24th European Conference on Artificial Intelligence - ECAI 2020
Santiago de Compostela, Spain

Embedding the Trust Degrees of Agents in Abstract Argumentation

Bettina Fazzinga[1] and **Sergio Flesca, Filippo Furfaro**[2]

Abstract. We propose a new paradigm for reasoning over abstract argumentation frameworks where the trustworthiness of the agents is taken into account. In particular, we study the problems of computing the minimum trust degree τ^* such that, if we discard the arguments said only by agents whose trust degree is not greater than τ^*, a given set of arguments S (resp., argument a), that is not necessarily an extension (resp., (credulously) accepted) over the original argumentation framework, becomes an extension (resp., (credulously) accepted). Solving these problems helps reason on how the robustness of sets of arguments and single arguments depends on what is considered trustworthy or not. We thoroughly characterize the computational complexity of the considered problems, along with some variants where a different aggregation mechanism is used to decide the arguments to discard.

1 INTRODUCTION

Since their introduction [16], *Abstract Argumentation Frameworks* (AAFs) have been a popular paradigm for reasoning on disputes between agents. An AAF models a dispute in terms of a directed graph, whose nodes are the *arguments* proposed by the agents participating the dispute, and whose edges represent *attack* relationships: an attack from an argument a to an argument b represents the fact that a undercuts/rebuts/undermines b. AAFs are used to reason on sets of arguments and/or single arguments to decide whether they are "robust". In particular, over a given AAF F, two fundamental problems have been studied in the literature:

- VER(F, S): Is the set of arguments S an "*extension*" of F (i.e., a set of arguments that can be considered "robust")?
- ACC(F, a): Is a a "*(credulously) accepted*" argument of F (i.e., does a belong to some extension of F)?

Herein, in order to decide on the "robustness" of a set of arguments, different semantics have been introduced, such as *admissible*, *preferred*, etc. For instance, a set S is an *admissible* extension if it is "*conflict-free*" (i.e., there is no attack between arguments in S), and every argument attacking arguments in S is counterattacked by an argument in S.

In order to make AAFs suitable for modeling disputes in scenarios with different characteristics, several variants have been proposed. In particular, *Weighted AAFs* are a variant of AAFs where the arguments and/or the attacks can be associated with weights. In this paper, we introduce *Trust-aware AAFs* (T-AAFs), a form of weighted AAFs where the weights are assigned to arguments and are representative of the trustworthiness of the agents who propose the arguments. The following example is inspired by the scenario of an e-commerce site whose customers share their reviews, where T-AAFs find a natural application.

Example 1 *Ann, Mary, Carl and John are reviewing a notebook. Their reviews contain the following six arguments:*
a='Since it contains up-to-date components, it is expensive'
b='Nowadays, it is easy to find cheap up-to-date components. Therefore, that aspect does not imply the price.'
c='Since its brand is not high quality, it does not contain up-to-date components'
d='Since its battery is lightweight, it is lightweight overall'
e='It is heavy'
f='The battery is very heavy'.
Figure 1 shows the corresponding argumentation graph, properly augmented to highlight who-claims-what (for instance, a and d are claimed by Mary, and e is claimed by both Ann and Carl). The numbers in brackets represent the trustworthiness scores, on a scale of 1 *to* 10*, assigned to the agents on the basis of their past reviews.*

As a matter of fact, reasoning on reviews (as in the scenario of the above example) is a hot topic attracting the interest of the research community [28, 35], owing to the popularity of commercial sites. Herein, the customers who publish their comments/reviews are often associated with reputation scores measuring their trustworthiness. In this context, reasoning on extensions is useful, since the fact that a set of arguments is an extension means that it provides a reasonable summary of the main features and critical aspects of the reviewed object. Analogously, reasoning on the acceptance of an argument helps understand if it can be reasonably considered representative of the object. Now, in the T-AAF F of Example 1, argument a does not belong to any extension. However, a is proposed by Mary, who has a high trust degree. Thus, the analyst can benefit from knowing that, although a is not accepted, it becomes accepted in the AAF F^τ (with $\tau = 2$) obtained from F by discarding what said only by agents whose trust degree is $\leq \tau$. This means that the analyst can choose now to consider a a robust argument, given that F^τ does not contain

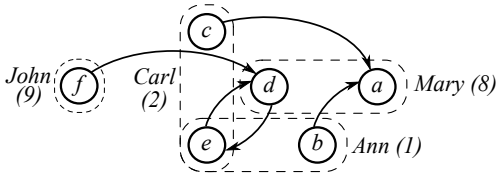

Figure 1: A T-AAF F, where the agents are assigned a trust degree

[1] ICAR - CNR, Rende (CS) - ITALY, email: fazzinga@icar.cnr.it
[2] DIMES - University of Calabria, Rende (CS) - ITALY, email: {flesca, furfaro}@dimes.unical.it

what said by agents with "low" trust degrees (we recall that we are in a scale from 1 to 10). Analogously, even if $S = \{a, f\}$ is not an (admissible) extension in F, it can be somehow considered a reasonable summary of the reviews, since it is an extension over the same F^τ. In general, denoting as "τ-extension" (resp., "τ-accepted") a set (resp., an argument) that is an extension (resp., accepted) over F^τ, the following two problems over a T-AAF F are of interest to the analyst:

- MIN-TVER$^\sigma(F, S)$: What is the minimum trust degree τ such that the set S is a τ-extension over F?
- MIN-TACC$^\sigma(F, a)$: What is the minimum trust degree τ such that the argument a is τ-accepted over F?

The rationale of searching for the minimum trust degree τ such that S is an extension (or a accepted) in F^τ is twofold. On the one hand, we aim at preserving as many as possible agents participating the dispute; on the other hand, we aim at understanding if discarding what said by some agents makes S an extension (or a accepted). In particular, the removal of agents is done by discarding them in ascending order of trust degree (so that less trustworthy agents are "sacrified" before), and in a "fair" manner: when an agent u is discarded, all the agents whose trustworthiness is equal to or less than u are discarded too.

Contribution. These points summarize our contribution:

- We characterize the complexity of MIN-TVER$^\sigma(F, S)$ and MIN-TACC$^\sigma(F, a)$ in terms of both upper and lower bounds. Since they are functional problems, we describe the complexity in terms of classes suitable for functional problems (such as *FP* and $FP^{\mathcal{C}}$), that are briefly reviewed in the last paragraph of Section 2. We also address the decisional counterparts TVER$^\sigma(F, S, \tau^*)$ and TACC$^\sigma(F, a, \tau^*)$, and discuss their practical relevance. A synopsis of the results is reported in Table 1;
- We consider two variants of MIN-TVER$^\sigma(F, S)$ and MIN-TACC$^\sigma(F, a)$ where the threshold-based reasoning is replaced by a different mechanism for discarding agents/arguments from the dispute (based on the sum of the trust degrees of the discarded agents). We discuss the impact of these changes on the complexity and their relationship with the literature of weighted AAFs.

σ	VER$^\sigma$ and TVER$^\sigma$	ACC$^\sigma$ and TACC$^\sigma$	MIN-TVER$^\sigma$	MIN-TACC$^\sigma$
ad,*st*,*co*	*P*	*NP-c*	*FP*	$FP^{NP[\log n]}$-*c*
gr	*P*	*P*	*FP*	*FP*
pr	*coNP-c*	*NP-c*	$FP^{NP[\log n]}$-*c*	$FP^{NP[\log n]}$-*c*

Table 1: Summary of the computational complexities

2 PRELIMINARIES

Abstract Argumentation Framework [16]. An *Abstract Argumentation Framework* (*AAF*) F is a pair $\langle A, D\rangle$, where A is a finite and non-empty set, whose elements are called *arguments*, and $D \subseteq A \times A$ is a binary relation over A, whose elements are called *attacks*. The graph having A and D as set of nodes and edges, respectively, is called *argumentation graph* of F. Given $a, b \in A$, we say that a *attacks* b iff $(a, b) \in D$. A set $S \subseteq A$ *attacks* an argument $b \in A$ iff there is $a \in S$ that *attacks* b. An argument a *attacks* S iff $\exists b \in S$ attacked by a.

A set $S \subseteq A$ of arguments is said to be *conflict-free* if there are no $a, b \in S$ such that a *attacks* b. An argument a is said to be *acceptable w.r.t.* $S \subseteq A$ iff $\forall b \in A$ such that b *attacks* a, there is $c \in S$ such that c *attacks* b.

Extension. An *extension* is a set of arguments that is considered "reasonable" according to some semantics. In particular, we consider the following semantics from the literature:

- *admissible* (*ad*): S is an *admissible extension* iff S is conflict-free and its arguments are acceptable w.r.t. S;
- *stable* (*st*): S is a *stable extension* iff S is conflict-free and S attacks each argument in $A \setminus S$;
- *complete* (*co*): S is a *complete extension* iff S is admissible and every argument acceptable w.r.t. S is in S;
- *grounded* (*gr*): S is a *grounded extension* iff S is a minimal (w.r.t. $\subseteq$) complete set of arguments;
- *preferred* (*pr*): S is a *preferred extension* iff S is a maximal (w.r.t. $\subseteq$) complete set of arguments.

Accepted arguments. An argument a is (credulously) *accepted* under a semantics σ iff a belongs to some σ extension of F. In some sense, checking the acceptability of an argument is a way of deciding whether a represents a robust point of view in the discussion modeled by F.

Classical problems: VER ***and*** ACC. Given an AAF F, a semantics σ, a set of arguments S and an argument a, the fundamental problems of verifying whether S is a σ extension and whether a is (credulously) accepted (under σ) will be denoted as VER$^\sigma(F, S)$ and ACC$^\sigma(F, a)$, respectively. These problems have been widely studied in the literature, and their complexity is reported in Table 1 [19, 14, 17, 10].

Functional complexity classes. Most of the results stated in the paper refer to *functional* complexity classes, that are more suitable for characterizing the complexity of MIN-TVER$^\sigma(F, S)$ and MIN-TACC$^\sigma(F, a)$, since they are intrinsically functional problems.

FP is the class of the functional problems that can be solved by a deterministic Turing machine in polynomial time (w.r.t. the size of the input of the problem). Basically, while the Turing machines behind the problems in P return a binary result, those behind FP return string of bits. In this paper, we in particular deal with the classes $FP^{||\mathcal{C}}$ and $FP^{\mathcal{C}[\log n]}$, where $\mathcal{C}$ is a decisional complexity class (such as NP or Σ_p^2). Specifically, $FP^{||\mathcal{C}}$ is the class of functions computable by a polynomial-time Turing machine with calls to an oracle for the class $\mathcal{C}$, where the calls are "non-adaptive" (this is equivalent to saying that the oracle invocations can take place in parallel). The meaning of "*call to an oracle for* $\mathcal{C}$" is that the computation performed by the oracle is considered as it requires constant time (thus, calling the oracle produces no computational overhead w.r.t. the complexity of the machine that performs the call). Analogously, $FP^{\mathcal{C}[\log n]}$ is the class of the function problems that can be solved by a deterministic Turing machine in polynomial time, where at most $\log n$ adaptive queries can be posed to an oracle for the class $\mathcal{C}$.

3 TRUST-AWARE AAFs

We here introduce an extension of AAFs that takes into account the trustworthiness of the agents who propose the arguments. This extension is a form of *weighted AAF*: specifically, we associate the weights to the arguments and the semantics of weights is that they represent trust degrees.

Let $F = \langle A, D\rangle$ be an AAF and U the set of agents proposing the arguments in A. The association between agents and arguments is modeled by the function $\omega : U \to 2^A$, returning, for each agent u, the set of arguments proposed by u. We assume that every argument is proposed by at least one agent, and the same argument can be proposed by several agents. The set of agents proposing an argument a is denoted as $\omega^{-1}(a)$.

We assume the presence of an *agent trust function* τ^U assigning to each agent $u \in U$ a trust degree $\tau^U(u)$, i.e., a **positive integer** providing a measure of how trustworthy u is considered. τ^U can be used to derive an assignment of trust values to the arguments. In fact, since a measure of the trustworthiness of an argument a should take into account how much trust can be put on the premises of a and the consequentiality of its claim from these premises, it is reasonable to derive a trust degree of a from the trust degrees of the agents who propose a. In this regard, we model the trustworthiness of arguments with the *argument trust function* T^{U,ω,τ^U} (or, more simply, T) assigning to each argument a the **positive integer** equal to the maximum trust degree of the agents that propose a [3], i.e., $T(a) = \max_{u\in\omega^{-1}(a)} \tau^U(u)$.

For the sake of simplicity, and without loss of generality, from now on we will only implicitly consider the set of users U and the functions ω and τ^U, and we will explicitly consider only the argument trust function T implied by them.

Given an abstract argumentation framework $\langle A, D\rangle$ and an argument trust function T over A, we call the triple $F = \langle A, D, T\rangle$ a *Trust-aware Abstract Argumentation Framework* (T-AAF). We denote as $\mathcal{T}(F)$ the set of distinct trust degrees of F's arguments augmented with 0.

Example 2 *(Continuing Example 1 - Fig. 1) From the users' trust degrees, we have* $T(e) = \max(\tau^U(\text{Ann}), \tau^U(\text{Carl})) = 2$, $T(a) = T(d) = 8$, $T(c) = 2$, $T(b) = 1$, $T(f) = 9$. *Moreover, we have:* $\mathcal{T}(F) = \{0, 1, 2, 8, 9\}$.

τ-restrictions, τ-extensions and τ-accepted arguments. Let $F = \langle A, D, T\rangle$ be a T-AAF, τ a trust value, and σ a semantics. We define the *τ-restriction* of F the T-AAF $F^\tau = \langle A', D', T'\rangle$ where $A' = \{a |\, a \in A \wedge T(a) > \tau\}$, $D' = D \cap (A' \times A')$, and T' is the restriction of T over A'. That is, F^τ is the T-AAF consisting of all and only the arguments of F with trust greater than τ and of all and only the attacks in F between these arguments. Basically, considering the τ-restriction of F means considering τ as a threshold, and then taking into account only what said by the agents whose trust degree is greater than τ, while discarding what said *only* by agents whose trust degree is $\leq \tau$. Observe that $F^\tau = F$ when $\tau = 0$, since the trust function assigns only positive values.

We now introduce the natural generalization of the classical notions of *extension* and *accepted argument* (reviewed in Section 2) to the case of T-AAFs. Given a T-AAF F and a trust degree τ, we define "τ-extension of F" (shorthand for "*trusted extension with trust level* τ") under the semantics σ any set of arguments that is an extension of F^τ under σ. Basically, a τ-extension for F is a set of arguments that meets the conditions of the semantics σ when discarding the arguments proposed by agents whose trust degree is $\leq \tau$. In turn, an argument a of F is said to be "*τ-accepted*" (shorthand for "*trustingly accepted with trust level* τ") under σ if a belongs to at least one τ-extension under σ. The rationale of τ-acceptance is analogous to τ-extension: An argument a may not be accepted in the classical sense, but it can still be τ-accepted for some τ, meaning that a turns out to be a "robust" argument when discarding what said by users not sufficiently trustworthy (w.r.t. the threshold τ). The reason is that the removal of arguments (and the consequent removal of the attacks involving the removed arguments) can change the number of extensions and their composition.

Example 3 *(Continuing examples 1, 2) Under σ = ad, $\{c, f\}$ is a τ-extension even with $\tau = 0$, while $\{a, f\}$ is a τ-extension for $\tau = 2$ but not for lower degrees in $\mathcal{T}(F)$. Under all the considered semantics, there is no $\tau \in \mathcal{T}(F)$ such that d is τ-accepted, while a is τ-accepted for $\tau = 2$, but not for any lower $\tau \in \mathcal{T}(F)$.*

4 PROBLEM STATEMENT

We consider the following problems over a given T-AAF F and under a semantics σ:

- MIN-TVER$^\sigma(F, S)$: Given a set S of arguments of F, what is the minimum trust degree τ in $\mathcal{T}(F)$ (if exists) such that S is a τ-extension of F under σ?
- MIN-TACC$^\sigma(F, a)$: Given an argument a of F, what is the minimum trust degree τ in $\mathcal{T}(F)$ (if exists) such that a is τ-accepted under σ?

The rationale behind these problems is this. Minimizing the value of τ required to make S a τ-extension and a τ-accepted aims at discarding as few agents as possible from the dispute. This way, we try to preserve as much as possible what the agents said, and whenever we discard some agent, we do this consistently with the trust degrees (as users are discarded in ascending order of their trust degree). Given this, the output τ^* of MIN-TVER and MIN-TACC can help reason on the dispute in the following sense. If τ^* is "low", it means that only users with low trust degree must be discarded to certify the robustness of S and a: hence, even if S is not an extension and a is not accepted in the traditional sense, they can be reasonably considered as robust, since, in some sense, the culprit of them being not robust is what said by agents with low trust degree. Vice versa, if the returned τ^* is "high", it means that things said by trustworthy agents must be discarded in order to make S an extension and a accepted, thus it may be risky to consider them "robust".

Example 4 *From the discussion in Example 3 regarding the set $\{c, f\}$ and the argument a, it follows that* MIN-TVER$^{\text{ad}}(F, \{c, f\}) = 0$ *and* MIN-TACC$^{\text{ad}}(F, a) = 2$.

MIN-TVER and MIN-TACC are the natural optimization counterparts of the following decision problems over a given T-AAF F and under a semantics σ:

- TVER$^\sigma(F, S, \tau^*)$: Is S a τ-extension of F under σ for some $\tau \leq \tau^*$?
- TACC$^\sigma(F, a, \tau^*)$: Is a τ-accepted for F under σ for some $\tau \leq \tau^*$?

We will address these problems in the preliminary phase of the study of the complexity of MIN-TVER and MIN-TACC, since the complexity characterization of the optimization counterparts is simplified by the knowledge of the complexity of the decisional counterparts. However, TVER and TACC are of independent interest, and an example of their practical relevance is the scenario where an analyst wants to reason on the dispute by examining what happens when a fixed trust degree is used as threshold. This happens when the analyst has in mind some trust degree τ^* above which s/he has a

[3] The framework is orthogonal to the aggregation trust function. Switching to another aggregate operator (such as sum, avg) would merely change the trust degrees of the arguments in the T-AAF. Hence, the complexity results presented in the paper remain valid.

strong motivation for considering the agents trustworthy, while s/he has no strong motivation for considering the agents ranked below τ^* trustworthy or not. For instance, τ^* can be the trust degree of an agent that the analyst considers trustworthy for personal knowledge.

Remark: on the non-monotone behavior w.r.t. the trust degrees used as thresholds for cutting arguments. It is worth noting that, by definition, the result τ^* of an instance of MIN-TVER$^\sigma(F, S)$ has the property that S is an extension over F^{τ^*} but not over any F^τ with $\tau \in [0..\tau^*)$. However, this does not imply that any trust degree τ higher than τ^* still warrants that S is an extension over F^τ. The analogous behavior can be observed when reasoning on MIN-TACC$^\sigma(F, a)$. This derives from the fact that increasing the trust threshold not only discards attacks, but also defenses, thus it is reasonable that an argument may move from "accepted" to "not accepted" (and vice versa) after increasing the trust threshold. In this setting, the reason for specifically searching for the minimum trust threshold is that the goal of making S an extension or a accepted is pursued under the requirement of taking into account as much as possible what was claimed by the agents (coherently with their trust degrees). However, if the analyst is not satisfied with this kind of analysis, as s/he wants to get a more precise and complete picture of the trust intervals wherein S is or is not an extension (or a is or is not accepted), the problems studied in the paper are still of practical relevance. In fact, the characterization of these trust intervals can be obtained by iteratively invoking a solver for MIN-TVER$^\sigma(F, S)$ (or MIN-TACC$^\sigma(F, a)$) according to this scheme:
1) The first invocation of the solver is over the T-AAF $F^0 = F$;
2) Then, at each of the following iterations, the solver is invoked over the T-AAF F^τ, where τ is the next value in $T(F)$ not already considered and greater than the threshold τ_i returned by the previous invocation.
The sequence of the solutions of these invocations can be straightforwardly used to re-construct the intervals containing the trust degrees for which S is (or is not) a τ-extension.

5 COMPLEXITY CHARACTERIZATION

We start by addressing the decisional variants TVER$^\sigma(F, S, \tau^*)$ and TACC$^\sigma(F, a, \tau^*)$. Due to space limitations, the proof of Theorem 4 is sketched. The practical relevance of the results of our complexity analysis is discussed in Section 8.

Theorem 1 TVER$^\sigma(F, S, \tau^*)$ *is in P for $\sigma \in \{ad, co, st, gr\}$ and is coNP-complete for $\sigma = pr$.*

Proof. Let $\{\tau_1, \ldots, \tau_x\} = \{\tau \in \mathcal{T}(F) \mid \tau \leq \tau^*\}$. The case $\sigma \neq pr$ is trivial: TVER$^\sigma(F, S, \tau^*)$ can be decided by iteratively invoking an algorithm solving VER$^\sigma(F^{\tau_i}, S)$ (that is in P, as reported in Table 1), where τ_i ranges over $\{\tau_1, \ldots, \tau_x\}$.

The statement TVER$^{pr}(F, S, \tau^*) \in$ *coNP* follows from the fact that a polynomial size witness for the answer "*false*" consists of x supersets $S_1, \ldots, S_x$ of S witnessing that S is not maximally admissible in $F^{\tau_1}, \ldots, F^{\tau_x}$, respectively. The hardness for TVER$^{pr}(F, S, \tau^*)$ is implied by the reducibility from VER^{pr} (that is *coNP*-complete). □

With similar arguments exploiting the fact that the counterpart ACC$^\sigma(F, a)$ is NP-complete (under all the considered semantics), the following theorem regarding TACC$^\sigma(F, a, \tau^*)$ can be easily proved.

Theorem 2 TACC$^\sigma(F, a, \tau^*)$ *is NP-complete for every $\sigma \in \{ad, co, st, pr\}$ and is in FP for $\sigma = gr$.*

Proof. The membership to P for $\sigma = gr$ straightforwardly follows from the fact that ACC$^\sigma(F, a)$ is in P for $\sigma = gr$. For $\sigma \in \{ad, st\}$, the membership to NP derives from the correctness of the guess-and-check strategy consisting of the following steps:
i) guess a value $\tau' \leq \tau^*$;
ii) compute $F^{\tau'}$;
iii) guess a subset S of the arguments in $F^{\tau'}$ containing a;
iv) check whether S is an admissible extension for $F^{\tau'}$ (if $\sigma = ad$) or S is a stable extension (if $\sigma = st$).
Obviously, the computational steps can be done in polynomial time. In particular, for step iv, this derives from the fact that VER$^{ad}(F, S)$ and VER$^{st}(F, S)$ are in P. Finally, for $\sigma \in \{co, pr\}$, the NP-membership follows from the fact that the acceptability problem under these semantics is equivalent to the case $\sigma = ad$.

The hardness trivially follows from the reducibility from ACC$^\sigma(F, a)$. □

We now focus on MIN-TVER$^\sigma(F, S)$ and MIN-TACC$^\sigma(F, a)$. We first present a construction that encodes any 3-CNF ϕ into an AAF F_ϕ and that will be used in the proofs of the main results.

Definition 1 (F_ϕ) *Let $\phi = C_1 \wedge C_2 \wedge \ldots C_k$ be a 3-CNF formula over the set $X = \{x_1, \ldots, x_n\}$ of propositional variables. Let every clause C_i be of the form $C_i = l_i^1 \vee l_i^2 \vee l_i^3$, where each l_i^u is a literal of the form x_j or $\neg x_j$. We define F_ϕ as the AAF $\langle A, D\rangle$ where:*

- *A consists of: (i) two arguments x_j and $\neg x_j$ for each propositional variable $x_j \in X$; (ii) an argument c_i for each clause C_i in ϕ; (iii) the two arguments ϕ and ψ;*
- *D contains, for each clause C_i, (i) an attack $\delta_{\phi_i} = (c_i, \phi)$; (ii) an attack $\delta_{\psi_i} = (\psi, c_i)$; (iii) for each literal l_i^u in C_i, either the attack $\delta_i^u = (x_j, c_i)$ or $\delta_i^u = (\neg x_j, c_i)$, depending on whether $l_i^u = x_j$ or $l_i^u = \neg x_j$, respectively. Moreover, D contains, for each x_j in X, the four attacks $\delta_{1_j} = (\psi, x_j)$, $\delta_{2_j} = (\psi, \neg x_j)$, $\delta_{3_j} = (x_j, \neg x_j)$, $\delta_{4_j} = (\neg x_j, x_j)$. Finally, D contains the attack $\delta_\phi = (\phi, \psi)$.*

Example 5 *Consider the $3-$CNF formula $\phi = (x_1 \vee x_2 \vee x_3) \wedge (\neg x_1 \vee x_2 \vee \neg x_3) \wedge (\neg x_1 \vee \neg x_2) \wedge (\neg x_2 \vee x_3) \wedge (\neg x_2 \vee \neg x_3)$. The corresponding F_ϕ is depicted in Figure 2.*

Lemma 1 states the relationship between the satisfiability of ϕ and the non-emptiness of the preferred extensions of F_ϕ.

Lemma 1 *$\emptyset$ is a preferred extension in F_ϕ iff there is no truth assignment for $x_1, \ldots, x_n$ satisfying ϕ.*

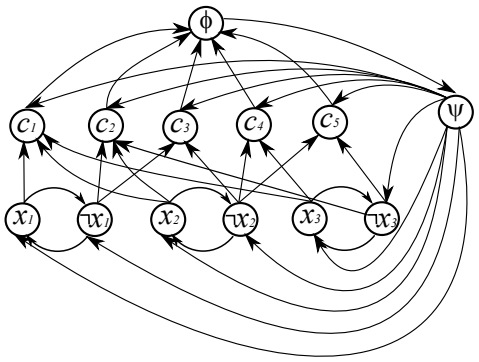

Figure 2: The AAF F_ϕ corresponding to $\phi = (x_1 \vee x_2 \vee x_3) \wedge (\neg x_1 \vee x_2 \vee \neg x_3) \wedge (\neg x_1 \vee \neg x_2) \wedge (\neg x_2 \vee x_3) \wedge (\neg x_2 \vee \neg x_3)$.

Proof. ($\Rightarrow$) We reason by contradiction. Assume that there is a truth assignment t making ϕ evaluate to true. We build the set L of n arguments, one for each variable in X, such that $x_m \in L$ iff $t(x_m)$= *true*, otherwise $\neg x_m \in L$, for each $x_m \in X$. The conflict-freeness of L follows from the fact that, for each variable, we put in L either x_m or $\neg x_m$. Since t makes ϕ evaluate to true, for each clause c_i we can find at least one variable $x_j \in X$ for which the truth value assigned to x_j by t makes c_i evaluate to true. This, in turn, means that L defends ϕ from every attack from $c_1, \ldots, c_k$. Since ϕ attacks ψ, the set L is defended from the attacks from ψ, thus the set $L \cup \{\phi\}$ is an admissible extension, and it is also a preferred extension, since no other argument is acceptable in it. Thus, $\emptyset$ is not a preferred extension.
($\Leftarrow$) We now show that, if there is no truth assignment t for $x_1, \ldots, x_n$ making ϕ evaluate to true, no admissible extension different from $\emptyset$ exists in F_ϕ. Since no satisfying truth assignment exists, there is no conflict-free set of arguments (composed by arguments of the form x_m or $\neg x_m$) defending ϕ from the attacks from $c_1, \ldots, c_k$. Then, since ϕ is in conflict also with ψ, is not acceptable in any extension. The arguments of the form x_m and $\neg x_m$ are not defended from the attacks from ψ, thus they are not acceptable. Every c_i is involved in attacks from/towards all the other arguments in $A \setminus \{c_1, \ldots, c_k\}$, thus it cannot be acceptable in any extension involving arguments in $A \setminus \{c_1, \ldots, c_k\}$. Furthermore, $\{c_1, \ldots, c_k\}$ is not an admissible extension as the attacks from ψ and from the arguments of the form x_m and $\neg x_m$ are not counterattacked. The argument ψ is in conflict with all the other arguments in A. Thus, in this case, $\emptyset$ is the unique admissible extension, implying that it is also a preferred extension. □

Example 6 *Continuing Example 5, it is easy to see that the set* $\{x_1, \neg x_2, \neg x_3, \phi\}$ *(resp.,* $\{\neg x_1, \neg x_2, x_3, \phi\}$*) is an admissible extension, and it corresponds to the truth assignment* $t = x_1/$true, $x_2/$false, $x_3/$false *(resp.,* $t = x_1/$false, $x_2/$false, $x_3/$true*) that makes* ϕ *evaluate to* true*. In this case, thus,* $\emptyset$ *is not a preferred extension. Consider now* $\phi' = \phi \wedge (x_2)$*. It is easy to see that* ϕ' *is not satisfiable (since in all the satisfying truth assignments for* ϕ *the variable* x_2 *was* false*). Hence,* $\emptyset$ *is a preferred extension for* $F_{\phi'}$*.*

Before stating the theorems characterizing MIN-TVER$^\sigma(F, S)$ and MIN-TACC$^\sigma(F, a)$ we characterize the complexity of the problem COUNT-ORD-INSTANCES, that is used in their proofs. Indeed, in their proofs, we exploit the $FP^{NP[\log n]}$-completeness of COUNT-ORD-INSTANCES: "Given a problem $\mathcal{P}$ complete for *NP* or *coNP* and n instances $I_1, \ldots, I_n$ of $\mathcal{P}$ that are ordered so that $\forall i \in [1..n-1]$ I_i *is false* $\Rightarrow$ I_{i+1} *is false*, compute how many instances evaluate to *true*". We start by considering its variant COUNT-INSTANCES, that is the following functional problem: "Given a problem $\mathcal{P}$ that is complete for NP or $coNP$ and n instances $I_1, \ldots, I_n$ of $\mathcal{P}$, compute the number n^T of instances that evaluate to *true*". The difference from COUNT-ORD-INSTANCES is that, in the latter, also an ordering criterion on the input instances $I_1, \ldots, I_n$ is imposed: $\forall i \in [1..n-1]$ I_i *is false* $\Rightarrow$ I_{i+1} *is false* (COUNT-ORD-INSTANCES). We characterize the two problems in the case where $\mathcal{P}$ is NP-complete (the case where $\mathcal{P}$ is coNP-complete follows straightforwardly). It is easy to see that COUNT-INSTANCES in $FP^{NP[\log n]}$: the value n^T can be obtained by performing a binary search over the interval $[1..n]$, where, at the generic iteration of the search, we consider an $x \in [1..n]$ and test whether at least x instances evaluate to *true*. This test can be done by invoking an *NP* oracle. In fact, since $\mathcal{P}$ is in NP, there is a polynomial size witness W^i for any *true* instance I_i of $\mathcal{P}$, thus there is also a polynomial size witness X for testing whether at least x instances are *true*: X consists of a set $\mathcal{I}$ of x distinct indices in $[1..n]$ and of the set W containing, for each $i \in \mathcal{I}$, the witness W^i. It is straightforward to see that this also shows the membership of COUNT-ORD-INSTANCES to $FP^{NP[\log n]}$. As for the hardness of COUNT-ORD-INSTANCES (which also implies the hardness of COUNT-INSTANCES), it can be easily proved by showing a reduction from MAX-CLIQUE, the $FP^{NP[\log n]}$-complete problem of computing the maximum size of a clique in a given graph G. In fact, MAX-CLIQUE can be reduced to the instance COI of COUNT-ORD-INSTANCES, where the underlying NP problem is CLIQUE (i.e., "*Given a graph G and an integer K, is there a clique of G with at least K nodes*?"), and the n instances $I_1, \ldots, I_n$ are defined as: $\forall i \in [1..n]$ I_i is the instance of CLIQUE over the same graph as MAX-CLIQUE where $K = i$. It is straightforward to see that this sequence of instances conforms to the ordering criterion of COUNT-ORD-INSTANCES (in fact, if G admits no clique with size i, it cannot contain any clique of size $i+1$) and that the output of COI is exactly what requested by MAX-CLIQUE.

We now introduce the theorems characterizing MIN-TVER$^\sigma(F, S)$ and MIN-TACC$^\sigma(F, a)$.

Theorem 3 MIN-TVER$^\sigma(F, S)$ *is in FP for* $\sigma \in \{ad, co, st, gr\}$ *and is* $FP^{NP[\log n]}$*-complete for* $\sigma = pr$

Proof. For $\sigma = gr$ we can reason as done in Theorem 1, by trying the trust degrees in $\mathcal{T}(F)$ in ascending order. The same strategy is correct also for $\sigma \in \{ad, co, st\}$, but we here provide, for these three semantics, more specific polynomial-time strategies, that work after checking that S is conflict-free (otherwise, MIN-TVER$^\sigma(F, S)$ has no solution, since there is no way of making S conflict-free by removing arguments outside S):

- $\sigma = ad$: a necessary and sufficient condition for making admissible a (conflict-free) set S by means of argument removals is removing from F the set A' of arguments attacking S without being counter-attacked by S. The answer of MIN-TVER$^\sigma(F, S)$ can be computed by evaluating A' and then computing the maximum trust degree τ' of the arguments in A'. If τ' is lower than the trust degrees of the arguments in S, then τ' is the answer of MIN-TVER$^\sigma(F, S)$. Otherwise, it means that removing A' implies the removal of arguments in S, thus MIN-TVER$^\sigma(F, S)$ has no solution.
- $\sigma = co$: we can reason analogously to the case $\sigma = ad$, but now the set of arguments to be removed is $A' \cup A''$, where A' is as above and A'' contains the arguments outside S that are acceptable w.r.t. S. Observe that A'' must be computed progressively, as removing acceptable arguments can make other arguments acceptable.
- $\sigma = st$: in this case, the arguments to be removed are all those outside S that are not attacked by S.

Under $\sigma = pr$, MIN-TVER$^{pr}(F, S)$ is in $FP^{NP[\log n]}$ since it can be solved by performing a binary search over $\mathcal{T}(F)$, where at each step a *coNP* oracle solving an instance of TVER^{pr} is called. As for the hardness, we show a reduction from COUNT-ORD-INSTANCES. Let I^{COI} be an instance of COUNT-ORD-INSTANCES over the instances $\phi_1, \ldots, \phi_h$ of UNSAT over 3-CNFs. We recall that the ordering imposed on the UNSAT instances means that $\forall i \in [1..h-1]$ if ϕ_i is *false* (i.e., the formula in ϕ_i is satisfiable) then ϕ_{i+1} is false too. For each ϕ_i we construct an AAF $F_{\phi_i} = \langle A_i, D_i \rangle$ as in Definition 1. We assume that a suitable renaming is performed to distinguish the arguments in each F_{ϕ_i} from the others. We define the T-AAF $F = \langle A, D, T \rangle$ where: $A = \{s\} \cup A_1 \cup \ldots \cup A_h$, $D = D_1 \cup \ldots \cup D_h$, and T assigns $h+1$ to s and $h+1-i$ to each argument belonging to A_i, for every $i \in [1..h]$.

Let $I^{\min}$ be the instance of MIN-TVER$^{pr}(F, S)$ with $S = \{s\}$. We show that $I^{\min}$ returns $h-i$ iff I^{COI} returns i.

($\Rightarrow$) Given that I^{COI} returns i, we take $\tau = h - i$ and consider F^{τ}. It is easy to see that F^{τ} is the restriction of F whose set of arguments consists of s along with the arguments of each F_{ϕ_j} with $j \in [1..i]$. Lemma 1 states that, for any j, if ϕ_j is unsatisfiable then $\emptyset$ is the unique preferred extension in F_{ϕ_j}. Hence, since $F_{\phi_1}, \ldots, F_{\phi_i}$ have no arguments/attacks in common, and since s in involved in no attack, it is easy to see that the fact that every ϕ_j is unsatisfiable (for $j \in [1..i]$) implies that $\{s\}$ is the unique preferred extension in F^{τ}. Furthermore, since I^{COI} returns i, for every $l \in [i+1..h]$ ϕ_l is satisfiable and thus F_{ϕ_l} admits a non-empty preferred extension (owing to Lemma 1). This means that $\{s\}$ is not a preferred extension in $F^{\tau'}$ for any $\tau' \leq h - i - 1$. Thus, $\tau = h - i$ is the minimum trust degree making $\{s\}$ a τ-extension under $\sigma = pr$.

($\Leftarrow$) Since $I^{\min}$ returns $\tau = h - i$, the set $\{s\}$ is a preferred extension for F^{τ}, and this, along with the fact that s is involved in no attack, implies that there is no other admissible extension in F^{τ}. This means that $\emptyset$ is the unique admissible extension in every F_{ϕ_j}, with $j \in [1..i]$. Then, Lemma 1 implies that the formulas $\phi_1, \ldots, \phi_i$ are unsatisfiable. Furthermore, since $I^{\min}$ returns $h - i$, $\{s\}$ is not a preferred extension in F^{h-i-1}. This means that $\emptyset$ is not the unique admissible extension in $F_{\phi_{i+1}}$. Hence, Lemma 1 implies that ϕ_{i+1} is satisfiable, thus I^{COI} returns i. □

Theorem 4 MIN-TACC$^{\sigma}(F, a)$ *is in FP for* $\sigma = gr$ *and* $FP^{NP[\log n]}$*-complete for* $\sigma \in \{ad, co, st, pr\}$.

Proof. The case $\sigma = gr$ can be solved with the same strategy as MIN-TVER$^{\sigma}(F, S)$ (see the proof of Theorem 3), but invoking a solver for ACC$^{gr}(F, a)$ (that is in P) instead of VER$^{gr}(F, S)$. As for the other semantics, the problem is in $FP^{NP[\log n]}$ since it can be solved via a binary search over $\mathcal{T}(F)$, where each step submits an instance of TACC$^{\sigma}(F, a, \tau^*)$ to an *NP* oracle. The hardness can be proved with a reduction from COUNT-ORD-INSTANCES, whose rationale is similar to the reduction in Theorem 3. □

6 VARIANTS OF THE FRAMEWORK

We discuss two variants of MINTVER and MINTACC that can support the reasoning over T-AAFs, especially when the arguments' trust degrees are not levels but additive measures. For instance, consider the case where the trust degrees result from a voting session, where each argument is associated with the number of people agreeing with it. Hence, when reasoning on alternative ways of making S an extension or a an accepted argument, removing a set of arguments X can be reasonably viewed as a weaker modification than removing a set of arguments Y if the sum of the weights in X is less than Y (since this corresponds to disregard fewer "votes" of the agents). This means that it makes sense to consider the problems MIN-SUMVER$^{\sigma}(F, S)$ and MIN-SUMACC$^{\sigma}(F, a)$, that search for the minimum sum of trust degrees (i.e., the minimum number of votes) that must be discarded to make S an extension and a accepted, respectively.

More formally, for a set of arguments X, we denote as $T(X) = \sum_{a \in X} T(a)$ the overall trust degree of the arguments in X, and as F^{-X} the T-AAF obtained from F by removing the arguments in X and the attacks involving them. Hence, MIN-SUMVER$^{\sigma}(F, S)$ and MIN-SUMACC$^{\sigma}(F, a)$ are:

- MIN-SUMVER$^{\sigma}(F, S)$: what is the minimum $T(X)$ such that S is an extension for F^{-X} under σ?
- MIN-SUMACC$^{\sigma}(F, a)$: what is the minimum $T(X)$ such that a is accepted in F^{-X} under σ?

Compared with MINTVER and MINTACC, this means that, if we discard an argument with trust degree τ, in MINSUMVER and MINSUMACC we are allowed to not discard arguments with trust degree lower than τ. Thus, the arguments are now considered independent when deciding on their removal, and the objective is to preserve as much as possible the overall trust of the arguments by minimizing the sum of the trust degrees of what is discarded.

It is worth noting that MIN-SUMACC$^{\sigma}(F, a)$ under $\sigma = gr$ is analogous to the problem MIN-BUDGET of [18]: the difference is that in MIN-BUDGET the weights are on the attacks and their semantics is different (a weight of an attack represents a measure of the inconsistency that we introduce if we discard the attack). When discussing the following two theorems, we will come back on this analogy, and explain both how the results in [18] help the characterization of our problems, and how our new results complete those in [18].

Theorem 5 MIN-SUMACC$^{\sigma}(F, a)$ *is* FP^{NP}*-complete for any* $\sigma \in \{ad, co, st, pr, gr\}$.

Theorem 5 is a straightforward consequence of the analogous result for the above-mentioned problem MIN-BUDGET (it can be proved, for all the semantics, with minor changes to the proof of FP^{NP}-completeness of MIN-BUDGET in [18]). As for MIN-SUMVER$^{\sigma}(F, S)$, its dual version over weighted attacks was mentioned (in the decisional version) but not characterized in [18]. Thus, the following theorem also completes the picture provided in [18], as it can be shown to hold also if the weights are associated with the attacks (as happens in the framework of [18]).

Theorem 6 MIN-SUMVER$^{\sigma}(F, S)$ *is:* 1) *in FP for* $\sigma \in \{ad, co, st\}$, 2) *in* $FP^{\Sigma_2^p}$ *and* $FP^{||NP}$*-hard for* $\sigma = pr$, 3) *in* FP^{NP} *for* $\sigma = gr$.

The fact that MIN-SUMVER$^{\sigma}(F, S)$ is in FP under $\sigma \in \{ad, co, st\}$ can be proved with the same reasoning used for proving that MIN-TVER$^{\sigma}(F, S)$ is in FP under the same semantics (Theorem 3). On the contrary, the strategy used in the proof of Theorem 3 for proving that MIN-TVER$^{\sigma}(F, S)$ is in FP under $\sigma = gr$ cannot be used to prove that MIN-SUMVER$^{gr}(F, S)$ is in FP. The reason is that the mechanism of MIN-SUMVER based on the sums of trust degrees introduces a form of complexity that is not present in MINTVER. In fact, once a value k for the sum of trust degrees is taken, several T-AAFs can be obtained from F by discarding a set of argument whose "overall" trust is k (while, in the threshold mechanism of MIN-TVER$^{\sigma}(F, S)$, a given threshold yields exactly one T-AAF). Given this, the precise characterization of MIN-SUMVER$^{gr}(F, S)$ remains an open problem. However, the following theorem states that MIN-SUMVER$^{gr}(F, S)$ is in FP at least when the argumentation graph is acyclic.

Proposition 1 *Over T-AAFs whose argumentation graph is acyclic,* MIN-SUMVER$^{gr}(F, S)$ *is in FP while* MIN-SUMACC$^{gr}(F, a)$ *is* FP^{NP}*-complete.*

Proof. The FP^{NP}-completeness of MIN-SUMACC$^{gr}(F, a)$ over acyclic T-AAFs can be proved with minor changes to the proof of FP^{NP}-completeness of MIN-BUDGET over acyclic AAFs in [18]. The fact that MIN-SUMVER$^{gr}(F, S) \in P$ can be shown by considering the classical algorithm $\mathcal{A}$ that computes the grounded extension S over an AAF by initializing S as the set of unattacked arguments

and then by iteratively incorporating in S the arguments defended by the version of S computed at the previous iteration, until a fixpoint is reached. It can be seen that, since F is acyclic, $\mathcal{A}$ can be made return S by removing from F, at each iteration of $\mathcal{A}$, all the arguments that are incorporated in the grounded extension but are not in S, along with the arguments not in S that recursively become unattacked after these removals. Denoting as X the set of arguments removed according to this strategy, it is easy to see the acyclicity of F makes removing X from F the minimal condition for S to be the grounded extension of F^{-X}. □

Proposition 1 highlights an asymmetry: if the argumentation graph is acyclic, MIN-SUMACC$^{gr}(F, a)$ remains FP^{NP}-complete, while MIN-SUMVER$^{gr}(F, S)$ is polynomial. In some sense, this means that checking if an argument becomes accepted after removing an amount of trust degrees Σ requires looking into all the restrictions F^{-X} with $T(X) = \Sigma$, independently from the cyclicity of the argumentation graph. On the contrary, if the argumentation graph is acyclic, computing the answer of MIN-SUMVER$^{gr}(F, S)$ can be done by looking only into the input extension and the argumentation graph.

7 RELATED WORK

There are a lot of works extending AAFs with the aim of representing the "strength" of arguments and/or attacks. These proposals differ in how this aspect is encoded, for instance via preferences [3], degrees of beliefs [32], importance of the values the arguments pertain to [8, 4] and probabilities [20, 21, 22, 26].

The reasonability of associating weights with arguments or attacks has been widely discussed in the literature, and, as observed in [18], depending on the scenarios and the semantics of the weights, there are cases where assigning weights to arguments is more reasonable than to attacks, and vice versa. An example of weighted AAF where weights represent trust degrees and are associated with the arguments is [15], where a fuzzy reasoning mechanism is embedded in *SMACk*, a system for analyzing arguments taken from disputes available in online commercial websites. The latter work, along with [5, 23, 25, 13, 31], belongs to the family of approaches where the reasoning yields acceptability degrees for the arguments, obtained by suitably revising the "initial" arguments' strengths. A second family of approaches [2, 8, 30, 26, 33, 24], instead, eventually produces a binary result for each argument, stating whether it is acceptable or not. In this regard, our framework can be viewed in between these two families: on the one hand, the mechanisms invoked to decide if S is an extension and a accepted produce a binary result; on the other hand, the results of MIN-TVER$^{\sigma}(F, S)$, MIN-TACC$^{\sigma}(F, a)$ and their variants could be also viewed as "strengths" of S and a. However, these strengths are not revisions of the initial weights. For instance, consider an argument a with the highest trust degree in $\mathcal{T}(A)$. If the answer of MIN-TACC$^{\sigma}(F, a)$ is 0, it means that even discarding no argument, a is accepted, that is a positive characteristics, and not a downgrading of $T(a)$. Thus, several properties listed in [1] regarding the output strength of arguments (such as *Weakening* and *Maximality*) make no sense on our semantics, as they are better tailored at reasoning paradigms belonging to the first family.

It is worth noting that our results still hold if the weights are associated to attacks: the difference in semantics does not correspond to a difference in computational complexity and solution strategies. Thus, in particular, as observed in Section 6, our work completes the framework in [18] (where the problem MIN-BUDGET, dual to MIN-TACC$^{gr}(F, a)$, was addressed). In fact, on the one hand, our results on MIN-SUMVER$^{\sigma}(F, S)$ can be used to solve the verification problem analogous to MIN-BUDGET (which deals only with the acceptance). On the other hand, our results on MIN-TVER$^{\sigma}(F, S)$ and MIN-TACC$^{\sigma}(F, a)$ can be used over the framework of [18] to use a different threshold-based mechanism tailored at the case where the weights denote levels instead of additive measures.

In this regard, the interest of the research community to extending the framework in [18] in the direction of our work is witnessed by [12], where the use of aggregate operators other than *sum* (including *min* and *max*) for reasoning on attacks to be discarded was formalized. However, no result on the computational complexity and no computational method has been proposed in [12] for these extensions.

Other variants of AAFs related to our T-AAFs are those based on a qualitative modeling of the uncertainty regarding the presence of arguments/attacks. These include incomplete AAFs [7], as well as the frameworks dealing with extension enforcement [6, 11, 34] and strategic argumentation [29]. When comparing our framework with these approaches, the trust degrees associated with the agents in our work can be abstractly viewed as values guiding the exploration of the alternative scenarios yielded by the uncertainty. For instance, our notion of τ-restriction corresponds to the notion of *completion* of incomplete AAFs (in the case where the incompleteness involves only the arguments), but while the reasoning over iAAFs considers all the possible completions (i.e., all the possible combinations of uncertain arguments), in our approach we consider only the scenarios filtered by the threshold mechanism. We also observe that, differently from our setting, the problems investigated in the literature of the various settings discussed so far are mainly decisional, with very few exceptions (for instance the *Optimal Extension Enforcement* problem).

8 CONCLUSIONS

We have studied some natural extensions of the verification and acceptance problems for reasoning over AAFs where the trustworthiness of the agents is encoded as a weight function over the arguments. The obtained results not only provide a framework for embedding the trust of agents in the traditional reasoning over AAFs, but also complete some results in the literature regarding similar problems over weighted AAFs.

It is worth noting that our results are relevant also from a practical standpoint. In fact, the proofs of the tractable cases contain the description of ad-hoc strategies for efficiently solving the addressed problems. As for the hard cases, the lower and upper bounds shown in the paper give a hint on suitable solving approaches. For instance, the problems that have been shown to be inside the class FP^{NP} can be solved by translating them into ILP instances and invoking a well-established ILP solver, as proposed in [18] for MIN-BUDGET (MIN-SUMVER$^{pr}(F, S)$ may not be solvable this way, as long as its membership to FP^{NP} is not proved). Generally speaking, resorting to ILP solvers (such as CPLEX) is a reasonable choice (if allowed by the expressiveness of ILP, that is bounded by FP^{NP}), as this exploits a number of heuristics implemented in the commercial solvers that in many cases enhance the efficiency of evaluating even hard instances. Future work will be devoted to implement ILP-based strategies and compare them with the usage of SAT-solvers, that are commonly used as tools for verifying/generating the extensions and deciding the acceptance of arguments in "classical" abstract argumentation.

Ongoing work is focused on characterizing the acceptance prob-

lem under the skeptical semantics, and embedding the implementation of the proposed computational models in a system supporting the analysis of the reviews published by the customers of commercial websites. In this regard, the combination of our framework with argumentation mining techniques [9, 27] will be investigated.

REFERENCES

[1] Leila Amgoud, Jonathan Ben-Naim, Dragan Doder, and Srdjan Vesic, 'Acceptability semantics for weighted argumentation frameworks', in *Proc. Int. Joint Conf. on Artificial Intelligence (IJCAI), Melbourne, Australia, Aug. 19-25, 2017*, pp. 56–62, (2017).

[2] Leila Amgoud and Claudette Cayrol, 'A reasoning model based on the production of acceptable arguments', *Ann. Math. Artif. Intell.*, **34**(1-3), 197–215, (2002).

[3] Leila Amgoud and Srdjan Vesic, 'A new approach for preference-based argumentation frameworks', *Ann. Math. Artif. Intell.*, **63**(2), 149–183, (2011).

[4] Katie Atkinson and Trevor Bench-Capon, 'Value based reasoning and the actions of others', in *Proc. European Conf. on Artificial Intelligence (ECAI), The Hague, The Netherlands*, pp. 680–688, (2016).

[5] Pietro Baroni, Marco Romano, Francesca Toni, Marco Aurisicchio, and Giorgio Bertanza, 'Automatic evaluation of design alternatives with quantitative argumentation', *Argument & Computation*, **6**(1), 24–49, (2015).

[6] Ringo Baumann and Gerhard Brewka, 'Expanding argumentation frameworks: Enforcing and monotonicity results', in *Proc. Computational Models of Argument (COMMA), Desenzano del Garda, Italy, Sept. 8-10, 2010*, pp. 75–86, (2010).

[7] Dorothea Baumeister, Daniel Neugebauer, Jörg Rothe, and Hilmar Schadrack, 'Verification in incomplete argumentation frameworks', *Artif. Intell.*, **264**, 1–26, (2018).

[8] Trevor J. M. Bench-Capon, 'Persuasion in practical argument using value-based argumentation frameworks', *J. Log. Comput.*, **13**(3), 429–448, (2003).

[9] Elena Cabrio and Serena Villata, 'Five years of argument mining: a data-driven analysis', in *Proc. Int. Joint Conference on Artificial Intelligence (IJCAI), July 13-19, 2018, Stockholm, Sweden*, pp. 5427–5433, (2018).

[10] Sylvie Coste-Marquis, Caroline Devred, and Pierre Marquis, 'Symmetric argumentation frameworks', in *Proc. of Symbolic and Quantitative Approaches to Reasoning with Uncertainty (ECSQARU), Barcelona, Spain*, pp. 317–328, (2005).

[11] Sylvie Coste-Marquis, Sébastien Konieczny, Jean-Guy Mailly, and Pierre Marquis, 'Extension enforcement in abstract argumentation as an optimization problem', in *Proc. Int. Joint Conference on Artificial Intelligence (IJCAI), Buenos Aires, Argentina, July 25-31, 2015*, pp. 2876–2882, (2015).

[12] Sylvie Coste-Marquis, Sébastien Konieczny, Pierre Marquis, and Mohand Akli Ouali, 'Weighted attacks in argumentation frameworks', in *Proc. Int. Conf. on Knowledge Representation and Reasoning (KR), Rome, Italy*, pp. 593–597, (2012).

[13] Célia da Costa Pereira, Andrea Tettamanzi, and Serena Villata, 'Changing one's mind: Erase or rewind?', in *Proc. Int. Joint Conf. on Artificial Intelligence (IJCAI), Barcelona, Catalonia, Spain, July 16-22*, pp. 164–171, (2011).

[14] Yannis Dimopoulos and Alberto Torres, 'Graph theoretical structures in logic programs and default theories', *Theor. Comput. Sci.*, **170**(1-2), 209–244, (1996).

[15] Mauro Dragoni, Célia da Costa Pereira, Andrea G. B. Tettamanzi, and Serena Villata, 'Combining argumentation and aspect-based opinion mining: The smack system', *AI Commun.*, **31**(1), 75–95, (2018).

[16] Phan Minh Dung, 'On the acceptability of arguments and its fundamental role in nonmonotonic reasoning, logic programming and n-person games', *Artif. Intell.*, **77**(2), 321–358, (1995).

[17] Paul E. Dunne and Trevor J. M. Bench-Capon, 'Coherence in finite argument systems', *Artif. Intell.*, **141**(1/2), 187–203, (2002).

[18] Paul E. Dunne, Anthony Hunter, Peter McBurney, Simon Parsons, and Michael Wooldridge, 'Weighted argument systems: Basic definitions, algorithms, and complexity results', *Artif. Intell.*, **175**(2), (2011).

[19] Paul E. Dunne and Michael Wooldridge, 'Complexity of abstract argumentation', in *Argumentation in Artificial Intelligence*, 85–104, (2009).

[20] Bettina Fazzinga, Sergio Flesca, and Filippo Furfaro, 'Probabilistic bipolar abstract argumentation frameworks: complexity results', in *Proc. Int. Joint Conf. on Artificial Intelligence(IJCAI), July 13-19, 2018, Stockholm, Sweden*, pp. 1803–1809.

[21] Bettina Fazzinga, Sergio Flesca, and Filippo Furfaro, 'Complexity of fundamental problems in probabilistic abstract argumentation: Beyond independence', *Artif. Intell.*, **268**, 1–29, (2019).

[22] Bettina Fazzinga, Sergio Flesca, and Francesco Parisi, 'On the complexity of probabilistic abstract argumentation frameworks', *ACM Trans. Comput. Log. (TOCL)*, **16**(3), 22, (2015).

[23] Dov M. Gabbay and Odinaldo Rodrigues, 'Equilibrium states in numerical argumentation networks', *Logica Universalis*, **9**(4), 411–473, (2015).

[24] Anthony Hunter, 'Probabilistic qualification of attack in abstract argumentation', *Int. J. Approx. Reasoning*, **55**(2), 607–638, (2014).

[25] João Leite and João Martins, 'Social abstract argumentation', in *Proc. Int. Joint Conf. on Artificial Intelligence, Barcelona, Catalonia, Spain, July 16-22, 2011*, pp. 2287–2292, (2011).

[26] Hengfei Li, Nir Oren, and Timothy J. Norman, 'Probabilistic argumentation frameworks', in *Proc. Int. Workshop on Theory and Applications of Formal Argumentation (TAFA), Barcelona, Spain*, pp. 1–16, (2011).

[27] Marco Lippi and Paolo Torroni, 'Argumentation mining: State of the art and emerging trends', *ACM Trans. Internet Technol.*, **16**(2), 10:1–10:25, (2016).

[28] Mengwen Liu, Yi Fang, Alexander G. Choulos, Dae Hoon Park, and Xiaohua Hu, 'Product review summarization through question retrieval and diversification', *Inf. Retr. Journal*, **20**(6), 575–605, (2017).

[29] Michael J. Maher, 'Resistance to corruption of strategic argumentation', in *Proc. Conf. on Artificial Intelligence (AAAI) Feb. 12-17, 2016, Phoenix, Arizona, USA*, pp. 1030–1036, (2016).

[30] Sanjay Modgil, 'Reasoning about preferences in argumentation frameworks', *Artif. Intell.*, **173**(9-10), 901–934, (2009).

[31] Antonio Rago, Francesca Toni, Marco Aurisicchio, and Pietro Baroni, 'Discontinuity-free decision support with quantitative argumentation debates', in *Proc. Int. Conf. on Principles of Knowledge Representation and Reasoning (KR), Cape Town, South Africa, April 25-29*, pp. 63–73, (2016).

[32] Francesco Santini, Audun Jøsang, and Maria Silvia Pini, 'Are my arguments trustworthy? abstract argumentation with subjective logic', in *Proc. Int. Conf. on Information Fusion (FUSION), Cambridge, UK, July 10-13*, pp. 1982–1989, (2018).

[33] Matthias Thimm, 'A probabilistic semantics for abstract argumentation', in *Proc. European Conf. on Artificial Intelligence (ECAI), Montpellier, France*, pp. 750–755, (2012).

[34] Johannes Peter Wallner, Andreas Niskanen, and Matti Järvisalo, 'Complexity results and algorithms for extension enforcement in abstract argumentation', *J. Artif. Intell. Res.*, **60**, 1–40, (2017).

[35] Rong Zhang, Wenzhe Yu, Chaofeng Sha, Xiaofeng He, and Aoying Zhou, 'Product-oriented review summarization and scoring', *Frontiers Comput. Sci.*, **9**(2), 210–223, (2015).