A summary of Community Detection from Coalitions through Argument Similarity

Paola D. Budán[1,2], Melisa G. Escañuela González[2], Maximiliano C. D. Budán [1,2], Maria Vanina Martinez[3] and Guillermo R. Simari[4]

[1]Institute for Computer Science and Engineering (CONICET-UNS), Bahía Blanca, Argentina
[2]Institute for Computer Science and Information Systems (ICSIS-UNSE), Universidad Nacional de Santiago del Estero, Santiago del Estero, Argentina
[3]Artificial Intelligence Research Institute (IIIA-CSIC), Carrer de Can Planas, Barcelona, Catalonia, Spain.
[4]Department of Computer Science and Engineering, Universidad Nacional del Sur (UNS), Bahía Blanca, Argentina

Abstract

We present a summary of an argumentation-based method for finding and analyzing communities in social media on the Web, where a community is regarded as a set of supported opinions that might be in conflict. First, we identify argumentative coalitions to define communities; then, we apply a similarity-based evaluation method over the set of arguments in the coalition to determine the level of cohesion inherent to each community, classifying them appropriately. Introducing conflict points and attacks between coalitions based on argumentative (dis)similarities to model the interaction between communities leads to considering a meta-argumentation framework where the set of coalitions plays the role of the set of arguments and where the attack relation between the coalitions is assigned a particular strength which is inherited from the arguments belonging to the coalition. Various semantics are introduced to consider attacks' strength to particularize the effect of the new perspective.

Keywords

Argument Similarity, Communities, Coalitions, Strength of Attacks between communities

1. Introduction

The identification of *communities* in social media and the *detection of stances* in Tweets has become increasingly important in recent times [1, 2, 3, 4, 5] as a result of the tangible effect that these platforms have on the public opinion. In this domain, identifying communities implies analyzing the position of contributing agents concerning a particular topic or their respective *argumentative stance*; several tools can be used for this purpose, for instance [1, 6, 7]. Most of these methods focus on analyzing tweets to characterize the relationship between messages.

The present work summarizes the *argumentation-based method* propose by Budán et al. [8] to analyze stances in a debate exchange and formally characterize the relationships between these stances using similarity, understanding the similarity as an attribute of the relationship. Although *Formal Argumentation Theory* provides several formalisms to model emerging behavior, e.g. [9, 10, 11, 12, 13]; this paper is based on the well-known *Abstract Argumentation Frameworks*

7th Workshop on Advances in Argumentation in Artificial Intelligence (AI³), November 06–09, 2023, Rome, Italy
pbudan1@gmail.com (P. D. Budán); meliesca@unse.edu.ar (M. G. E. González); mbudan@unse.edu.ar (M. C. D. B.); vmartinez@iiia.csic.es (M. V. Martinez); grsimari@gmail.com (G. R. Simari)

(AFs) proposed by Dung [14], and in the Cayrol and Lagasquie-Schiex [15] extended Dung's framework that considers two independent types of interactions between arguments: *attack* and *support*. This formalism is called *Bipolar Argumentation Frameworks* (BAFs), and models situations where arguments may give support for other arguments.

Based on the bipolar formalism, in [16] the authors presented an approach to use a similarity degree measure between arguments to characterize the attack and support relations in a BAF [17, 15]. In [8], the detection of communities in social media is based on the support and similarity relations between arguments, while the classification of the communities was done according to the similarity degree between the stances that conform to them, taking advantage of the notion of coalition presented in [18] and the framework proposed in [16]. During the whole work, the context where the argumentation discussion is put into play is considered. Note that we mainly refer to discursive communities. This clarification is necessary because it will allow us to regard communities as subgroups with cohesive thinking.

The following example will illustrate the ideas involved in this research. Consider the set of opinions extracted from the *ProCon* website[§] in favor of (pro) or against (con) the following proposition *"Is Human Activity Primarily Responsible for Global Climate Change?"*:

- A *Con1: More than one thousand scientists disagree that human activity is primarily responsible for global climate change.*
- B *Con2: The Cook review of 11,944 peer-reviewed studies found that 66.4% of the studies had no stated position on anthropogenic global warming, and while 32.6% of the studies implied or stated that humans are contributing to climate change, only 65 papers (0.5%) explicitly stated: "that humans are the primary cause of global warming."*
- C *Pro1: The rise in atmospheric CO2 over the last century was caused by human activity, as it occurred at a rate much faster than natural climate changes could produce.*
- D *Pro2: A National Climate Assessment report said human-caused climate changes, such as increased heat waves and drought, "are visible in every state".*
- E *Undef1: A 2012 Purdue University survey found that 47% of climatologists challenge the idea that humans are primarily responsible for climate change and instead believe that climate change is caused by an equal combination of humans and the environment (37%), mostly by the environment (5%), or that there's not enough information to say (5%).*

We can roughly distinguish three communities that give opinions regarding the responsibility of humans for climate change: one of them supports the idea that human activity is responsible for climate change (arguments `C` and `D`), another confronts the previous one with the opposite position (arguments `A`, `B`), and lastly, there is argument `E` representing an intermediate posture between the other two. By analyzing these well-defined general postures, we will obtain the details of the beliefs each community backs; but, by closely examining the opinions in each community, we can determine each community's inherent *strength*.

Given a system that represents knowledge as arguments and considers the existing conflicts and supports between these arguments, it is feasible to create maximal cohesive sets of arguments

[§]See https://climatechange.procon.org. The ProCon website states that its goal is "To promote civility, critical thinking, education, and informed citizenship by presenting the pro and con arguments to debatable issues in a straightforward, nonpartisan, freely accessible way."

by taking advantage of the mechanism proposed in [16] to collect in a set as many conflict-free and related-by-support arguments as possible, ensuring coherence of the whole set.

2. Background

In [15], Cayrol and Lagasquie-Schiex proposed an approach known as *Bipolar Argumentation Framework*, to model the *support* and the *attack* between arguments:

Definition 1 (Bipolar Argumentation Framework (BAF)). *A Bipolar Argumentation Framework is a 3-tuple $\Theta = \langle \mathtt{Args}, \mathtt{R}_a, \mathtt{R}_s \rangle$, where $\mathtt{Args}$ is a set of arguments, and $\mathtt{R}_a$ and $\mathtt{R}_s$ are two disjoint binary relations defined on $\mathtt{Args}$ called attack and support, respectively.*

Cayrol *et al.* [15] presented the extensions of the attack and support notions introducing the *supported* and *secondary* defeats, denoting with G_Θ the bipolar argumentation graph:

Definition 2 (Defeat in BAF). *Let $\Theta = \langle \mathtt{Args}, \mathtt{R}_a, \mathtt{R}_s \rangle$ be a BAF, and $\mathtt{A}, \mathtt{B}$ two arguments in $\mathtt{Args}$. Then, it is said that:*

- *$\mathtt{A}$ is a* supported defeat *for $\mathtt{B}$ iff there exists a sequence $\mathtt{A}_1\ \mathtt{R}_1\ \ldots\ \mathtt{R}_n\ \mathtt{A}_{n+1}$, with $n \geq 1$, where $\mathtt{A}_1 = \mathtt{A}$ and $\mathtt{A}_{n+1} = \mathtt{B}$, such that $\mathtt{R}_i = \mathtt{R}_s, 1 \leq i \leq n-1$, and $\mathtt{R}_n = \mathtt{R}_a, \mathtt{A}_i \in \mathtt{Args}, 1 \leq i \leq n+1$.*
- *$\mathtt{A}$ is a* secondary defeat *for $\mathtt{B}$ iff there exists a sequence $\mathtt{A}_1\ \mathtt{R}_1\ \ldots\ \mathtt{R}_n\ \mathtt{A}_{n+1}$, with $n \geq 2$, where $\mathtt{A}_1 = \mathtt{A}$ and $\mathtt{A}_{n+1} = \mathtt{B}$, such that $\mathtt{R}_1 = \mathtt{R}_a$, and $\mathtt{R}_i = \mathtt{R}_s, 2 \leq i \leq n, \mathtt{A}_i \in \mathtt{Args}, 1 \leq i \leq n+1$.*

Considering the simplest case of defeat in any BAF, a sequence of two arguments $\mathtt{A}\ \mathtt{R}_a\ \mathtt{B}$ is also regarded as a supported defeat from $\mathtt{A}$ to $\mathtt{B}$, i.e., a direct defeat is also a supported defeat.

A set of arguments must keep a minimum of coherence not containing an argument that attacks another one in the same set [15]. And an *external* coherence requiring that the set does not include both a supporter and an attacker of the same argument:

Definition 3 (Conflict-freeness and Safety Properties in BAF). *Let $\Theta = \langle \mathtt{Args}, \mathtt{R}_a, \mathtt{R}_s \rangle$ be a BAF, and $\mathtt{S} \subseteq \mathtt{Args}$ be a set of arguments. We say that $\mathtt{S}$ is* conflict-free *iff $\nexists\, \mathtt{A}, \mathtt{B} \in \mathtt{S}$ s.t. there is an attack (direct, or supported, or secondary) from $\mathtt{A}$ to $\mathtt{B}$. We say that $\mathtt{S}$ is* safe *iff $\nexists\, \mathtt{A} \in \mathtt{Args}$ and $\nexists\, \mathtt{B}, \mathtt{C} \in \mathtt{S}$ s.t. there is an attack (direct, or supported, or secondary) from $\mathtt{B}$ to $\mathtt{A}$, and either there is a sequence of support from $\mathtt{C}$ to $\mathtt{A}$, or $\mathtt{A} \in \mathtt{S}$.*

The closure under $\mathtt{R}_s$ introduced in [15] is a requirement that only concerns the support relation.

Definition 4 (Closure Property in BAF). *Let $\Theta = \langle \mathtt{Args}, \mathtt{R}_a, \mathtt{R}_s \rangle$ be an BAF. $\mathtt{S} \subseteq \mathtt{Args}$ be a set of arguments. $\mathtt{S}$ is closed under $\mathtt{R}_s$ iff $\forall\ \mathtt{A} \in \mathtt{S}, \forall\ \mathtt{B} \in \mathtt{Args}$ if $\mathtt{A}\ \mathtt{R}_s\ \mathtt{B}$ then $\mathtt{B} \in \mathtt{S}$.*

Definition 5 (Coalitions in a BAF). *Let $\Theta = \langle \mathtt{Args}, \mathtt{R}_a, \mathtt{R}_s \rangle$ be a BAF, and G_Θ be a bipolar argumentation graph. A subset $\mathcal{C}_\Theta \subseteq \mathtt{Args}$ is a* coalition *in Θ iff $\mathcal{C}_\Theta$ is a maximal conflict free set in Θ such that the subgraph G'_Θ induced by $\mathcal{C}_\Theta$ is connected only by support relations.*

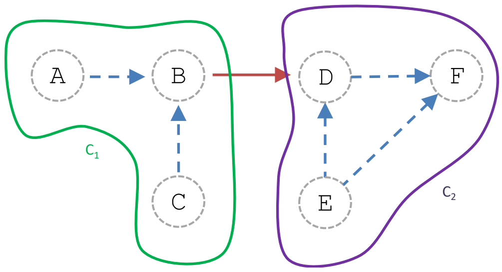

Figure 1: Coalitions in BAF

A coalition represents a relationship on the set of arguments; therefore, the notion of attack between them introduces a *meta-argumentation framework*:

Definition 6 (Attack between coalitions in BAF). *Let $\mathcal{C}_1$ and $\mathcal{C}_2$ be two coalitions over Θ. If there exist $\mathtt{A} \in \mathcal{C}_1$ and $\mathtt{B} \in \mathcal{C}_2$ such that $\mathtt{A}\ \mathtt{R}_a\ \mathtt{B}$, then the coalition $\mathcal{C}_1$ c-attacks (or just attacks) $\mathcal{C}_2$.*

Example 1. *In the figure 1 we present a BAF example described as $\Theta = \langle \mathtt{Args}, \mathtt{R}_a, \mathtt{R}_s \rangle$, where: $\mathtt{Args} = \{\mathtt{A}, \mathtt{B}, \mathtt{C}, \mathtt{D}, \mathtt{E}, \mathtt{F}, \mathtt{G}\}$; $\mathtt{R}_a = \{(\mathtt{B}, \mathtt{D})\}$; $\mathtt{R}_s = \{(\mathtt{A}, \mathtt{B}), (\mathtt{C}, \mathtt{B}), (\mathtt{E}, \mathtt{D}), (\mathtt{E}, \mathtt{F}), (\mathtt{D}, \mathtt{F})\}$. In the bipolar argumentation graph G_Θ of this particular BAF we have that the argument $\mathtt{G}$ is a secondary defeater for $\mathtt{F}$, while $\mathtt{C}$ and $\mathtt{A}$ are supported defeaters for argument $\mathtt{D}$. Moreover, we distinguish the following coalitions: $\mathcal{C}_1 = \{\mathtt{A}, \mathtt{B}, \mathtt{C}\}$, highlighted green and $\mathcal{C}_2 = \{\mathtt{E}, \mathtt{D}, \mathtt{F}\}$, highlighted purple. These are maximal conflict-free sets and $\mathcal{C}_1, \mathcal{C}_2$ are maximal sets closed under $\mathtt{R}_s$. Additionally, we have that: $\mathcal{C}_1$ attacks $\mathcal{C}_2$, because $\mathtt{B}$ attacks $\mathtt{D}$.*

2.1. A Similarity Valued Argumentation Framework

In [16] the authors presented a *Similarity-based Bipolar Argumentation Framework* (or S-BAF), which is a mechanism for considering the context of the comparison between arguments, introducing *enriched arguments*, that is, arguments decorated with additional information. They assume a set $\mathcal{D}$ of available descriptors corresponding to the discursive domain. Each descriptor has a set of values associated; for a descriptor $d \in \mathcal{D}$, $\mathcal{V}_d$ is the set of semantic values.

Definition 7 (Enriched Argument). *Let $\Theta = \langle \mathtt{Args}, \mathtt{R}_a, \mathtt{R}_s \rangle$ be a BAF, $\mathtt{A}$ be an abstract argument in Θ, and $\mathcal{D}$ be a set of descriptors. An* enriched argument *is a pair $\mathrm{A} = \langle \mathtt{A}, \delta_\mathtt{A} \rangle$, where $\delta_\mathtt{A}$ is a finite non-empty set of pairs $(d, \mathcal{V}_d^\mathtt{A})$, where $d \in \mathcal{D}$ and $\mathcal{V}_d^\mathtt{A} \subseteq \mathcal{V}_d$. The set of all enriched arguments will be denoted as* Args.

Next, we introduce the notion of context of the argumentation.

Definition 8 (Context). *Let $\mathcal{D}$ be a set of descriptors, a* context *$\mathbb{C}$ will be represented as $\mathbb{C} = \{(d, w_d) \mid d \in \mathcal{D}, w_d \in [0, 1]\}$, i.e., a context is a set of ordered pairs where $d \in \mathcal{D}$ is a descriptor and $w_d \in [0, 1]$ is the weight associated with d. We denote with $\mathcal{D}^\mathbb{C}$ the set of descriptors mentioned in the context $\mathbb{C}$, i.e., $\mathcal{D}^\mathbb{C} = \{d \mid (d, w_d) \in \mathbb{C}\}$.*

Given a context $\mathbb{C}$, for any argument $\mathrm{X} \in \mathrm{Args}$, we denote the descriptors in X that appear on the context $\mathbb{C}$ as $\mathcal{D}_\mathrm{X}^\mathbb{C}$, *i.e.*, $\mathcal{D}_\mathrm{X}^\mathbb{C} = \mathcal{D}_\mathrm{X} \cap \mathcal{D}^\mathbb{C}$.

Definition 9 (Similarity coefficient for a descriptor). *Let* Args *be a set of enriched arguments,* $\mathrm{A} = \langle \mathtt{A}, \delta_{\mathtt{A}} \rangle$ *and* $\mathrm{B} = \langle \mathtt{B}, \delta_{\mathtt{B}} \rangle$ *two enriched arguments in* Args*, and* $\mathbb{C}$ *a context. We define the similarity coefficient for each descriptor* $d \in \mathcal{D}_{\mathrm{A}}^{\mathbb{C}} \cap \mathcal{D}_{\mathrm{B}}^{\mathbb{C}}$ *with weight* w_d*, denoted* $\mathsf{Coef}_d(\mathrm{A}, \mathrm{B})$*, as follows:*

$$\mathsf{Coef}_d(\mathrm{A}, \mathrm{B}) = \begin{cases} \dfrac{\mid \mathcal{V}_d^{\mathtt{A}} \cap \mathcal{V}_d^{\mathtt{B}} \mid}{\mid \overline{\mathcal{V}_d^{\mathtt{A}} \cap \mathcal{V}_d^{\mathtt{B}}} \mid} \cdot w_d & \text{if } \mid \overline{\mathcal{V}_d^{\mathtt{A}} \cap \mathcal{V}_d^{\mathtt{B}}} \mid \neq 0 \\ \\ w_d & \text{otherwise} \end{cases}$$

Intuitively, to determine the similarity between two arguments based on a specific descriptor, you count the shared semantic values in that descriptor, divide by the non-shared values, and adjust it based on the descriptor's relevance in the given context, as mentioned in several references [19, 20, 21].

Definition 10 (Similarity degree between arguments). *Let* Args *be a set of enriched arguments,* $\mathrm{A} = \langle \mathtt{A}, \delta_{\mathtt{A}} \rangle$ *and* $\mathrm{B} = \langle \mathtt{B}, \delta_{\mathtt{B}} \rangle$ *be two enriched arguments in* Args*, and* $\mathbb{C}$ *be a context. The similarity degree between arguments in a context* $\mathbb{C}$*, denoted* $\mathsf{Sim}_{\mathbb{C}}$*, is defined as a function* $\mathsf{Sim}_{\mathbb{C}} : \mathsf{Args} \times \mathsf{Args} \to [0, 1]$*, such that:*

$$\mathsf{Sim}_{\mathbb{C}}(\mathrm{A}, \mathrm{B}) = \begin{cases} \alpha_n & \text{if } \mathcal{D}_{\mathrm{A}}^{\mathbb{C}} \cap \mathcal{D}_{\mathrm{B}}^{\mathbb{C}} = \{d_1, \ldots, d_n\} \neq \emptyset \\ 0 & \text{otherwise} \end{cases}$$

where $\alpha_1 = \mathsf{Coef}_{d_1}(\mathrm{A}, \mathrm{B})$ *and* $\alpha_i = \odot(\alpha_{i-1}, \mathsf{Coef}_{d_i}(\mathrm{A}, \mathrm{B}))$ *with* $2 \leq i \leq n$*, and, the operator* $\odot$ *should be either a* T-norm *satisfying the following properties: commutative, associative, monotonically increasing, and with* 1 *as its neutral element; or* $\odot$ *should be a* T-conorm*, satisfying commutative, associative, monotonically decreasing with* 0 *as its neutral element.*

The abstract concepts presented earlier will be illustrated in the following example.

Example 2. *Suppose that we are analyzing the arguments* $\mathtt{A}$ *and* $\mathtt{C}$ *of our example, which have the following descriptors and values:* $\delta_{\mathtt{A}} = \{(climate\_change, \{yes\});(refers\_evidence, \{yes\})\}$*;* $\delta_{\mathtt{C}} = \{(climate\_change, \{yes\});(refers\_evidence, \{no\});(human\_responsability, \{yes\})\}$*. Now, suppose that the context for the arguments comparison is the following:* $\mathbb{C} = \{(climate\_change, 0.4);(human\_responsability, 0.4);(non\_human\_causes, 0.2)\}$*. For* $climate\_change$ *descriptor, we have that the two arguments have a single value in common and no different ones. So, the* $\mathsf{Coef}_d(\mathrm{A}, \mathrm{C}) = 0.4$*; for* $human\_responsability$*, arguments have different values for this descriptor, and no common value. So that, according to the similarity coefficient definition, the* $\mathsf{Coef}_d(\mathrm{A}, \mathrm{C}) = 0$*; and for* non_human_causes*, it is not mentioned in the arguments under analysis. Now, considering the bounded sum T-conorm, we have that the* $\mathsf{Sim}_{\mathbb{C}}(\mathrm{A}, \mathrm{C}) = 0.4$*, given that:* $min(0.4 + 0, 1) = 0.4$*. The similarity value obtained reflects that the both arguments refer to climate change, but each argument in different way.*

The following definition introduces the enriched BAF framework based on the original BAF:

Definition 11 ((Induced) Enriched BAF). *Let* $\Theta = \langle \mathtt{Args}, \mathtt{R}_a, \mathtt{R}_s \rangle$ *be a BAF, the enriched BAF induced is defined as* $\overline{\Theta} = \langle \mathsf{Args}, \mathbb{R}_s, \mathbb{R}_a \rangle$*, where* Args *is the set of enriched arguments corresponding to arguments in* $\mathtt{Args}$*, and* $\mathbb{R}_a$ *and* $\mathbb{R}_s$ *are the attack and support relationships in* Args *that are induced by* $\mathtt{R}_a$ *and* $\mathtt{R}_s$*, respectively.*

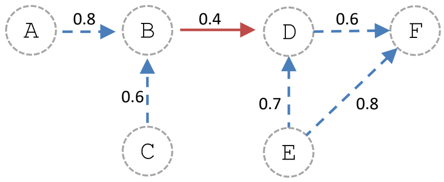

Figure 2: Similarity in the Bipolar Argumentation Framework (Fig 1)

Now, we introduce the *cohesion degree* of a set of supporting enriched arguments and the *controversy degree* associated with a set of attacking enriched arguments.

Definition 12 (Cohesion & Controversy degrees). *Given a set of enriched arguments* $\mathbb{S} \subseteq$ Args *and a context* $\mathbb{C}$*, let* $\mathsf{Sim}_{\mathbb{C}}$ *be a similarity degree function for* $\mathbb{C}$*, and* $\mathbb{R}_a^{\mathbb{S}} = \{(\mathrm{X}, \mathrm{Y}) \in \mathbb{R}_a \mid \mathrm{X}, \mathrm{Y} \in \mathbb{S}\}$ *be the subset of* $\mathbb{R}_a$ *restricted to the arguments of* $\mathbb{S}$ *and* $\mathbb{R}_s^{\mathbb{S}} = \{(\mathrm{X}, \mathrm{Y}) \in \mathbb{R}_s \mid \mathrm{X}, \mathrm{Y} \in \mathbb{S}\}$ *be the subset of* $\mathbb{R}_s$ *restricted to the arguments of* $\mathbb{S}$ *then we have:*

- *The* cohesion degree *for* $\mathbb{S}$*, denoted as* $\mathsf{Coh}_{\mathbb{C}}(\mathbb{S})$*, is defined as:*

$$\mathsf{Coh}_{\mathbb{C}}(\mathbb{S}) = \begin{cases} \beta_n & \textit{if}\, \mathbb{R}_s^{\mathbb{S}} = \{(\mathrm{A}_1, \mathrm{B}_1), \ldots, (\mathrm{A}_n, \mathrm{B}_n)\} \neq \emptyset \\ 0 & \textit{otherwise} \end{cases}$$

  where $\beta_1 = \mathsf{Sim}_{\mathbb{C}}(\mathrm{A}_1, \mathrm{B}_1)$ *and* $\beta_i = \oplus(\beta_{i-1}, \mathsf{Sim}_{\mathbb{C}}(\mathrm{A}_i, \mathrm{B}_i))$ *with* $2 \leq i \leq n$.
- *The* controversy degree *for* $\mathbb{S}$*, denoted as* $\mathsf{Cont}_{\mathbb{C}}(\mathbb{S})$*, is defined as:*

$$\mathsf{Cont}_{\mathbb{C}}(\mathbb{S}) = \begin{cases} \gamma_n & \textit{if}\, \mathbb{R}_a^{\mathbb{S}} = \{(\mathrm{A}_1, \mathrm{B}_1), \ldots, (\mathrm{A}_n, \mathrm{B}_n)\} \neq \emptyset \\ 0 & \textit{otherwise} \end{cases}$$

  where $\gamma_1 = \mathsf{Sim}_{\mathbb{C}}(\mathrm{A}_1, \mathrm{B}_1)$ *and* $\gamma_i = \otimes(\gamma_{i-1}, \mathsf{Sim}_{\mathbb{C}}(\mathrm{A}_i, \mathrm{B}_i))$ *with* $2 \leq i \leq n$.

Both $\mathsf{Coh}_{\mathbb{C}}(\cdot)$ and $\mathsf{Cont}_{\mathbb{C}}(\cdot)$ can be obtained independently using a recursive function instantiated with T-norms or T-conorms, in the same manner as with the similarity function $\mathsf{Sim}_{\mathbb{C}}$, depending on the user modeling intentions [22]. Next, we present our example where the similarity degree associated with each relationship was previously established (for more details, see [16]).

Example 3. *We continue with our abstract example, the graph in Figure 2 shows the similarity degree associated with the arguments in each relation. Intuitively, we can observe that the weakest support is between the arguments* C *and* B *and* D *and* F*. Based on the similarity degree obtained in each relation, we compute the cohesion coefficient associated with the set of supporting arguments (considering a product T-norm) and the controversy coefficient associated with attacking arguments (considering a max T-conorm). Thus, we have that:* $\mathsf{Coh}_{\mathbb{C}}(\{(\mathrm{E}, \mathrm{D}), (\mathrm{D}, \mathrm{F})\}) = 0.42$*;* $\mathsf{Coh}_{\mathbb{C}}(\{(\mathrm{A}, \mathrm{B})\}) = 0.8$*;* $\mathsf{Coh}_{\mathbb{C}}(\{(\mathrm{C}, \mathrm{B})\}) = 0.6$*;* $\mathsf{Coh}_{\mathbb{C}}(\{(\mathrm{E}, \mathrm{F})\}) = 0.8$*; and* $\mathsf{Cont}_{\mathbb{C}}(\{(\mathrm{B}, \mathrm{D})\}) = 0.4$*. Observe that, in this particular case, the cohesion associated with the support relation is analyzed considering the support chain presented in the argumentation model (see Figure 2). At the same time, the controversy measure is obtained by analyzing the pairs of attacking arguments.*

The enriched BAF $\overline{\Theta}$ will be extended to include the degrees just defined.

Definition 13 (Similarity-based BAF). *Let* $\overline{\Theta} = \langle \mathrm{Args}, \mathbb{R}_s, \mathbb{R}_a \rangle$ *be an enriched BAF and* $\mathbb{C}$ *a context, a* Similarity-Based Bipolar Argumentation Framework *(or S-BAF) is defined as a tuple* $\Phi = \langle \overline{\Theta}, \mathsf{Sim}_{\mathbb{C}}, \mathsf{Coh}_{\mathbb{C}}^{\overline{\Theta}}, \mathsf{Cont}_{\mathbb{C}}^{\overline{\Theta}} \rangle$*, where* $\mathsf{Sim}_{\mathbb{C}}$ *is a similarity degree function for enriched arguments in* Args*, and* $\mathsf{Coh}_{\mathbb{C}}^{\overline{\Theta}}$ *and* $\mathsf{Cont}_{\mathbb{C}}^{\overline{\Theta}}$ *are, respectively, the cohesion and controversy degree functions defined over* $\overline{\Theta}$ *in the context* $\mathbb{C}$*.*

When no confusion may arise, we will avoid mentioning the $\overline{\Theta}$ enriched BAF as a superscript of the cohesion and controversy degree operators, writing instead $\langle \overline{\Theta}, \mathsf{Sim}_{\mathbb{C}}, \mathsf{Coh}_{\mathbb{C}}, \mathsf{Cont}_{\mathbb{C}} \rangle$, making the notation more straightforward. Additionally, in S-BAF, the support and attack relations will have a particular interpretation since a threshold $\tau \in [0, 1]$ will be considered in the specification of the type of attack being analyzed. Consequently, the attacks in an S-BAF will be of two types: (i) *strong*, when the cohesion and controversy values associated with the attack are greater than the threshold τ; in this situation, we have *strong-direct attack*, *strong-supported attack*, and *strong-secondary attack*, or (ii) *weak*, if at least one of the values is less than τ; then, in this case, we have *weakly-direct attack*, *weakly-supported attack*, and *weakly-secondary attack*:

Definition 14 (Conflict-freeness and Safety properties in a S-BAF). *Let* $\Phi = \langle \overline{\Theta}, \mathsf{Sim}_{\mathbb{C}}, \mathsf{Coh}_{\mathbb{C}}, \mathsf{Cont}_{\mathbb{C}} \rangle$ *be a S-BAF, where* $\overline{\Theta} = \langle \mathrm{Args}, \mathbb{R}_s, \mathbb{R}_a \rangle$ *is the enriched BAF, and* $\tau \in [0, 1]$ *be a given threshold. Then:*

- $\mathbb{S}$ *is a* strongly-conflict-free *set iff there is no* $\mathrm{A}, \mathrm{B} \in \mathbb{S}$ *such that there exists a strong or weak attack from* A *to* B*.*
- $\mathbb{S}$ *is a* τ-conflict-free *set iff there is no* $\mathrm{A}, \mathrm{B} \in \mathbb{S}$ *such that there exists a strong attack from* A *to* B *and* $\mathsf{Cont}_{\mathbb{C}}(\mathbb{S}) > \tau$*.*
- $\mathbb{S}$ *is a* weakly-conflict-free *set iff there is no* $\mathrm{A}, \mathrm{B} \in \mathbb{S}$ *such that there exists a strong attack from* A *to* B*.*
- $\mathbb{S}$ *is a* strongly-safe *set iff there is no* $\mathrm{A} \in \mathrm{Args}$ *and no pair* $\mathrm{B}, \mathrm{C} \in \mathbb{S}$ *such that there exists a strong or weak attack from* B *to* A*, and either there is a sequence of support from* C *to* A*, or* $\mathrm{A} \in \mathbb{S}$*.*
- $\mathbb{S}$ *is* τ-safe *set iff there is no* $\mathrm{A} \in \mathrm{Args}$ *and no pair* $\mathrm{B}, \mathrm{C} \in \mathbb{S}$ *such that there exists a strong attack from* B *to* A*,* $\mathsf{Cont}_{\mathbb{C}}(\mathbb{S} \cup \{\mathrm{A}\}) > \tau$*, and either there is a sequence of support from* C *to* A *such that* $\mathsf{Coh}_{\mathbb{C}}(\{\mathrm{C}, \ldots, \mathrm{A}\}) > \tau$*, or* $\mathrm{A} \in \mathbb{S}$*.*
- $\mathbb{S}$ *is* weakly-safe *set iff there is no* $\mathrm{A} \in \mathrm{Args}$ *and no pair* $\mathrm{B}, \mathrm{C} \in \mathbb{S}$ *such that there is a strong attack from* B *to* A *and either there is a sequence of support from* C *to* A *such that* $\mathsf{Coh}_{\mathbb{C}}(\{\mathrm{C}, \ldots, \mathrm{A}\}) > \tau$*, or* $\mathrm{A} \in \mathbb{S}$*.*

In the following step, in [16] the authors extended the notions of defense for an argument with respect to a set of arguments.

Definition 15. *Let* $\mathbb{S} \subseteq \mathrm{Args}$ *be a set of arguments, and* $\mathrm{A} \in \mathrm{Args}$ *an argument. Then:*

- $\mathbb{S}$ *is a* strong defense *for* A *iff for all* $\mathrm{B} \in \mathrm{Args}$ *such that if* B *is a strong or weak attacker of* A *then there exists* $\mathrm{C} \in \mathbb{S}$ *where* C *is a strong attacker of* B*.*
- $\mathbb{S}$ *is a* weak defense *for* A *iff for all* $\mathrm{B} \in \mathrm{Args}$ *such that if* B *is a strong or weak attacker attacker of* A *then there exists* $\mathrm{C} \in \mathbb{S}$ *where* C *is a weak attacker attacker of* B*.*

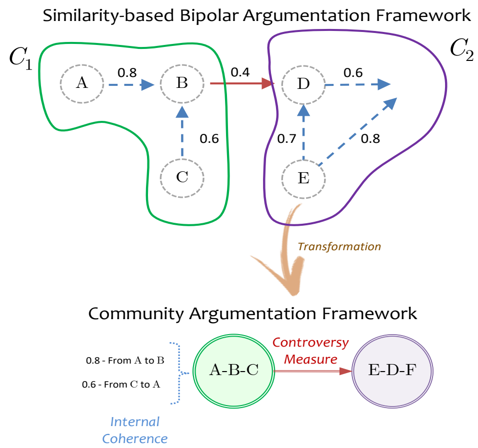

Figure 3: Communities in Bipolar Argumentation Frameworks

Next, we analyze our running example to obtain the different types of acceptable argument sets, where the properties of conflict-freeness and safety are considered.

Example 4. *We continue analyzing the Example 3 presented in Figure 2, introducing a threshold* $\tau = 0.48$. *With that addition we obtain: a weakly-direct attacks, with controversy coefficient lower than* τ*, are from* B *to* D*; a weakly-supported attacks are from* C *to* D *(since* $\mathsf{Coh}_{\mathbb{C}}(\{(\mathrm{C},\mathrm{B})\}) \geq \tau$ *and* $\mathsf{Cont}_{\mathbb{C}}(\{(\mathrm{B},\mathrm{D})\}) < \tau$*), from* A *to* D *(because* $\mathsf{Coh}_{\mathbb{C}}(\{(\mathrm{A},\mathrm{B})\}) \geq \tau$ *and* $\mathsf{Cont}_{\mathbb{C}}(\{(\mathrm{B},\mathrm{D})\}) < \tau$ *). Additionally, we have:* $S_1 = \{\mathrm{A},\mathrm{B},\mathrm{C}\}$*, is* strongly-conflict-free*,* strongly-safe*, because there are no elements in the set that simultaneously support and attack external arguments.*

In the next, we explore how these arguments can be organized into *communities* or coalitions.

3. Communities from Valuated-Similarity Coalitions

The *community* term definition is a complex task that is being approached from different perspectives [23, 24, 22, 25, 26]. For us, a *community* is a group of agents presenting different postures through a set of arguments expressing supporting and conflicting positions in a setting akin to a debate (see Figure 3). Support signifies a relationship based on common opinions, forming coalitions that represent communities. Additionally, we measure the internal cohesion of a community. Conflict between communities indicates the degree of controversy on specific statements. In a knowledge system that represents arguments and their conflicts and supports, we can find cohesive sets of arguments by employing a mechanism from Budan et al. [16] to gather conflict-free and support-related arguments. We also introduce a degree of controversy by considering the addition of attacks while maintaining coherence. This results in sets of arguments or stances forming a *community*—a group with consistent stances on a topic. We define a threshold with dual significance: it represents the maximum controversy a community can have without losing coherence or the minimum coherence required for a community to have a solid position. To identify communities, we use a bipolar argumentation graph where arguments are labeled with similarity degrees between related arguments.

Definition 16 (S-valued bipolar argumentation graph). *Given* $\Phi = \langle \overline{\Theta}, \mathsf{Sim}_{\mathbb{C}}, \mathsf{Coh}_{\mathbb{C}}, \mathsf{Cont}_{\mathbb{C}} \rangle$*, an* S-BAF *where* $\overline{\Theta} = \langle \text{Args}, \mathbb{R}_s, \mathbb{R}_a \rangle$ *is the underlying bipolar argumentation framework. An* s-valued Bipolar Argumentation Graph*, denoted* G_Φ*, is the argumentation graph where the nodes are the elements of* Args *and the arcs between nodes depict the* $\mathbb{R}_s$ *(dashed arcs) and* $\mathbb{R}_a$ *(full arcs) relationships, where the arcs are decorated with the similarity degree* $\mathsf{Sim}_{\mathbb{C}}$*.*

Now, it is necessary to revisit the concept of coalitions [15] to extend it by formalizing how a similarity degree can influence the support relations.

Definition 17 (S-coalitions). *Given an* S-BAF $\Phi = \langle \overline{\Theta}, \mathsf{Sim}_{\mathbb{C}}, \mathsf{Coh}_{\mathbb{C}}, \mathsf{Cont}_{\mathbb{C}} \rangle$*, where* $\overline{\Theta} = \langle \text{Args}, \mathbb{R}_s, \mathbb{R}_a \rangle$ *is the underlying bipolar argumentation framework, let* G_Φ *be the s-valued bipolar interaction graph over* Φ*, and* $\mathcal{C} \subseteq \text{Args}$ *be a set of enriched arguments. Then, we say that* $\mathcal{C}$ *is an s-coalition iff it is a maximally strongly-conflict-free set such that the sub-graph* G'_Φ *induced by* $\mathcal{C}$ *is connected only by support relations. We will denote as* $\mathcal{C}_\Phi$ *the set of coalitions obtained from* Φ*.*

Note that self-attacking arguments are disregarded in this approach according to the classic definition of a coalition where no attacks are permitted (Definition 5).

Proposition 1. *Each enriched argument, which is not self-attacking, belongs to an s-coalition.*

Given that a s-coalition is a set of enriched arguments, we can use the cohesion function established in Definition 12 to determine a cohesion measure associated with that s-coalition.

Definition 18 (Types of s-coalitions). *Let* $\Phi = \langle \overline{\Theta}, \mathsf{Sim}_{\mathbb{C}}, \mathsf{Coh}_{\mathbb{C}}, \mathsf{Cont}_{\mathbb{C}} \rangle$ *be an* S-BAF*,* $\mathcal{C} \in \mathcal{C}_\Phi$ *be a coalition obtained from* Φ*, and* $\tau \in [0, 1]$ *be a threshold. Then:*

- $\mathcal{C}$ *is a* strong-coalition *iff* $\mathsf{Coh}_{\mathbb{C}}(\mathcal{C}) = 1$.
- $\mathcal{C}$ *is a* τ*-coalition iff* $\tau \leq \mathsf{Coh}_{\mathbb{C}}(\mathcal{C}) < 1$.
- $\mathcal{C}$ *is a* weak-coalition *iff* $0 \leq \mathsf{Coh}_{\mathbb{C}}(\mathcal{C}) < \tau$.

Intuitively, in a *strong-coalition* the pieces of knowledge refer to the same aspects of the argumentation process, without conflict. In a τ*-coalition*, the opinions allude to the same values for each descriptor considered but contain some descriptors whose values differ. Lastly, in a *weak-coalition*, although the arguments do not contradict each other, they can refer to the same aspects differently, or some of them might refer to different aspects of the issue.

Example 5. *Continuing our Example 4, using a product T-norm to obtain the cohesion value, considering a* $\tau = 0.48$*, and analyzing the abstract argumentation framework represented in Figure 3, we have that the coalitions* $\mathcal{C}_1 = \{\text{A}, \text{B}, \text{C}\}$*,* $\mathcal{C}_2 = \{\text{E}, \text{D}, \text{F}\}$ *are* weak *because the* $\mathsf{Coh}_{\mathbb{C}}(\mathcal{C}_1) = 0.2 < \tau$*,* $\mathsf{Coh}_{\mathbb{C}}(\mathcal{C}_2) = 0.34 < \tau$*. However, if we choose a different function to obtain the cohesion of the sets, the max T-conorm for instance, we have that:* $\mathsf{Coh}_{\mathbb{C}}(\mathcal{C}_1) = 0.8 > \tau$*; the same occurs with* $\mathcal{C}_2$*. Under this perspective, all the communities are* τ-coalitions*. At the level of semantic analysis, by using a product T-norm to obtain the cohesion value, we find that the* $\mathcal{C}_1$ *and* $\mathcal{C}_2$ *are* sd-fragile-communities*, while in the second interpretation that relies on a max T-conorm, we conclude that* $\mathcal{C}_1$ *and* $\mathcal{C}_2$ *are* sd-moderate-communities*.*

When we find a strong support relation inside a coalition, it is natural to think that the cohesion associated with the supported arguments would be high.

Proposition 2. *Let $\Phi = \langle\overline{\Theta}, \mathsf{Sim}_{\mathbb{C}}, \mathsf{Coh}_{\mathbb{C}}, \mathsf{Cont}_{\mathbb{C}}\rangle$ be an S-BAF, where $\overline{\Theta}$ is the underlying bipolar argumentation framework $\overline{\Theta} = \langle \mathsf{Args}, \mathbb{R}_s, \mathbb{R}_a\rangle$, and let $\mathrm{A}, \mathrm{B} \in \mathsf{Args}$ two enriched arguments such that $(\mathrm{A}, \mathrm{B}) \in \mathbb{R}_s$, and $\mathbb{R}_a$ does not contain (A, B) or (B, A). If $\mathrm{A}\ \mathbb{R}_s\ \mathrm{B}$ is either a strong-support or weak-support relation, then there exists at least a weak-coalition containing both A and B.*

Now, it is necessary to introduce an attack relationship between conflicting coalitions:

Definition 19 (Internal attacks between s-coalitions). *Given the S-BAF $\Phi = \langle\overline{\Theta}, \mathsf{Sim}_{\mathbb{C}}, \mathsf{Coh}_{\mathbb{C}}, \mathsf{Cont}_{\mathbb{C}}\rangle$, where $\overline{\Theta} = \langle \mathsf{Args}, \mathbb{R}_s, \mathbb{R}_a\rangle$ is the underlying bipolar argumentation framework, let $\mathcal{C}_\Phi$ be the set of s-coalitions obtained from Φ, and $\mathcal{C}, \mathcal{C}' \in \mathcal{C}_\Phi$ be two s-coalitions. We will say that there exists an attack point from $\mathcal{C}$ to $\mathcal{C}'$ iff there are two enriched arguments $\mathrm{A} \in \mathcal{C}$ and $\mathrm{B} \in \mathcal{C}'$ such that $(\mathrm{A}, \mathrm{B}) \in \mathbb{R}_a$. We will denote as $\mathbb{R}_a^{[\mathcal{C},\mathcal{C}']}$ the set of all attacks points between $\mathcal{C}$ and $\mathcal{C}'$.*

Intuitively, it is possible to say that if there is an attack between two arguments that *belong* to two different coalitions, then it is natural to raise this conflict to the coalition level and define now an attack *between* these coalitions.

Definition 20 (S-coalitions Attacks). *Let $\Phi = \langle\overline{\Theta}, \mathsf{Sim}_{\mathbb{C}}, \mathsf{Coh}_{\mathbb{C}}, \mathsf{Cont}_{\mathbb{C}}\rangle$ be an S-BAF, where $\overline{\Theta} = \langle \mathsf{Args}, \mathbb{R}_s, \mathbb{R}_a\rangle$ is the underlying bipolar argumentation framework, let $\mathcal{C}_\Phi$ be the set of s-coalitions obtained from Φ. We will define the attack relation between s-coalitions derived from Φ, denoted $\mathbb{R}_a^{\mathcal{C}_\Phi}$, as*

$$\mathbb{R}_a^{\mathcal{C}_\Phi} = \{(\mathcal{C}, \mathcal{C}') \mid \mathcal{C}, \mathcal{C}' \in \mathcal{C}_\Phi \text{ and } \mathbb{R}_a^{[\mathcal{C},\mathcal{C}']} \neq \emptyset\}$$

Furthermore, it is interesting to study the strength of the attack from one coalition to another by considering the strength of the attacks that define the existing points of conflict. Formally:

Definition 21 (Strength of attack between s-coalitions). *Let $\mathcal{C}_\Phi$ be the set of s-coalitions obtained from Φ, $\mathcal{C}, \mathcal{C}' \in \mathcal{C}_\Phi$ be two s-coalitions, and $\mathbb{R}_a^{[\mathcal{C},\mathcal{C}']} = \{(\mathrm{A}_1, \mathrm{B}_1), \ldots, (\mathrm{A}_n, \mathrm{B}_n)\} \subseteq \mathbb{R}_a$ be the set of all attack points between $\mathcal{C}$ and $\mathcal{C}'$ with $\mathbb{R}_a^{[\mathcal{C},\mathcal{C}']} \neq \emptyset$. The attack strength, or* attack degree *between $\mathcal{C}$ and $\mathcal{C}'$, denoted $\mathsf{Str}_{\mathcal{C}}^{\Phi}(\mathcal{C}, \mathcal{C}')$, is defined as:*

$$\mathsf{Str}_{\mathcal{C}}^{\Phi}(\mathcal{C}, \mathcal{C}') = \delta_n,$$

where δ_n is defined as $\delta_1 = \mathsf{Sim}_{\mathbb{C}}(\mathrm{A}_1, \mathrm{B}_1)$ and $\delta_i = \otimes(\delta_{i-1}, \mathsf{Sim}_{\mathbb{C}}(\mathrm{A}_i, \mathrm{B}_i))$ with $2 \leq i \leq n$.

The attack degree can be obtained by instantiating the $\mathsf{Sim}_{\mathbb{C}}(\cdot, \cdot)$ similarity function with T-norms or T-conorms, considering the user modeling preferences. Once the attacks between s-coalitions are identified, and their strength is computed, we begin by using the attack degree to distinguish between strong and weak attacks. This classification can be employed to define different semantics by using different forms of acceptability.

Definition 22 (Classification of attacks between s-coalitions). *Given an S-BAF $\Phi = \langle\overline{\Theta}, \mathsf{Sim}_{\mathbb{C}}, \mathsf{Coh}_{\mathbb{C}}, \mathsf{Cont}_{\mathbb{C}}\rangle$, with $\overline{\Theta} = \langle \mathsf{Args}, \mathbb{R}_s, \mathbb{R}_a\rangle$ as the underlying bipolar argumentation framework, let $\mathcal{C}_\Phi$ be the set of s-coalitions obtained from Φ, $\mathcal{C}, \mathcal{C}' \in \mathcal{C}_\Phi$ be two s-coalitions such that $(\mathcal{C}, \mathcal{C}') \in \mathbb{R}_a^{\mathcal{C}_\Phi}$, and $\tau \in [0, 1]$ be a threshold. We say that:*

- $\mathcal{C}$ *strongly-attacks* $\mathcal{C}'$ *iff* $\mathsf{Coh}_{\mathbb{C}}(\mathcal{C}) \geq \tau$ *and* $\mathsf{Str}_{\mathcal{C}}^{\Phi}(\mathcal{C}, \mathcal{C}') \geq \tau$,
- $\mathcal{C}$ *weakly-attacks* $\mathcal{C}'$ *iff* $\mathsf{Coh}_{\mathbb{C}}(\mathcal{C}) < \tau$ *or* $\mathsf{Str}_{\mathcal{C}}^{\Phi}(\mathcal{C}, \mathcal{C}') < \tau$.

The previous definition formalizes the intuition that a strong attack considers two necessary elements: the strength of attack and the s-coalition internal cohesion measure applied to the set of the enriched arguments in the s-coalition. Now, we can formalize a new meta-argumentation framework to analyze a new kind of semantics concerning the set of communities.

Definition 23 (Meta-argumentation framework). *Given* $\Phi = \langle \overline{\Theta}, \mathsf{Sim}_{\mathbb{C}}, \mathsf{Coh}_{\mathbb{C}}, \mathsf{Cont}_{\mathbb{C}} \rangle$*, an* s-*BAF with* $\overline{\Theta} = \langle \mathrm{Args}, \mathbb{R}_s, \mathbb{R}_a \rangle$ *as the underlying bipolar argumentation framework, we define the* meta-argumentation framework associated with Φ*, as a 3-tuple* $\Omega^{\mathcal{C}} = \langle \mathcal{C}_\Phi, \mathbb{R}_a^{\mathcal{C}_\Phi}, \mathsf{Str}_{\mathcal{C}}^{\Phi} \rangle$*, where* $\mathcal{C}_\Phi$ *is the set of s-coalitions obtained from* Φ*,* $\mathbb{R}_a^{\mathcal{C}_\Phi}$ *is an attack relation between s-coalitions derived from* Φ*,* $\mathsf{Str}_{\mathcal{C}}^{\Phi}$ *is the attack strength function defined over* Φ*.*

Note that in the new meta-argumentation framework, the set $\mathcal{C}_\Phi$ of coalitions plays the role of the argument set, and the relation $\mathbb{R}_a^{\mathcal{C}_\Phi}$ represents the set of attacks. Henceforth, we will describe this meta-argumentation framework $\Omega^{\mathcal{C}} = \langle \mathcal{C}_\Phi, \mathbb{R}_a^{\mathcal{C}_\Phi}, \mathsf{Str}_{\mathcal{C}}^{\Phi} \rangle$ through a weighted directed graph $\mathsf{G}_{\mathcal{C}_\Phi}$, called meta-argumentation graph, with a unique kind of edge representing attacks between coalitions. Furthermore, each edge is assigned a weight representing the strength behind the attack it represents under the interpretation of attack strength.

Next, we will introduce the measure of controversy associated with a set of s-coalitions, where the different types of attacks are analyzed to specify how contradictory they are.

Definition 24 (Controversy degree for a s-coalition set). *Given a meta-argumentation framework* $\Omega^{\mathcal{C}} = \langle \mathcal{C}_\Phi, \mathbb{R}_a^{\mathcal{C}_\Phi}, \mathsf{Str}_{\mathcal{C}}^{\Phi} \rangle$*, where* $\mathcal{C}_\Phi$ *is a set of s-coalitions,* $\mathbb{R}_a^{\mathcal{C}_\Phi}$ *is an attack relation,* $\mathsf{Str}_{\mathcal{C}}^{\Phi}$ *is the attack strength function,* $\mathcal{S} \subseteq \mathcal{C}_\Phi$ *be a set of s-coalitions, and* $\mathbb{R}_a^{\mathcal{S}} = \{(\mathcal{C}_1, \mathcal{C}_2), \ldots, (\mathcal{C}_{n-1}, \mathcal{C}_n)\} \subseteq \mathbb{R}_a^{\mathcal{C}_\Phi}$*. The controversial measure for* $\mathcal{S}$*, denoted* $\mathsf{Cont}_{\mathcal{C}}^{\Phi}(\mathcal{S})$*, is defined as:*

$$\mathsf{Cont}_{\mathcal{C}}^{\Phi}(\mathcal{S}) = \begin{cases} \lambda_n & \textit{if}\, \mathbb{R}_a^{\mathcal{S}} \neq \emptyset \\ 0 & \textit{otherwise} \end{cases}$$

where $\lambda_1 = \mathsf{Str}_{\mathcal{C}}^{\Phi}(\mathcal{C}_1, \mathcal{C}_2)$ *and* $\lambda_i = \otimes(\lambda_{i-1}, \mathsf{Str}_{\mathcal{C}}^{\Phi}(\mathcal{C}_{i-1}, \mathcal{C}_i))$ *with* $2 \leq i \leq n$*.*

The instantiation of the controversy degree function is a design decision. Two possible choices are the T-norms and T-conorms.

Proposition 3. *Let* $\mathcal{S} \subseteq \mathcal{C}_\Phi$ *be a set of coalitions, and* $\mathbb{S} \subseteq \mathrm{Args}$ *be the enriched arguments involved in* $\mathcal{S}$*, then* $\mathsf{Cont}_{\mathbb{C}}(\mathbb{S}) = \mathsf{Cont}_{\mathcal{C}}^{\Phi}(\mathcal{S})$*.*

Given that the controversy associated with a set of coalitions is the same as the controversy associated with the set of enriched arguments involved, the previous result establishes a common point between the s-BAF and the meta-argumentation framework. Now, based on the semantic analysis done in [16], we introduce the notions of *conflict-free* s-coalition sets in our meta-argumentation framework $\Omega^{\mathcal{C}}$. Thus, it is possible to determine the set of communities that can coexist within an argumentative model.

Definition 25 (Conflict-freeness in $\Omega^{\mathcal{C}}$). *Given a meta-argumentation framework* $\Omega^{\mathcal{C}} = \langle \mathcal{C}_\Phi, \mathbb{R}_a^{\mathcal{C}_\Phi}, \mathsf{Str}_c^\Phi \rangle$, *where* $\mathcal{C}_\Phi$ *is the set of s-coalitions,* $\mathbb{R}_a^{\mathcal{C}_\Phi}$ *is an attack relation between s-coalitions, and* Str_c^Φ *the strength of attack function. Given a controversy degree function* Cont_c^Φ *defined over* $\Omega^{\mathcal{C}}$. *Let* $\mathcal{S} \subseteq \mathcal{C}_\Phi$ *be a subset of coalitions, and* τ *be a threshold. Then:*

- $\mathcal{S}$ *is a* strongly-conflict-free *set iff there is no* $\mathcal{C}_1, \mathcal{C}_2 \in \mathcal{S}$ *such that there exists a strong or weak attack from* $\mathcal{C}_1$ *to* $\mathcal{C}_2$.
- $\mathcal{S}$ *is a* τ-conflict-free *set iff there is no* $\mathcal{C}_1, \mathcal{C}_2 \in \mathcal{S}$ *such that there exists a strong attack from* $\mathcal{C}_1$ *to* $\mathcal{C}_2$, *and* $\mathsf{Cont}_c^\Phi(\mathcal{S}) \leq \tau$.
- $\mathcal{S}$ *is a* weakly-conflict-free *set iff there is no* $\mathcal{C}_1, \mathcal{C}_2 \in \mathcal{S}$ *such that there exists a strong attack from* $\mathcal{C}_1$ *to* $\mathcal{C}_2$.

The following proposition establishes the semantic connections between the meta-argumentation framework dealing with coalitions of arguments and the subjacent similarity-based argumentation framework.

Proposition 4. *Let* $\Omega^{\mathcal{C}} = \langle \mathcal{C}_\Phi, \mathbb{R}_a^{\mathcal{C}_\Phi}, \mathsf{Str}_c^\Phi \rangle$ *be the meta-argumentation framework associated with* Φ, Cont_c^Φ *a controversy degree function defined over* $\Omega^{\mathcal{C}}$, $\{\mathcal{C}_1, \ldots, \mathcal{C}_n\}$ *be a finite set of coalitions, and* $\tau \in [0, 1]$ *be a threshold. Then:*

i) $\{\mathcal{C}_1, \ldots, \mathcal{C}_n\}$ *is strongly-conflict-free for* $\Omega^{\mathcal{C}}$ *iff* $\mathcal{C}_1 \cup \cdots \cup \mathcal{C}_n$ *is strongly-conflict-free for* Φ.
ii) $\{\mathcal{C}_1, \ldots, \mathcal{C}_n\}$ *is strongly-conflict-free for* $\Omega^{\mathcal{C}}$ *iff* $\mathcal{C}_1 \cup \cdots \cup \mathcal{C}_n$ *is strongly-safe for* Φ.
iii) If $\mathcal{C}_1 \cup \cdots \cup \mathcal{C}_n$ *is* τ*-conflict-free for* Φ, *then* $\{\mathcal{C}_1, \ldots, \mathcal{C}_n\}$ *is* τ*-conflict-free for* $\Omega^{\mathcal{C}}$.
iv) If $\mathcal{C}_1 \cup \cdots \cup \mathcal{C}_n$ *is at least* τ*-safe for* Φ *then* $\{\mathcal{C}_1, \ldots, \mathcal{C}_n\}$ *is* τ*-conflict-free for* $\Omega^{\mathcal{C}}$.
v) $\{\mathcal{C}_1, \ldots, \mathcal{C}_n\}$ *is weakly-conflict-free for* $\Omega^{\mathcal{C}}$ *iff* $\mathcal{C}_1 \cup \cdots \cup \mathcal{C}_n$ *is weakly-conflict-free for* Φ.
vi) $\{\mathcal{C}_1, \ldots, \mathcal{C}_n\}$ *is weakly-conflict-free for* $\Omega^{\mathcal{C}}$ *iff* $\mathcal{C}_1 \cup \cdots \cup \mathcal{C}_n$ *is at least weakly-safe for* Φ.

The following example exercises the concepts just introduced:

Example 6. *Continuing the analysis of Example 5, and recalling that the threshold set is* $\tau = 0.48$ *we have that there is a conflict point between the s-coalitions* $\mathcal{C}_1$ *and* $\mathcal{C}_2$*: the pair* (B, D). *In this case,* $\mathsf{Str}_c^\Phi(\mathcal{C}_1, \mathcal{C}_2) = 0.4 < \tau$, *therefore,* $\mathcal{C}_1$ *weakly-attacks* $\mathcal{C}_2$.

The characterization of the attack relationship between coalitions and considering the associated strength of attacks allows us to establish the following property.

Proposition 5. *Let* $\mathcal{C}_1, \mathcal{C}_2 \in \mathcal{C}_\Phi$ *be two s-coalitions. If* $\mathcal{C}_1$ *and* $\mathcal{C}_2$ *are two disjoint s-coalitions that are connected by the attack relation, then there exists at least a weak-attack between* $\mathcal{C}_1$ *and* $\mathcal{C}_2$.

Now, we will introduce the notions of defense for coalitions by extrapolating from the defense relationship between the arguments gathered in the coalitions.

Definition 26. *Let* $\Omega^{\mathcal{C}} = \langle \mathcal{C}_\Phi, \mathbb{R}_a^{\mathcal{C}_\Phi}, \mathsf{Str}_c^\Phi \rangle$ *be the meta-argumentation framework associated with* Φ, Cont_c^Φ *a controversy degree function defined over* $\Omega^{\mathcal{C}}$, $\mathcal{S} \subseteq \mathcal{C}_\Phi$ *be a set of coalitions over* Φ, *and* $\mathcal{C}_1 \in \mathcal{C}_\Phi$ *a s-coalitions. Then:*

- *The set* $\mathcal{S}$ *is a* strong defense *for* $\mathcal{C}_1$ *iff for all* $\mathcal{C}_2 \in \mathcal{C}_\Phi$ *such that if* $\mathcal{C}_2$ *is a strong or weak attacker of* $\mathcal{C}_1$ *then there exists* $\mathcal{C}_3 \in \mathcal{S}$ *where* $\mathcal{C}_3$ *is a strong attacker of* $\mathcal{C}_2$.

– *The set $\mathcal{S}$ is a* weak defense *for $\mathcal{C}_1$ iff for all $\mathcal{C}_2 \in \mathcal{C}_\Phi$ such that if $\mathcal{C}_2$ is a strong or weak attacker of $\mathcal{C}_1$ then there exists $\mathcal{C}_3 \in \mathcal{S}$ where $\mathcal{C}_3$ is a weak attacker of $\mathcal{C}_2$.*

Example 7. *Continuing with the running example, we observe that $\mathcal{C}_1$ does not receive any attack. Furthermore, there is no defense for the attacks of $\mathcal{C}_1$ to $\mathcal{C}_2$.*

The concept of a coalition introduced provides a valuable framework for defining communities within argumentation-supported debates. It is worth noting that the principles of *conflict-freeness* and *safety* can also be applied to these communities. These characteristics are particularly useful for analyzing intricate debates, such as those often encountered on social networks. This expanded formal argumentation theory equips us with tools to enhance the analysis of debates.

4. Related Works, Conclusion and Future Work

Puertas *et al.* [27] used Twitter data to detect social communities. They employed expert knowledge, computational linguistics, and AI techniques to extract vocabulary-based community features, and explore language-related relationships within the social network. One notable difference from our work is that the s-coalition detection method here doesn't focus on individual opinions. Instead, it considers relationships among opinions to find and characterize communities or coalitions. However, both approaches involve language-related features. Puertas et al. employ term frequency techniques, while our method relies on enriched arguments using descriptors. Lenine [28] propose a work to categorize coalition into three types: conceptual (math-based), quasi-conceptual (deductive empirical), and extrapolative (statistical). Our work falls into the first category, building on s-BAF [16]. Another notable approach by Vassiliades et al. [29] introduces an Abstract Argumentation Framework (AF) with domain-specific arguments, allowing the determination of argument acceptance scope. It differs from Budan et al. [16] in considering only attack relations, and using different argument modeling tools. Bistarelli at al. [30] detail a set of semantics based on weighted defences. A threshold α obtained from collective attacks received by an argument is used to define a threshold representing the conflict permitted in the AF semantics without alter its coherence, providing flexibility. In our approach, the threshold is a given value that fulfills the same role, but it is applied both direct attacks and attacks that involve the support relationships necessary to express the coalition strength.

The detection of communities has broad applications today, such as identifying common research areas in collaboration networks, targeting like-minded users for marketing, or predictions in political areas. This work introduces a novel approach to identify meta-structures (coalitions) based on the similarity between supported arguments. We utilize similarity to characterize attacks between coalitions and assess controversy. However, these methods rely on specialized argument mining techniques and argument descriptors, and computational costs vary based on tree traversal. Future research directions include implementing this approach for coalition detection and community modeling, particularly in decision support systems with user preferences. Additionally, applying the proposed conceptualization to enhance argument schemes, especially those involving analogies represents an intriguing avenue for development.

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 grants PGI 24/N046 and PGI 24/ZN057), Agencia Nacional de Promoción Científica y Tecnológica under grants PICT-2018-0475 (PRH-2014-0007) and PICT-2020-SERIEA-01481, and CONICET under grants PIP 11220200101408CO and PIP 11220170100871CO. The authors also acknowledge support by the Spanish project PID2022-139835NB-C21 funded by MCIN/AEI/10.13039/501100011033, MCIN/AEI/ 10.13039/501100011033 and by "European Union NextGenerationEU/PRTR".

References

[1] D. Furman, S. Marro, C. Cardellino, D. Popa, L. A. Alemany, You can simply rely on communities for a robust characterization of stances, Florida Artificial Intelligence Research Society 34 (2021).
[2] D. Jin, X. Wang, M. Liu, J. Wei, W. Lu, F. Fogelman-Soulié, Identification of generalized semantic communities in large social networks, IEEE Transactions on Network Science and Engineering 7 (2020) 2966–2979.
[3] O. Fraisier, G. Cabanac, Y. Pitarch, R. Besançon, M. Boughanem, Stance classification through proximity-based community detection, in: Proc of the 29th on Hypertext and Social Media, Association for Computing Machinery, New York, NY, USA, 2018, p. 220–228.
[4] M. Lai, V. Patti, G. Ruffo, P. Rosso, Stance evolution and twitter interactions in an italian political debate, in: International Conference on Applications of Natural Language to Information Systems, Springer, 2018, pp. 15–27.
[5] M. Dias, K. Becker, INF-UFRGS-OPINION-MINING at SemEval-2016 task 6: Automatic generation of a training corpus for unsupervised identification of stance in tweets, in: Proceedings of the 10th International Workshop on Semantic Evaluation (SemEval-2016), Association for Computational Linguistics, San Diego, California, 2016, pp. 378–383.
[6] E. W. Pamungkas, V. Basile, V. Patti, Stance classification for rumour analysis in Twitter: Exploiting affective information and conversation structure, in: A. Cuzzocrea, F. Bonchi, D. Gunopulos (Eds.), Proc. of the CIKM 2018 Workshops co-located with 27th ACM Int. Conf. on Information and Knowledge Management (CIKM 2018), Torino, Italy, volume 2482 of *CEUR Workshop Proceedings*, CEUR-WS.org, 2018.
[7] A. Misra, B. Ecker, T. Handleman, N. Hahn, M. Walker, A semi-supervised approach to detecting stance in tweets, arXiv preprint arXiv:1709.01895 (2017).
[8] P. D. Budán, M. G. Escañuela Gonzalez, M. C. D. Budán, M. V. Martinez, G. R. Simari, Strength in coalitions: Community detection through argument similarity, Argument & Computation 2023 (2023) 1–51. doi:10.3233/AAC-220006.
[9] T. Bench-Capon, P. Dunne, Argumentation in artificial intelligence, AI 171 (2007) 619–641.
[10] P. Besnard, A. Hunter, Elements of argumentation, volume 47, MIT press, 2008.
[11] I. Rahwan, G. Simari, Argumentation in artificial intelligence, Springer, 2009.

[12] P. Baroni, D. Gabbay, M. Giacomin, L. van der Torre, Handbook of Formal Argumentation, volume 1, College Publications, London, 2018.
[13] D. Gabbay, M. Giacomin, G. R. Simari, M. Thimm, Handbook of Formal Argumentation, volume 2, College Publications, London, 2021.
[14] P. Dung, On the acceptability of arguments and its fundamental role in nonmonotonic reasoning, logic programming and n-person games, Artif. Intell. 77 (1995) 321–357.
[15] C. Cayrol, M.-C. Lagasquie-Schiex, On the acceptability of arguments in bipolar argumentation frameworks, in: Symbolic and quantitative approaches to reasoning with uncertainty, Springer, 2005, pp. 378–389.
[16] P. D. Budán, M. G. Escañuela Gonzalez, M. C. D. Budán, M. V. Martinez, G. R. Simari, Similarity notions in bipolar abstract argumentation, Argum. Comput. 11 (2020) 103–149.
[17] L. Amgoud, C. Cayrol, M.-C. Lagasquie-Schiex, On the bipolarity in argumentation frameworks., in: Nmr, volume 4, 2004, pp. 1–9.
[18] C. Cayrol, M.-C. Lagasquie-Schiex, Coalitions of arguments: A tool for handling bipolar argumentation frameworks, International Journal of Intelligent Systems 25 (2010) 83–109.
[19] D. Lin, An information-theoretic definition of similarity, in: J. W. Shavlik (Ed.), Proceedings of the Fifteenth International Conference on Machine Learning (ICML 1998), USA, Morgan Kaufmann, 1998, pp. 296–304.
[20] A. Huang, Similarity measures for text document clustering, in: Proceedings of the sixth new zealand computer science research student conference (NZCSRSC2008), Christchurch, New Zealand, 2008, pp. 49–56.
[21] Q. Le, T. Mikolov, Distributed representations of sentences and documents, in: International Conference on Machine Learning, 2014, pp. 1188–1196.
[22] E. P. Klement, R. Mesiar, E. Pap, Triangular Norms, volume 8 of *Trends in Logic,* Studia Logica Library, reprint ed., Springer, 2010.
[23] T. K. Bradshaw, The post-place community: Contributions to the debate about the definition of community, Community Development 39 (2008) 5–16.
[24] S. Fortunato, Community detection in graphs, Physics reports 486 (2010) 75–174.
[25] J. Prodnik, Post-fordist communities and cyberspace: A critical approach, in: Cybercultures, Brill, 2012, pp. 73–100.
[26] N. Chouchani, M. Abed, Online social network analysis: detection of communities of interest, Journal of Intelligent Information Systems 54 (2020) 5–21.
[27] E. Puertas, L. G. Moreno-Sandoval, J. Redondo, J. A. Alvarado-Valencia, A. Pomares-Quimbaya, Detection of sociolinguistic features in digital social networks for the detection of communities, Cognitive Computation 13 (2021) 518–537.
[28] E. Lenine, Modelling coalitions: From concept formation to tailoring empirical explanations, Games 11 (2020) 55.
[29] A. Vassiliades, T. Patkos, G. Flouris, A. Bikakis, N. Bassiliades, D. Plexousakis, Abstract argumentation frameworks with domain assignments, in: Proc. IJCAI, 2021, pp. 2076–2082.
[30] S. Bistarelli, F. Rossi, F. Santini, A novel weighted defence and its relaxation in abstract argumentation, International Journal of Approximate Reasoning 92 (2018) 66–86.