studied to integrate uncertainty into abstract argumentation frameworks. Gabbay and Rodrigues[72] propose an equational method to introduce probabilistic semantics into argumentation networks by translating them into classical propositional logic. Their approach defines probabilistic semantics by solving numerical equations over the real interval $[0, 1]$, allowing a formal connection between probabilistic reasoning and argumentation structures. In the same direction, Hunter and Thimm[73] extend probabilistic argumentation by addressing incomplete probability assignments, focusing on developing constraints for probability functions within argumentation frameworks and propose a method for completing partial probability assignments using the principle of maximum entropy. Furthermore, they introduce several probabilistic constraints, such as coherence, justification, and rationality, to formalize different probabilistic argumentation models. Similarly, probabilistic argumentation frameworks can be categorized into those extending abstract models[74–76] and those integrating probabilities into structured frameworks. The latter approach dates back nearly two decades,[77,78] followed by a period of limited activity and a more recent resurgence of interest.[79–81] In our work, we instantiate LAF using their flexible operations, in order to represent qualitative and quantitative features that characterize the strength of claims within the knowledge base. Unlike purely probabilistic frameworks, our approach employs LAF to capture argument interactions, conceptual relationships, and argumentative schemes. This allows for a more detailed analysis of discourse structures, moving beyond mere probability assignments. That is, our work explicitly detects communities based on argumentative schemes, particularly ad-populum argumentation, and characterizes their internal cohesion.

9.5 LLMs and argument mining

In recent years, LLMs have significantly advanced natural language processing, and one of the fields that has benefited from these advancements is argumentation mining. In this direction, Gorur et al.[82] address relation-based argument mining (RbAM), which focuses on identifying agreement (support) and disagreement (attack) relations among arguments. They note that RbAM is a challenging classification task, with existing methods failing to perform satisfactorily, and show that general-purpose LLMs, appropriately primed and prompted, can significantly outperform the best-performing baseline. They demonstrate this empirically, using two open-source LLMs (Llama-2 and Mistral), and conclude that LLMs can be used for the RbAM task using appropriate priming and prompting, but also highlight the downsides of slower inference time and greater GPU requirements. These results are relevant for future implementations of LAF, because until today we manually evaluate the kind of relation between two arguments. Furthermore, in our work we do not focus on the detection of support and attack relationships, but rather on how to characterize the dynamics inside of a structured debate. However, we employ the LLMs as tools to obtain the necessary elements to evaluate the formalism. Freedman et al.[83] introduce argumentative LLMs (ArgLLMs) to address the limitations of standard LLMs, specifically their lack of explainability and contestability. The ArgLLM pipeline aims to augment LLMs with argumentative reasoning to produce more transparent and contestable outputs by constructing argumentation frameworks, which represent arguments and their relationships, using the LLM. These structured frameworks serve as the basis for formal reasoning, enabling decisions made by the ArgLLM to be traceable and understandable. This work constitutes a relevant precedent to consider in the future implementation of our framework, as it allows us to evaluate how incorporating labeled knowledge pieces, as currently proposed, can enhance LLMs' explainability. Trajano et al.[15] present an approach for translating natural language arguments into computational representations using LLMs. The main objective of this work is to bridge the gap between human communication and computational reasoning, particularly in multi-agent systems and symbolic AI frameworks. This work leverages argumentation schemes to classify natural language arguments based on a predefined set of AS and employs retrieval-augmented generation (RAG) to provide contextual information to LLMs, guiding them in the translation task. The system retrieves relevant argumentation schemes and example mappings from a vector database to provide context for the LLMs. Then, the retrieved context is used to generate a computational representation of the argument while preserving its logical structure. Unlike other approaches, our work does not use AS to identify underlying reasoning patterns in a debate. Instead, we formalized these AS through our instantiation of LAF, and LLMs are used solely to extract concepts, attributes, if-then rules, and associated arguments in structured debates. Zhang et al.[84] propose LLM-UIE (efficient unified information extraction model based on LLMs) a parameter-efficient framework that enhance LLMs' performance in information extraction by introducing an answer selection module to align fuzzy outputs with structured answers. This contribution is particularly relevant for argument extraction, where identifying structured components from natural language is essential.

Ruiz-Dolz et al.[85] investigate the automatic identification of argumentative relations using transformer-based models, providing a cross-domain evaluation that highlights both the potential and the limitations of purely data-driven approaches. Agarwal et al.[86] introduce GraphNLI, a graph-based natural language inference model designed to predict polarity in online debates, combining neural representations with explicit graph structures. More recent contributions focus on the integration of LLMs as interactive or assistive components in argumentative settings. Faugier et al.[87] propose an LLM-assisted debate builder aimed at supporting users during the construction and exploration of argumentative structures. Similarly, Savigny and Yun[88] present AMELIA, a family of multi-task end-to-end language models addressing several argumentation tasks within a unified neural architecture. These approaches primarily rely on end-to-end learning, statistical inference, or interactive assistance to extract, predict, or generate argumentative structures from natural language. In contrast, our proposal is not concerned with argument extraction or linguistic classification, but with the formal representation and evaluation of arguments once they are instantiated as a graph. The algebraic labeling framework we introduce provides explicit semantics for aggregation, conflict resolution, and proponent-related reasoning, offering transparency and explainability that are typically absent from purely neural approaches. As such, our work is complementary to LLM-based argument mining systems and can naturally serve as a downstream reasoning and evaluation layer over argument graphs produced by such models.

9.6 *Computational argumentation and its applications*

Donadello et al.[89] investigate how automated persuasion systems can enhance their effectiveness by predicting user-specific utility functions through machine learning. The authors introduce two methods, EAI and EDS, to learn these functions based on user data. While EAI relies on a fixed amount of information, EDS dynamically selects the most relevant data to identify user subpopulations, proving more efficient in real-world applications, such as promoting healthy eating habits. In another direction, Rago et al.[90] address the growing demand for interactive explanations in Explainable AI. The study introduces Argumentative eXchanges (AXs), a framework that enables structured argumentation between AI models and humans to resolve conflicts over AI decisions. By leveraging quantitative bipolar argumentation frameworks, AXs facilitate interactive dialogues where AI can adapt explanations based on user responses. Experiments in simulated environments reveal that the most logically robust argument is not always the most persuasive, highlighting the importance of adapting argumentation strategies. Furthermore, Lawrence et al.[91] examine the evolution of argument technology and the development of the Argument Web, an integrated platform designed to analyze, visualize, and manipulate debates. This research highlights collaborations with the BBC to scale the platform for public discourse analysis, gamification, and educational initiatives. By making argumentation tools accessible to broader audiences, the study demonstrates how computational argumentation can transition from academic research to real-world applications, enhancing public engagement in structured debates and decision-making processes. In contrast to this contributions, which prioritize practical deployment, the present work adopts a more theoretical and formal approach. It centers on the refinement of AS, with particular emphasis on logical rigor and computational soundness in the representation of argument structures. Zhong and Negre[92] present CA-FATA, a framework that integrates matrix factorization with argumentation frameworks to achieve both predictive accuracy and interpretability in recommender systems. Their approach models user-item interactions as arguments, where contextual information and feature importance dynamically shape recommendation explanations. While the work of Zhong and Negre[92] applies argumentation to enhance the transparency of personalized recommendations, are focus is on online debates. CA-FATA employs argumentation to clarify algorithmic outputs, whereas the formalism proposed in this work leverages it to uncover collective reasoning in social interactions.

9.7 *Tools for argument mining*

Bikakis et al.[93] propose a prototype of ArgSE (Argument Search Engine), a tool for retrieve arguments both for and against a claim, adapted to specific audiences. With ArgSE the arguments can be filtered by credibility, structure and persuasive style. The ArgSE goal is to empower users to better understand debates, make decisions based on more information, and participate in a more effective way in political debates. However, ArgSE needs to be improve in the following aspects: (i) the tool stills performs an incomplete argument extraction, specially when the arguments do not follow a standard format or include specific assumptions, (ii) has a limited argument quality assessment, failing in detecting fallacies or biases; and (iii) has a lack od rich interaction and explainability, among other issues.

10 Conclusion and future work

In this article, we have devised a mechanism based on the concept of argumentative community to enhance Walton's ad-populum AS in order to render them as unambiguously defined and computable elements. To achieve this, we leveraged a labeling process approach, using labels that represent the similarity between opinions, and the specific argumentative group to which the proponents belong. First, we employ labels to measure the similarity between popular opinions and the proposed claim, which enables us to quantify how representative the popular opinion is in relation to the claim. Then, we take into account the argumentative context by assigning labels that reflect the argumentative group to which the proponent belongs. This is important because popular opinions and the quality of the proponent can vary depending on the context, as well as on the targeted group. The robustness of this mechanism is supported by Caminada et al.'s principles.[40] Defined in an abstract way, these properties serve as design criteria to ensure that an argumentative system is not affected by external interference or inconsistencies in its knowledge base. In this regard, an approach based on LAF would allow for explicitly modeling the influence and validity of arguments within a computational system, ensuring an effective transition from theoretical framework to functional implementation. In addition, to demonstrate that the proposed mechanism is implementable using existing tools, we document the results obtained in the different stages of the procedure using the aid of LLMs. However, these results can be improved by applying more specialized tools or frameworks such LLM-UIE[84] that tackles the limitations of LLMs in information extraction by applying parameter-efficient fine-tuning and incorporating an answer selection mechanism. Besides, we are currently working on developing automatic tools capable of generating the argumentation graph from if-then rules. The exercise carried out through the construction of the use case can be helpful in designing a pipeline that seamlessly integrates the work of knowledge engineers, LLMs, and tools for argumentative mining and reasoning.

Future work includes conducting additional experiments to further validate and assess the effectiveness of our formalism, as well as exploring the implementation of an automated tool for building the labeled graph. These experiments would help evaluate how our labels and approaches contribute to enhancing the capability of ad-populum AS in representing and analyzing arguments based on popular opinions. More broadly, this line of research aims to advance the understanding and application of tools for argumentative reasoning, particularly by improving the clarity and definition of Walton's schemes.

Acknowledgments

This work was supported by funds provided by Universidad Nacional de Santiago del Estero (UNSE) under grant 23/C201PIP-2023, Universidad Nacional del Sur (UNS) under grant PGI 24/ZN057, Agencia Nacional de Promoción Científica y Tecnológica under grant PICT-2018-0475 (PRH-2014-0007), and CONICET under grant PIP 11220210100577CO. The authors also acknowledge support from project PCI2022-135010-2 funded by MICIU/AEI 10.13039/501100011033 and by the European Union NextGenerationEU/PRTR.

The complete application, together with usage instructions, is publicly available at https://laf.ditomatto.com.ar/ and https://goo.su/r4aC3. The deployment provides all the functionality required to reproduce the analysis reported in this work, together with the corresponding usage instructions.

Funding

The authors received no financial support for the research, authorship and/or publication of this article.

Conflicting Interests

The authors declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Supplemental material

Supplemental material for this article is available online.

ORCID iDs

Paola D Budán https://orcid.org/0000-0002-7688-6181
Maximiliano J Debórtoli https://orcid.org/0009-0007-1995-7645
Maximiliano CD Budán https://orcid.org/0000-0002-9900-6511

Gerardo I Simari https://orcid.org/0000-0003-3185-4992
Maria Vanina Martinez https://orcid.org/0000-0003-2819-4735

References

1. Alsinet T, Argelich J, Béjar R, et al. Measuring polarization in online debates. *Appl Sci* 2021; 11: 24.
2. Kumar S, Hamilton WL, Leskovec J, et al. Community interaction and conflict on the web. In: *Proceedings of the 2018 world wide web conference, WWW '18, ACM*, 2018; pp.933–943.
3. Budán PD, Escañuela Gonzalez MG, Budán MCD, et al. Strength in coalitions: Community detection through argument similarity. *Argum Comput* 2023; 14: 275–325.
4. Irani A, Faloutsos M and Esterling K ArguSense. Argument-centric analysis of online discourse. *Proc Int AAAI Conf Web Soc Med* 2024; 18: 663–675.
5. Cetin P and Amrahov EŞ. A new network-based community detection algorithm for disjoint communities. *Turkish J Elect Eng Comp Sci* 2022; 30: 2190–2205.
6. Chi H and Liao B. A quantitative argumentation-based automated explainable decision system for fake news detection on social media. *Knowl Based Syst* 2022; 242: 108378.
7. Walton D. Justification of argumentation schemes. *Austral J Log* 2005; 3: 1–13.
8. Walton D. *Fundamentals of critical argumentation*. Cambridge University Press, 2006.
9. Walton D, Reed C and Macagno F. *Argumentation Schemes*. Cambridge University Press, Cambridge, UK, 2008.
10. Walton D. *Appeal to popular opinion*. Penn State Press, 1999.
11. Walton D and Reed CA. Argumentation schemes and enthymemes. *Synthese* 2005; 145: 339–370.
12. Budán MC, Lucero MJG and Simari GR. An aif-based labeled argumentation framework. In: *Foundations of information and knowledge systems*, 2014, pp.117–135. Springer.
13. Budán PD, Escañuela Gonzalez MG, Budán MCD, et al. Similarity notions in bipolar abstract argumentation. *Argument Comput* 2020; 11: 103–149.
14. Budán PD, Martinez MV, Budán MCD, et al. Introducing analogy in abstract argumentation. *WL4AI-2015* 2015; 25: 25–32.
15. Trajano G, Engelmann DC, Bordini RH, et al. Translating natural language arguments to computational arguments using LLMs. In: *Computational models of argument*, 2024, pp.289–300. IOS Press.
16. Collins P and Hahn U. Arguments and their sources. In: *The psychology of argument: Cognitive approaches to argumentation and persuasion* 2016, pp.129–149. College Publications.
17. Macagno F. Arguments of interpretation and argumentation schemes. In: Manzin M, Puppo F and Tomasi S (eds) *Studies on argumentation and legal philosophy: Further steps towards a pluralist approach* 2015. pp.51–80.
18. Chesñevar CI, McGinnis J, Modgil S, et al. Towards an argument interchange format. *Knowl Eng Rev* 2006; 21: 293–316.
19. Garg H and Rani D. New generalised Bonferroni mean aggregation operators of complex intuitionistic fuzzy information based on Archimedean t-norm and t-conorm. *J Exp Theoret Artif Intell* 2020; 32: 81–109.
20. Zhao S. Comparative semantics of fuzzy argumentation under the Łukasiewicz t-Norm. *Engineering* 2022; 14: 43–53.
21. Budán M, Lucero MG, Viglizzo I, et al. A labeled argumentation framework. *J Appl Log* 2015; 13: 534–553.
22. Budán MC, Cobo ML, Martinez DC, et al. An approach to characterize topic-centered argumentation. In: *PRIMA 2017: Principles and practice of multi-agent systems: 20th International Conference, Nice, France, October 30–November 3, 2017, Proceedings 20*, 2017, pp.459–467. Springer.
23. Gonzalez MGE, Budán MC, Simari GI, et al. Labeled bipolar argumentation frameworks. *J Artif Intell Res* 2021; 70: 1557–1636.
24. Budán M, Lucero MG and Simari G. An AIF-based labeled argumentation framework. In: *International symposium on foundations of information and knowledge systems*, 2014, pp.117–135. Springer.
25. Gottifredi S, García AJ and Simari GR. A study of AIF argument networks anomalies and a characterization of its solutions. In: *XIII Argentine symposium on Artificial Intelligence (ASAI 2012)(XLI JAIIO, La Plata, 27 y 28 de agosto de 2012)*, 2012.
26. Lucero MJG, Chesñevar CI and Simari GR. Modelling argument accrual with possibilistic uncertainty in a logic programming setting. *Inf Sci (Ny)* 2013; 228: 1–25.
27. Prakken H. A study of accrual of arguments, with applications to evidential reasoning. In: *ICAIL '05: Proceedings of the 10th International Conferences on Artificial Intelligence and Law*, 2005, pp.85–94, New York, NY, USA. ACM.
28. Verheij B. Accrual of arguments in defeasible argumentation. In: *Proceedings of the 2nd Dutch/German workshop on nonmonotonic reasoning*, 1995, pp.217–224, Utrecht.
29. Čebirić Š., Goasdoué F, Kondylakis H, et al. Summarizing semantic graphs: a survey. *VLDB J* 2019; 28: 295–327.
30. Opdahl AL, Al-Moslmi T, Dang-Nguyen D-T, et al. Semantic knowledge graphs for the news: a review. *ACM Comput Surv* 2022; 55: 1–38.
31. Guarino N, Oberle D and Staab S. What is an ontology? In: *Handbook on ontologies*, 2009, pp.1–17, Springer.

32. Baclawski K, Kokar MK, Kogut PA, et al. Extending UML to support ontology engineering for the semantic web. In: *International conference on the unified modeling language*, 2001, pp.342–360. Springer.
33. Chen X, Jia S and Xiang Y. A review: knowledge reasoning over knowledge graph. *Expert Syst Appl* 2020; 141: 112948.
34. Zhu G and Iglesias C. Computing semantic similarity of concepts in knowledge graphs. *IEEE Trans Knowl Data Eng* 2016; 29: 72–85.
35. Rada R, Mili H, Bicknell E, et al. Development and application of a metric on semantic nets. *IEEE Trans Syst Man Cybern* 1989; 19: 17–30.
36. David V and Mailly JDJ-G. Similarity measures between order-sorted logical arguments.
37. Amgoud L and David V. Measuring similarity between logical arguments. In: *Sixteenth international conference on principles of knowledge representation and reasoning (Proceeding KR)*, 2018.
38. Amgoud L. Towards a formal model for task allocation via coalition formation. In: *Proceeding of the 4th International joint conference on Auton. Agents Multiagent Systems*, 2005, pp.1185–1186.
39. Coronel IM, Gonzalez MGE, Martinez DC, et al. Neighborhood-based argumental community support in the context of multi-topic debates. *Int J Approx Reason* 2024; 170: 109189.
40. Caminada M, Carnielli W and Dunne P. Semi-stable semantics. *J Log Comput* 2012; 22: 1207–1254.
41. Derczynski L. Complementarity, f-score, and NLP evaluation. In: *Proceedings of the Tenth international conference on language resources and evaluation (LREC'16)*, 2016, pp.261–266.
42. Wang R and Li J. Bayes test of precision, recall, and f1 measure for comparison of two natural language processing models. In: *Proceedings of the 57th Annual meeting of the association for computational linguistics*, 2019, pp.4135–4145.
43. Debórtli JM. *Evaluación de la calidad de la minería de argumentos en debates públicos realizada por los Modelos de Lenguajes de Gran Escala (LLMs)* Tesis Licenciatura en Sistemas de Información – Universidad Nacional de Santiago del Estero, Argentina. 2026.
44. Zhao X, Liang J and Wang J. A community detection algorithm based on graph compression for large-scale social networks. *Inf Sci (Ny)* 2021; 551: 358–372.
45. Paul A and Dutta A. Community detection using local group assimilation. *Expert Syst Appl* 2022; 206: 117794.
46. Zhang W, Zhao K and Shang R. Evolutionary multi-objective attribute community detection based on similarity fusion strategy with central nodes. *Appl Soft Comput* 2024; 150: 111101.
47. Amgoud L, Cayrol C and Lagasquie-Schiex M. On the bipolarity in argumentation frameworks. In: *Nmr*, 2004, Vol. 4, pp.1–9.
48. Malik MSI, Nawaz A and Jamjoom MM. Hate speech and target community detection in Nastaliq Urdu using transfer learning techniques. *IEEE Access* 2024; 12: 116875–116890.
49. Lawrence J, Visser J and Reed C. An online annotation assistant for argument schemes. In: *Proceedings of the 13th linguistic annotation workshop*, 2019, pp.100–107. Association for Computational Linguistics.
50. Musi E, Ghosh D and Muresan S. Towards feasible guidelines for the annotation of argument schemes. In: *Proceeding ArgMining2016*, 2016, pp.82–93.
51. Visser J, Lawrence J, Reed C, et al. Annotating argument schemes. In: *Argumentation through languages and cultures*, 2020, pp.101–139. Springer.
52. Visser J, Lawrence J, Wagemans J, et al. Revisiting computational models of argument schemes: Classification, annotation, comparison. In: *7th International conference on computational models of argument, COMMA 2018*, 2018, pp.313–324. ios Press.
53. Jo Y, Bang S, Reed C, et al. Classifying argumentative relations using logical mechanisms and argumentation schemes. *Trans Assoc Comput Linguist* 2021; 9: 721–739.
54. Benini M, Gobbo F and Wagemans J. The logic of the arguer. representing natural argumentative discourse in adpositional argumentation. *J Appl Log* 2023; 2631: 451.
55. Cayrol C and Lagasquie-Schiex M-C. Graduality in argumentation. *J Artif Intell Res.(JAIR)* 2005; 23: 245–297.
56. Amgoud L and Ben-Naim J. Evaluation of arguments in weighted bipolar graphs. *Int J Approx Reason* 2018; 99: 39–55.
57. Potyka N. Continuous dynamical systems for weighted bipolar argumentation. *KR* 2018; 2018: 148–157.
58. Amgoud L and Cayrol C. On the acceptability of arguments in preference-based argumentation. In: *Proceedings of the Fourteenth conference on Uncertainty in artificial intelligence*, 1998, pp.1–7. Morgan Kaufmann Publishers Inc.
59. Teze JCL, Gottifredi S, García AJ, et al. An approach to generalizing the handling of preferences in argumentation-based decision-making systems. *Knowl Based Syst* 2020; 189: 105112.
60. Kaci S and van der Torre L. Preference-based argumentation: arguments supporting multiple values. *Int J Approx Reason* 2008; 48: 730–751.
61. Leite J and Martins J. Social abstract argumentation. In: *Proceedings of the IJCAI*, 2011, pp.2287–2292.
62. Rago A and Toni F. Quantitative argumentation debates with votes for opinion polling. In: *PRIMA 2017: Principles and practice of multi-agent systems: 20th International Conference, Nice, France, October 30–November 3, 2017, Proceedings 20*, 2017, pp.369–385. Springer.

63. De Tarlé LD, Bonzon E and Maudet N. Multiagent dynamics of gradual argumentation semantics. In: *21st International conference on autonomous agents and multiagent systems (AAMAS 2022)*, 2022.
64. Baroni P, Rago A and Toni F. From fine-grained properties to broad principles for gradual argumentation: a principled spectrum. *Int J Approx Reason* 2019; 105: 252–286.
65. Mossakowski T and Neuhaus F. Modular semantics and characteristics for bipolar weighted argumentation graphs. *arXiv preprint arXiv:1807.06685*, 2018.
66. Janssen J, De Cock M and Vermeir D. Fuzzy argumentation frameworks. In: *Information processing and management of uncertainty in knowledge-based systems*, 2008, pp.513–520, 2008.
67. Tamani N and Croitoru M. Fuzzy argumentation system for decision support. In: *Information Processing and management of uncertainty in knowledge-based systems*, 2014, pp.77–86. Springer.
68. Alsinet T, Chesñevar CI, Godo L, et al. Formalizing argumentative reasoning in a possibilistic logic programming setting with fuzzy unification. *Int J Approx Reason* 2008; 48: 711–729.
69. Chesñevar CI, Simari GR, Alsinet T, et al. A Logic Programming Framework for Possibilistic Argumentation with Vague Knowledge. In: *Proceeding of the International Conference in Uncertainty in Artificial Intelligence (UAI 2004). Banff, Canada*, 2004, pp.76–84.
70. Schroeder M and Schweimeier R. Fuzzy argumentation and extended logic programming. In: *Proceedings of ECSQARU workshop adventures in argumentation*, 2001.
71. Schweimeier R and Schroeder M. Fuzzy unification and argumentation for well-founded semantics. In: *SOFSEM 2004: Theory and practice of computer science, 30th conference on current trends in theory and practice of computer science*, 2004, pp.102–121.
72. Gabbay DM and Rodrigues O. Probabilistic argumentation: an equational approach. *Log Univers* 2015; 9: 345–382.
73. Hunter A and Thimm M. Probabilistic argumentation with incomplete information. In: *ECAI 2014*, 2014, pp.1033–1034. IOS Press.
74. Fazzinga B, Flesca S and Parisi F. On the complexity of probabilistic abstract argumentation frameworks. *ACM Trans Comput Log* 2015; 16: 22.
75. Hunter A. Some foundations for probabilistic abstract argumentation. In: *Proceeding COMMA*, 2012, pp.117–128.
76. Thimm M. A probabilistic semantics for abstract argumentation. In: *Proceeding ECAI*, 2012, pp.750–755.
77. Haenni R, Kohlas J and Lehmann N. *Probabilistic argumentation systems*. Springer, 1999.
78. Krause P, Ambler S, Elvang-Gøransson M, et al. A logic of argumentation for reasoning under uncertainty. *Comput Intell* 1995; 11: 113–131.
79. Hunter A. A probabilistic approach to modelling uncertain logical arguments. *Int J Approx Reason* 2013; 54: 47–81.
80. Shakarian P, Simari GI, Moores G, et al. Cyber attribution: An argumentation-based approach. In: *Cyber Warfare: Building the scientific foundation*, volume 56 of *Advances in Information Security*, 2015, pp.151–171. Springer.
81. Simari GI, Shakarian P and Falappa MA. A quantitative approach to belief revision in structured probabilistic argumentation. *Ann Math Artif Intell* 2016; 76: 375–408.
82. Gorur D, Rago A and Toni F. Can large language models perform relation-based argument mining? *arXiv preprint arXiv:2402.11243*, 2024.
83. Freedman G, Dejl A, Gorur D, et al. Argumentative large language models for explainable and contestable decision-making. *arXiv preprint arXiv:2405.02079* 2024.
84. Zhang X, Cai S, Shen X, et al. Efficient unified information extraction model based on large language models. *Appl Soft Comput* 2025; 180: 113302.
85. Ruiz-Dolz R, Plaza L and Gervás P. Transformer-based models for automatic identification of argument relations A cross-domain evaluation. *Expert Syst Appl* 2021; 184: 115496.
86. Agarwal S and Bhargava PM. GraphNLI: A Graph-Based Natural Language Inference Model for Polarity Prediction in Online Debates. In: *Proceedings of the 60th Annual meeting of the association for computational linguistics (ACL)*, 2022, pp.6970–6981.
87. Faugier E, Armetta F, Bonifati A, et al. Assisted Debate Builder with Large Language Models. In: *Proceedings of the 27th European conference on artificial intelligence (ECAI)*, 2024, pp.4447–4450.
88. Savigny H and Yun B. AMELIA: A Family of Multi-task End-to-end Language Models for Argumentation. *CoRR*, abs/2508.17926, 2025.
89. Donadello I, Hunter A, Teso S, et al. Machine learning for utility prediction in argument-based computational persuasion. In: *proceedings of the AAAI conference on artificial intelligence*, 2022, Vol. 36, pp.5592–5599.
90. Rago A, Li H and Toni F. Interactive explanations by conflict resolution via argumentative exchanges. In: *arXiv preprint arXiv:2303.15022*, 2023.
91. Lawrence J, Visser J and Reed C. Translational argument technology: Engineering a step change in the argument web. *J Web Semant* 2023; 77: 100786.

92. Zhong J and Negre E. Argumentation meets matrix factorization: a dual perspective for explainable recommendations. *Appl Soft Comput* 2025; 183: 113652.
93. Bikakis A, Flouris G, Patkos T, et al. Sketching the vision of the web of debates. *Front Artif Intell* 2023; 6: 1124045.