



**HAL**  
open science

## On the Aggregation of Argumentation Frameworks

Jérôme Delobelle, Sébastien Konieczny, Srdjan Vesic

► **To cite this version:**

Jérôme Delobelle, Sébastien Konieczny, Srdjan Vesic. On the Aggregation of Argumentation Frameworks. 24th International Joint Conference on Artificial Intelligence (IJCAI'15), 2015, Buenos aires, Argentina. hal-01936948

**HAL Id: hal-01936948**

**<https://hal.science/hal-01936948>**

Submitted on 27 Nov 2018

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# On the Aggregation of Argumentation Frameworks

Jérôme Delobelle Sébastien Konieczny Srdjan Vesic

CRIL, CNRS – Université d’Artois

Lens, France

{delobelle, konieczny, vesic}@cril.fr

## Abstract

We study the problem of aggregation of Dung’s abstract argumentation frameworks. Some operators for this aggregation have been proposed, as well as some rationality properties for this process. In this work we study the existing operators and new ones that we propose in light of the proposed properties, highlighting the fact that existing operators do not satisfy a lot of these properties. The conclusions are that on one hand none of the existing operators seem fully satisfactory, but on the other hand some of the properties proposed so far seem also too demanding.

## 1 Introduction

Argumentation is based on the exchange and the evaluation of interacting arguments. In Dung’s abstract argumentation theory [Dung, 1995], argumentation frameworks are modeled by binary graphs, where the nodes represent the arguments, and the edges represent the attacks between them. Then the question is to know what are the set of arguments that can be conjointly accepted from such graphs.

Dynamics of argumentation frameworks is an interesting question. In the last years there have been a lot of contributions of the problem of evolution (revision) of argumentation frameworks [Cayrol *et al.*, 2008; Boella *et al.*, 2009; Amgoud and Vesic, 2009; da Costa Pereira *et al.*, 2011; Baumann, 2012; Coste-Marquis *et al.*, 2014a; 2014b]. But there are quite a few contributions on the aggregation of argumentation frameworks. This problem is an important one for multi-agent systems: suppose that each member of a group of agents has his own argumentation framework, that represents his beliefs, then the problem is to define a (social) argumentation framework that represents the beliefs of the group. This is an important question on its own, but also because this computed aggregated argumentation framework could be considered as the result of an *ideal* negotiation process, with respect to which the results of *practical* negotiation protocols (see e.g. the work by Bonzon and Maudet [2011]) could be evaluated.

Contributions on this issue have been mainly proposals of particular aggregation methods [Coste-Marquis *et al.*, 2007; Tohmé *et al.*, 2008; Cayrol and Lagasque-Schiex, 2011;

Gabbay and Rodrigues, 2012]. We focus on the methods that comply with classical Dung’s setting, that means that they take as input classical abstract argumentation frameworks, and return as output either an abstract argumentation framework (or more generally a set of abstract argumentation frameworks), or a set of extensions. So we do not study here the proposal by Gabbay and Rodrigues [2012] since the result is an ordered set of arguments.

One interesting contribution is the work by Dunne *et al.* [2012], where the authors define “rationality” properties for this kind of aggregation methods, but do not check if existing methods satisfy them.

This is what we propose in this work. We study existing methods of the literature in the light of the proposed properties. We also propose an additional method based on WAFs (Weighted Argumentation Frameworks) [Dunne *et al.*, 2011; Coste-Marquis *et al.*, 2012a]. In these works, one of the possible interpretations of the weights on the attacks that is mentioned is that they may represent the number of agents in a group that agree with this attack. So we endorse this view, and check how to use these works to define aggregation methods.

The conclusions of this study are that, on one hand, none of the existing operators seem fully satisfactory, but, on the other hand, some of the properties proposed so far seem also too demanding. So this suggests that some further work is still required both on the definition of operators as well as on the study of their properties.

The remainder of this paper is as follows. In the next section we provide the necessary background on argumentation frameworks. Section 3 is a reminder about the weighted argumentation frameworks, and Section 4 recalls the proposals for aggregating argumentation frameworks. Section 5 studies the properties of these aggregation methods. In Section 6 we propose a new aggregation method (more precisely three variations), and study their properties. In Section 7 we sum up and discuss the obtained results, before concluding in Section 8.

## 2 Preliminaries

In this section, we briefly recall some key elements of abstract argumentation frameworks as proposed by Dung [1995].

**Definition 1** An argumentation framework (AF) is a pair

$F = \langle A, R \rangle$  with  $A$  a set of arguments and  $R$  a binary relation on  $A$ , i.e.  $R \subseteq A \times A$ , called the **attack relation**. A set of arguments  $S \subseteq A$  attacks an argument  $b \in A$ , if there exists  $a \in S$ , such that  $(a, b) \in R$ . For an AF  $F = \langle A, R \rangle$ , we note  $Arg(F) = A$  and  $Att(F) = R$ .

In this paper we will focus on the standard semantics defined by Dung [1995]: We note  $\mathcal{E}_\sigma(F)$  the set of extensions of  $F$  for the semantics  $\sigma \in \{\mathbf{comp}(lete), \mathbf{pref}(erred), \mathbf{sta}(ble), \mathbf{gr}(ounded)\}$ .

An argument  $a$  is **skeptically** accepted iff there is at least one extension and  $a$  belongs to all extensions. An argument is **credulously** accepted iff it belongs to at least one extension. We denote by  $sa_\sigma(F)$  (respectively  $ca_\sigma(F)$ ) the set of skeptically (resp. credulously) accepted arguments in  $F$ .

### 3 Weighted Argumentation Frameworks

Let us now turn to an extension of Dung’s framework that allows to put weights on the attacks [Dunne *et al.*, 2011; Coste-Marquis *et al.*, 2012a; 2012b].

**Definition 2** A Weighted Argumentation Framework is a triple  $WF = \langle A, R, w \rangle$  where  $\langle A, R \rangle$  is a Dung abstract argumentation framework, and  $w : A \times A \rightarrow \mathbb{N}$  is a function assigning a natural number to each attack (i.e.  $w(a, b) > 0$  iff  $(a, b) \in R$ ), and a null value otherwise ( $w(a, b) = 0$  iff  $(a, b) \notin R$ ).

Let us note  $\overline{WF}$  the standard argumentation framework obtained from a weighted argumentation framework  $WF$  by “forgetting” the weights, i.e. if  $WF = \langle A, R, w \rangle$  then  $\overline{WF} = \langle A, R \rangle$ . We later show how to use WAFs to aggregate several AFs. For the moment let us recall how one can use them for relaxing extensions and for selecting the best amongst several extensions.

#### 3.1 Relaxing Extensions

Initially WAFs were introduced with the idea to make sure to obtain non-empty extensions [Dunne *et al.*, 2011]. Roughly, the idea is to delete some attacks in order to obtain extensions. The following definitions come from the work by Coste-Marquis *et al.* [2012b]:

**Definition 3** We say that  $\oplus$  is an aggregation function if for every  $n \in \mathbb{N}$ ,  $\oplus$  is a mapping from  $\mathbb{N}^n$  to  $\mathbb{N}$  such that:

- if  $x_i \geq x'_i$ , then  $\oplus(x_1, \dots, x_i, \dots, x_n) \geq \oplus(x_1, \dots, x'_i, \dots, x_n)$
- $\oplus(x_1, \dots, x_n) = 0$  iff for every  $i$ ,  $x_i = 0$
- $\oplus(x) = x$ .

**Definition 4** Let  $WF = \langle A, R, w \rangle$  be a weighted argumentation framework,  $\sigma$  be a semantics, and  $\otimes$  be an aggregation function. The aggregation of the weights of the attacks in a set  $S \subseteq R$  is  $w_\otimes(S, w) = \otimes_{(a,b) \in S} w(a, b)$ .

The function  $Sub(R, w, \beta)$  returns the set of subsets of  $R$  whose total aggregated weight does not exceed  $\beta$ :  $Sub(R, w, \beta) = \{S \mid S \subseteq R \text{ and } w_\otimes(S, w) \leq \beta\}$ . The set of  $\sigma_\otimes^\beta$ -extensions of  $WF$ , denoted  $\mathcal{E}_\sigma^{\otimes, \beta}(WF)$ , is defined as:  $\mathcal{E}_\sigma^{\otimes, \beta}(WF) = \{E \in \mathcal{E}_\sigma(\langle A, R \setminus S \rangle) \mid S \in Sub(R, w, \beta)\}$ .

In contrast to what happens in Dung’s setting, several grounded extensions may exist when relaxing extensions are considered. In particular it may happen that the empty set belongs to a set of relaxed extensions, which can be problematic because it trivializes skeptical inference relation, so the next definition removes this empty set from the set of extensions.

The most interesting value of  $\beta$  is the smallest one that leads to a non-empty extension (for the semantics under consideration):

**Definition 5** Given a weighted argumentation framework  $WF = \langle A, R, w \rangle$ , a semantics  $\sigma$ , and an aggregation function  $\otimes$ , the set of  $\sigma_\otimes$ -extensions of  $WF$ , denoted by  $\mathcal{E}_\sigma^\otimes(\langle A, R, w \rangle)$  is defined as  $\mathcal{E}_\sigma^\otimes(\langle A, R, w \rangle) = \underline{\mathcal{E}_\sigma^{\otimes, \beta}(\langle A, R, w \rangle)}$  where<sup>1</sup>:

- $\mathcal{E}_\sigma^{\otimes, \beta}(\langle A, R, w \rangle)$  is non-trivial<sup>2</sup>,
- There is no  $\beta' < \beta$  s.t.  $\mathcal{E}_\sigma^{\otimes, \beta'}(\langle A, R, w \rangle)$  is non-trivial,
- For a set  $\mathcal{E}$  of extensions,  $\underline{\mathcal{E}} = \mathcal{E} \setminus \{\emptyset\}$

#### 3.2 Best Extensions

In general, an argumentation framework may admit a large number of extensions for some semantics. Within the WAF setting, it is possible to take advantage of the available weights, in order to select the “best” extensions. In the paper by Coste-Marquis *et al.* [2012a] this selection process goes through a comparison of the extensions scores, expressing intuitively how good they are.

**Definition 6** Let  $WF = \langle A, R, w \rangle$  be a weighted argumentation framework. Let  $E$  and  $F$  be two extensions of  $\overline{WF}$  for a given semantics  $\sigma$  and  $\oplus$  be an aggregation function.

The  $\oplus$ -attack from  $E$  on  $F$  is:  $S_\oplus(E \rightarrow F) = \oplus_{a \in E, b \in F} w(a, b)$ . Then  $E >_\oplus F$  iff  $S_\oplus(E \rightarrow F) > S_\oplus(F \rightarrow E)$ .

Let us introduce an additional parameter to the best function, that will be useful for the operators we define in Section 7.

**Definition 7** Let  $WF$  be a weighted argumentation framework and  $X$  be a set such that  $\forall x \in X, x \subseteq Arg(WF)$ . Let  $\oplus$  be an aggregation function. Then :

- $best_1^\oplus(X, WF) = \{E \in X \mid \nexists E' \in X, E' >_\oplus E\}$
- $best_2^\oplus(X, WF) = argmax_{E \in X} |\{E' \in X \mid E >_\oplus E'\}|$
- $best_3^\oplus(X, WF) = argmax_{E \in X} |\{E' \in X \mid E >_\oplus E'\}| - |\{E' \in X \mid E' >_\oplus E\}|$
- $best_4^\oplus(X, WF) = argmax_{E \in X} KS_\oplus(E)$ , where  $KS_\oplus(E) = \min_{E' \in X, E' \neq E} (S_\oplus(E \rightarrow E'))$

Four natural ways to define the ordering  $>_\oplus$  are proposed in [Coste-Marquis *et al.*, 2012a] :

**Definition 8** Let  $WF$  be a weighted argumentation framework. Let  $\oplus$  be an aggregation function. Then,  $\forall i \in \{1, 2, 3, 4\}$ ,  $best_i^{\sigma, \oplus}(WF) = best_i^\oplus(\mathcal{E}_\sigma(\overline{WF}), WF)$ .

<sup>1</sup>We add a third condition compared to the original definition [Coste-Marquis *et al.*, 2012b], in order to directly encode the non-trivial skeptical inference [Coste-Marquis *et al.*, 2012b].

<sup>2</sup>A set of extensions is non-trivial if it has at least one non-empty extension.

## 4 Aggregation Operators

Some merging operators for argumentation frameworks, inspired by propositional logic merging operators [Koniczny and Pino Pérez, 2002], have been defined [Coste-Marquis *et al.*, 2007]. From now on, we make the hypothesis that all the argumentation frameworks are defined on the same set of arguments  $\mathcal{X}$ <sup>3</sup>. We use the notation  $\mathbb{AF}$  for the set of all argumentation frameworks defined from the set of arguments  $\mathcal{X}$  used by the agents.

**Definition 9** Let  $p = \langle AF_1, \dots, AF_n \rangle$  be a profile,  $d$  be a distance between AFs, and  $\oplus$  be an aggregation function. The merging operator  $\Delta_d^\oplus$  is defined as :  $\Delta_d^\oplus(\langle AF_1, \dots, AF_n \rangle) = \{AF \in \mathbb{AF} \mid AF \text{ minimizes } \oplus_{i=1}^n d(AF, AF_i)\}$ .

The distance used by Coste-Marquis *et al.* [2007] as an example is an edit distance (de), which is in our case equivalent to the cardinality of the symmetrical difference between the two attack relations. Typical examples of aggregation functions are *sum* ( $\Sigma$ ) and *leximax*. In what follows, we focus on these two functions.

Inspired by voting methods, other aggregation operators have been defined by Tohmé *et al.* [2008]. In particular, they propose a *qualified voting* method :

**Definition 10** Let  $p = \langle AF_1, \dots, AF_n \rangle$  be a profile and  $U \subset \{AF_1, \dots, AF_n\}$ . *Qualified voting* is defined as  $QV(\langle AF_1, \dots, AF_n \rangle) = (A, R)$ , where  $A$  is the set of arguments used by agents, and  $R = \{(a, b) \mid a, b \in A \text{ and } |\{AF_i : (a, b) \in \text{Att}(AF_i)\}| > \max(|\{AF_i : (b, a) \in \text{Att}(AF_i)\}|, |\{AF_i : (a, b) \notin \text{Att}(AF_i) \text{ and } (b, a) \notin \text{Att}(AF_i)\}|)\}$  and  $U \subseteq \{AF_i \mid (a, b) \in \text{Att}(AF_i)\}$ .

## 5 Properties of Aggregation Function

Dunne *et al.* [2012] propose some rationality properties for characterizing aggregation of sets of argumentation frameworks, based on translations of properties coming from social choice theory to the argumentation setting.

Recall that we note by  $\mathbb{AF}$  the set of all argumentation frameworks defined from a (finite) set of arguments and that we suppose that all agents have the same set of arguments. We denote by  $\mathcal{N}$  the set of all agents. An aggregation function  $\gamma$  is defined by  $\gamma : \mathbb{AF}^n \rightarrow \mathbb{AF}$ . We note  $\hat{F} = (F_1, \dots, F_n)$  a tuple in  $\mathbb{AF}^n$ . Unless stated explicitly all the properties are defined  $\forall \hat{F} \in \mathbb{AF}^n$ .

**Anonymity.** The aggregation function  $\gamma$  is **anonymous** if it produces the same argumentation framework for all permutations  $\Pi(\hat{F})$  of the input.

(ANON)  $\forall \hat{F}' \in \Pi(\hat{F}) : \gamma(\hat{F}') = \gamma(\hat{F})$

**Non-Triviality.** An argumentation framework is non-trivial, for a semantics  $\sigma$ , if it has at least one non-empty extension:  $|\mathcal{E}_\sigma(F)| \geq 1$  and  $\mathcal{E}_\sigma(F) \neq \{\emptyset\}$ . Let us note  $\mathbb{AF}_{NT_\sigma}$  the set

of non-trivial (for the semantics  $\sigma$ ) argumentation frameworks. The aggregation function  $\gamma$  is  **$\sigma$ -strongly non-trivial** if the output is always non-trivial :

( $\sigma$ -SNT)  $\gamma(\hat{F}) \in \mathbb{AF}_{NT_\sigma}$

The aggregation function  $\gamma$  is  **$\sigma$ -weakly non-trivial** if, when all the input frameworks are non-trivial, then the output framework is non-trivial:

( $\sigma$ -WNT)  $\forall \hat{F} \in \mathbb{AF}_{NT_\sigma}^n : \gamma(\hat{F}) \in \mathbb{AF}_{NT_\sigma}$

**Decisiveness.** An argumentation framework is decisive, for a semantics  $\sigma$ , if it has exactly one non-empty extension:  $|\mathcal{E}_\sigma(F)| = 1$  and  $\mathcal{E}_\sigma(F) \neq \{\emptyset\}$ . Let us note  $\mathbb{AF}_{D_\sigma}$  the set of decisive (for the semantics  $\sigma$ ) argumentation frameworks. The aggregation function  $\gamma$  is  **$\sigma$ -strongly decisive** if the output is always decisive :

( $\sigma$ -SD)  $\gamma(\hat{F}) \in \mathbb{AF}_{D_\sigma}$

The aggregation function  $\gamma$  is  **$\sigma$ -weakly decisive** if when all the input frameworks are decisive, then the output framework is decisive:

( $\sigma$ -WD)  $\forall \hat{F} \in \mathbb{AF}_{D_\sigma}^n : \gamma(\hat{F}) \in \mathbb{AF}_{D_\sigma}$

**Unanimity.** This property specifies that if all agents are unanimous with respect to some aspect of the domain (extensions, attacks, ...), for a semantics  $\sigma$ , then this unanimity should be reflected in the social outcome.

• **Unanimous attack** checks attacks between arguments:

(UA)  $\bigcap_{k=1}^n \text{Att}(F_k) \subseteq \text{Att}(\gamma(\hat{F}))$

•  **$\sigma$ -unanimity** concerns extensions:

( $\sigma$ -U)  $\bigcap_{k=1}^n \mathcal{E}_\sigma(F_k) \subseteq \mathcal{E}_\sigma(\gamma(\hat{F}))$

•  **$ca_\sigma$ -unanimity** concerns credulous inference:

( $ca_\sigma$ -U)  $\bigcap_{k=1}^n ca_\sigma(F_k) \subseteq ca_\sigma(\gamma(\hat{F}))$

•  **$sa_\sigma$ -unanimity** concerns skeptical inference:

( $sa_\sigma$ -U)  $\bigcap_{k=1}^n sa_\sigma(F_k) \subseteq sa_\sigma(\gamma(\hat{F}))$

**Majority.** If a strict majority of agents agree on something, then this should be reflected in the social outcome:

• **Majority attack** concerns attacks between arguments:

(MAJ-A)  $(|\{F_i : a \in \text{Att}(F_i)\}| > \frac{n}{2}) \Rightarrow a \in \text{Att}(\gamma(\hat{F}))$

•  **$\sigma$ -majority** concerns extensions:

( $\sigma$ -MAJ)  $(|\{F_i : S \in \mathcal{E}_\sigma(F_i)\}| > \frac{n}{2}) \Rightarrow S \in \mathcal{E}_\sigma(\gamma(\hat{F}))$

•  **$ca_\sigma$ -majority** concerns credulous inference:

( $ca_\sigma$ -MAJ)  $(|\{F_i : x \in ca_\sigma(F_i)\}| > \frac{n}{2}) \Rightarrow x \in ca_\sigma(\gamma(\hat{F}))$

•  **$sa_\sigma$ -majority** concerns skeptical inference:

( $sa_\sigma$ -MAJ)  $(|\{F_i : x \in sa_\sigma(F_i)\}| > \frac{n}{2}) \Rightarrow x \in sa_\sigma(\gamma(\hat{F}))$

**Closure.** These properties say that the aggregation function must not invent some entity which does not exist in the input.

• **Closure** says that the AF in output must match exactly one AF in input :

(CLO)  $\exists i \in \mathcal{N} : \text{Att}(\gamma(\hat{F})) = \text{Att}(F_i)$

• **Attack closure** says that if one attack is in the AF in output, this attack must be present in at least one AF in input:

(AC)  $\text{Att}(\gamma(\hat{F})) \subseteq \text{Att}(F_1) \cup \dots \cup \text{Att}(F_n)$

•  **$\sigma$ -closure** is related to extensions:

<sup>3</sup>This allows to simplify the definitions for input AFs. Namely, we are not obliged to use the notion of partial argumentation framework [Coste-Marquis *et al.*, 2007] that allows to represent the *ignorance* of an agent towards the fact that an attack holds or not, and that requires a first completion step before the proper merging.

$$(\sigma\text{-C}) \quad \forall S \in \mathcal{E}_\sigma(\gamma(\hat{F})) : S \in \bigcup_{k=1}^n \mathcal{E}_\sigma(F_k)$$

•  $ca_\sigma$ -closure is related to credulous inference :

$$(ca_\sigma\text{-C}) \quad \forall x \in ca_\sigma(\gamma(\hat{F})) : x \in \bigcup_{k=1}^n ca_\sigma(F_k)$$

•  $sa_\sigma$ -closure is related to skeptical inference :

$$(sa_\sigma\text{-C}) \quad \forall x \in sa_\sigma(\gamma(\hat{F})) : x \in \bigcup_{k=1}^n sa_\sigma(F_k)$$

Tohmé et al. [2008] propose some properties for characterizing good aggregation operators, inspired from social choice theory. They propose a property that they call *Pareto condition*, that is exactly Unanimous attack (UA), and a non-dictatorship property, that is satisfied by all reasonable aggregation operators. We give below the two other properties that they propose, that are translations of meaningful social choice theory properties:

**Positive responsiveness.** This property says that increasing the number of agents that have an attack, should not decrease the chance for that attack to appear in the social outcome:

**(PR)** Let  $\hat{F}$  and  $\hat{F}'$  be two profiles of  $\mathbb{AF}^n$ . If  $\{F_i \in \hat{F} \mid (a, b) \in Att(F_i)\} \subseteq \{F'_i \in \hat{F}' \mid (a, b) \in Att(F'_i)\}$ , and  $(a, b) \in Att(\gamma(\hat{F}'))$ , then  $(a, b) \in Att(\gamma(\hat{F}))$

**Independence of irrelevant alternatives.** Deciding whether an attack holds or not should be concerned only with the attacks between these two arguments in the input profile.

**(IIA)** Let  $\hat{F}$  and  $\hat{F}'$  be two profiles of  $\mathbb{AF}^n$ . If  $(\forall i (a, b) \in Att(F_i) \text{ iff } (a, b) \in Att(F'_i))$ , then  $((a, b) \in Att(\gamma(\hat{F}))) \text{ iff } (a, b) \in Att(\gamma(\hat{F}'))$

This is not mentioned in [Tohmé et al., 2008], but one can easily show that PR implies IIA:

**Proposition 1** *Positive responsiveness implies Independence of Irrelevance Alternatives.*

To this set of properties from the literature we want to add a very simple property that is missing. Indeed, if all the AFs at the input coincide, the result of merging should be identical to this AF.

• **Identity attack** on the attacks :

$$(\mathbf{A-ID}) \quad Att(\gamma(F, \dots, F)) = Att(F)$$

•  $\sigma$ -Identity on the extensions :

$$(\sigma\text{-ID}) \quad \forall F \in \mathbb{AF}_{NT_\sigma} : \mathcal{E}_\sigma(\gamma(F, \dots, F)) = \mathcal{E}_\sigma(F)$$

•  $ca_\sigma$ -Identity on the credulous inference :

$$(ca_\sigma\text{-ID}) \quad \forall F \in \mathbb{AF}_{NT_\sigma} : ca_\sigma(\gamma(F, \dots, F)) = ca_\sigma(F)$$

•  $sa_\sigma$ -Identity on the skeptical inference :

$$(sa_\sigma\text{-ID}) \quad \forall F \in \mathbb{AF}_{NT_\sigma} : sa_\sigma(\gamma(F, \dots, F)) = sa_\sigma(F)$$

These four intuitive properties are particular cases of the property of unanimity. In the case where we merge a single AF (which represents the beliefs of one agent) then this property implies that the operator should not change this AF, that, as we will see, is not ensured by all aggregation methods.

## 6 Properties of Existing Operators

Let us first check what are the properties satisfied by the merging operators of Coste-Marquis et al. [2007]. Recall that the properties introduced by Dunne et al. are defined for

a unique AF as output, whereas the merging operators may have several AFs as output. We generalize the properties from the previous section as follows: instead of asking that a property holds for *the* AF at the output, we ask that the same property is satisfied by *all* the output AFs. By following this idea, most of the properties can be generalized in a straightforward way. There exists no straightforward way to generalize the definition of IIA to the case when several AFs are allowed in the output. That is why we do not consider it in the remainder of the paper.

**Proposition 2**  $\Delta_{de}^\Sigma$  satisfies Anonymity (ANON), the properties of Identity (A-ID,  $\sigma$ -ID,  $ca_\sigma$ -ID,  $sa_\sigma$ -ID), the properties of Unanimity (UA), Majority (MAJ-A), Attack closure (AC) and Positive responsiveness (PR) for every semantics  $\sigma \in \{comp, pref, sta, gr\}$ . The other properties are not satisfied.

Let us check now if there are more properties satisfied when we use the leximax as aggregation function.

**Proposition 3**  $\Delta_{de}^{leximax}$  satisfies Anonymity (ANON), the properties of Identity (A-ID,  $\sigma$ -ID,  $ca_\sigma$ -ID,  $sa_\sigma$ -ID), the properties of Unanimity (UA), Attack closure (AC) and Positive responsiveness (PR) for every semantics  $\sigma \in \{comp, pref, sta, gr\}$ . The other properties are not satisfied.

Let us now turn to the properties satisfied by the qualified voting method [Tohmé et al., 2008].

**Proposition 4** *QV* satisfies *gr-weak non triviality* (*gr-WNT*), *gr-weak decisiveness* (*gr-WD*), Attack closure (AC) and Positive responsiveness (PR). The other properties are not satisfied.

## 7 Using WAFs for Aggregating AFs

Let us now propose new aggregation methods based on WAFs. When WAFs were introduced [Dunne et al., 2011; Coste-Marquis et al., 2012a; 2012b] one of the possible interpretations of the weights on the attack relation that was proposed was that it could represent the number of agents in a group that agree with the attack. So we endorse this interpretation and study how we can define operators that aggregate a set of AFs using techniques developed for WAFs.

### 7.1 $FUS_{All}$

The first method, noted  $FUS_{All}$ , consists in simply building a WAF where the weights represent the number of agents that agree with (i.e. that have) this attack. Once built, we use one of four *best* methods [Coste-Marquis et al., 2012a] in order to obtain, as output, a set of extensions representing the result of the aggregation of the profile.

**Definition 11**  $FUS_{All}^{best_i^{\sigma, \oplus}}(\hat{A}F) = best_i^{\sigma, \oplus}(waf(\hat{A}F))$  where  $waf(\hat{A}F) = \langle A, R, w \rangle$ , with:  $A = \mathcal{X}$ ,  $R = \bigcup_{i=1}^n R_i$ , and  $w(a, b) = |\{AF_i \in \hat{A}F \mid (a, b) \in R_i\}|$ .

Note that the construction of  $waf(\hat{A}F)$  is exactly the one proposed by Cayrol and Lagasque-Schiex [2011] up to a normalization of the weights, but in that work nothing is said about what to do with the obtained WAF. We propose to use the *best* methods in order to find extensions as output.

Let us now check which properties are satisfied by  $FUS_{All}$ . Note that these operators produce as result a set of extensions. Hence, some of the properties, dealing with attacks relation of the result, are not applicable in this framework, namely Unanimous attack (UA), Majority attack (MAJ-A), Closure (CLO), Attack closure (AC), Identity attack (ID-A) and Positive Responsiveness (PR). We recall also that in this paper we focus on the main semantics defined by Dung :  $\sigma \in \{pref, gr, sta, comp\}$ . Finally, concerning the best extensions, we choose to study the four best rules with the sum and the max as aggregation function ( $\oplus \in \{\Sigma, max\}$ ).

**Proposition 5** *Let  $\sigma \in \{comp, pref, sta, gr\}$  be a semantics. Let  $\oplus \in \{\Sigma, max\}$  be an aggregation function.  $FUS_{All}$  satisfies Anonymity (ANON) and properties gr-Identity (gr-ID),  $ca_{gr}$ -Identity ( $ca_{gr}$ -ID) and  $sa_{gr}$ -Identity ( $sa_{gr}$ -ID) for each rule best. The other properties are not satisfied.*

In particular, one of the properties that is not satisfied is non-triviality, meaning that with these operators we do not ensure to always have at least one result as output, which can be considered as an important drawback.

## 7.2 $FUS_{AllNT}$

A solution to satisfy non-triviality is to ensure that the set of extensions of a  $\overline{WF}$  is always non-empty, by using the relaxing extensions techniques [Coste-Marquis *et al.*, 2012b].

### Definition 12

$$FUS_{AllNT}^{\sigma, best_i^{\oplus}, \otimes}(\hat{AF}) = best_i^{\oplus}(\mathcal{E}_{\sigma}^{\otimes}(\overline{waf(\hat{AF})}, waf(\hat{AF})))$$

where  $waf(\hat{AF}) = \langle A, R, w \rangle$ , with:  $A = \mathcal{X}$ ,  $R = \bigcup_{i=1}^n R_i$ , and  $w : w(a, b) = |\{AF_i \in \hat{AF} \mid (a, b) \in R_i\}|$ .

Concerning the aggregation function used for relaxing extensions, we only focus on the sum<sup>4</sup> ( $\otimes = \Sigma$ ).

**Proposition 6** *Let  $\sigma \in \{comp, pref, sta, gr\}$  be a semantics. Let  $\oplus \in \{\Sigma, max\}$  and  $\otimes = \Sigma$  be two aggregation functions.  $FUS_{AllNT}$  satisfies Anonymity (ANON),  $\sigma$ -strong non-triviality ( $\sigma$ -SNT),  $\sigma$ -weak non triviality ( $\sigma$ -WNT) and properties gr-Identity (gr-ID),  $ca_{gr}$ -Identity ( $ca_{gr}$ -ID) and  $sa_{gr}$ -Identity ( $sa_{gr}$ -ID) for each rule best. The other properties are not satisfied.*

This operator is useful if we want to take into account all the attacks given by the agents. However, the result does not represent the opinion of the majority. For instance, suppose that we have nine AFs with  $A = \{a, b\}$  and  $R = \{\}$  and one AF with the same arguments but  $R = \{(a, b)\}$ . If we merge these ten AFs by using an aggregation operator from this family, then the attack  $(a, b)$ , only given by one agent, is present in the resulting system. This is clearly against the opinion of the majority.

## 7.3 $FUS_{MajNT}$

So, it seems that a more natural way of constructing the WAF corresponding to the set of AFs should take into account the notion of *majority* during the construction of the WAF. This

means that, instead of representing all the attacks of the profile, we only select the attacks accepted by a majority of agents.

### Definition 13

$$FUS_{MajNT}^{\sigma, best_i^{\oplus}, \otimes}(\hat{AF}) = best_i^{\oplus}(\mathcal{E}_{\sigma}^{\otimes}(\overline{mwf(\hat{AF})}, mwf(\hat{AF})))$$

where  $mwf(\hat{AF}) = \langle A, R, w \rangle$ , with:  $A = \mathcal{X}$ ,  $R = \{(a, b) \mid |\{AF_i \mid (a, b) \in Att(AF_i)\}| > \frac{n}{2}\}$ , and  $w(a, b) = |\{AF_i \in \hat{AF} \mid (a, b) \in R_i\}|$  if  $(a, b) \in R$ , and  $= 0$  otherwise.

Let us now check what properties are satisfied by  $FUS_{MajNT}$ .

**Proposition 7** *Let  $\sigma \in \{comp, pref, sta, gr\}$  be a semantics. Let  $\oplus \in \{\Sigma, max\}$  and  $\otimes = \Sigma$  be two aggregation functions.  $FUS_{MajNT}$  satisfies Anonymity (ANON),  $\sigma$ -strong non-triviality ( $\sigma$ -SNT),  $\sigma$ -weak non triviality ( $\sigma$ -WNT) and properties gr-Identity (gr-ID),  $ca_{gr}$ -Identity ( $ca_{gr}$ -ID) and  $sa_{gr}$ -Identity ( $sa_{gr}$ -ID) for each rule best. The other properties are not satisfied.*

Note the surprising fact that the properties of *majority* are not satisfied by  $FUS_{MajNT}$ .

## 8 Discussion

Let us sum up our results and discuss their impact. Table 1 summarizes the properties satisfied by all aggregation methods we considered in this paper. A cross  $\times$  means that the property is not satisfied, symbol  $\checkmark$  means that the property is satisfied,  $\checkmark^{\sigma}$  means that the property is satisfied for the semantics  $\sigma$ , and symbol  $-$  means that the property can not be applied to the operator (because the output of the operator is not compatible with the constraint given by the rule).

Properties	$\Delta_{dc}^{\Sigma}$	$\Delta_{dc}^{leximax}$	QV	$FUS_{All}$	$FUS_{AllNT}$	$FUS_{MajNT}$
ANON	$\checkmark$	$\checkmark$	$\times$	$\checkmark$	$\checkmark$	$\checkmark$
$\sigma$ -SNT	$\times$	$\times$	$\times$	$\times$	$\checkmark$	$\checkmark$
$\sigma$ -WNT	$\times$	$\times$	$\checkmark^{gr}$	$\times$	$\checkmark$	$\checkmark$
$\sigma$ -SD	$\times$	$\times$	$\times$	$\times$	$\times$	$\times$
$\sigma$ -WD	$\times$	$\times$	$\checkmark^{gr}$	$\times$	$\times$	$\times$
UA	$\checkmark$	$\checkmark$	$\times$	-	-	-
$\sigma$ -U	$\times$	$\times$	$\times$	$\times$	$\times$	$\times$
$ca_{\sigma}$ -U	$\times$	$\times$	$\times$	$\times$	$\times$	$\times$
$sa_{\sigma}$ -U	$\times$	$\times$	$\times$	$\times$	$\times$	$\times$
MAJ-A	$\checkmark$	$\times$	$\times$	-	-	-
$\sigma$ -MAJ	$\times$	$\times$	$\times$	$\times$	$\times$	$\times$
$ca_{\sigma}$ -MAJ	$\times$	$\times$	$\times$	$\times$	$\times$	$\times$
$sa_{\sigma}$ -MAJ	$\times$	$\times$	$\times$	$\times$	$\times$	$\times$
CLO	$\times$	$\times$	$\times$	-	-	-
AC	$\checkmark$	$\checkmark$	$\checkmark$	-	-	-
$\sigma$ -C	$\times$	$\times$	$\times$	$\times$	$\times$	$\times$
$ca_{\sigma}$ -C	$\times$	$\times$	$\times$	$\times$	$\times$	$\times$
$sa_{\sigma}$ -C	$\times$	$\times$	$\times$	$\times$	$\times$	$\times$
A-ID	$\checkmark$	$\checkmark$	$\times$	-	-	-
$\sigma$ -ID	$\checkmark$	$\checkmark$	$\times$	$\checkmark^{gr}$	$\checkmark^{gr}$	$\checkmark^{gr}$
$ca_{\sigma}$ -ID	$\checkmark$	$\checkmark$	$\times$	$\checkmark^{gr}$	$\checkmark^{gr}$	$\checkmark^{gr}$
$sa_{\sigma}$ -ID	$\checkmark$	$\checkmark$	$\times$	$\checkmark^{gr}$	$\checkmark^{gr}$	$\checkmark^{gr}$
PR	$\checkmark$	$\checkmark$	$\checkmark$	-	-	-

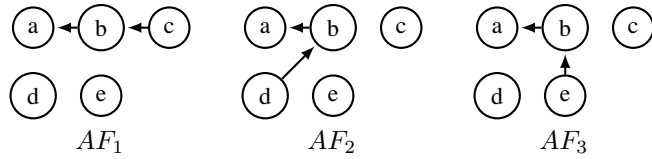
Table 1: Properties of aggregation methods

<sup>4</sup>That is the original definition by Dunne *et al.* [2012].

It is clear that there are few properties satisfied by the existing aggregation operators. There are two possible (non-exclusive) explanations: either the existing operators are not good enough, or the “rationality” properties are too demanding. Our point of view is that both are true to some extent; this means that more work is needed both in defining a set of rationality properties that capture more adequately the desirable behaviour of an aggregation operator, and on defining aggregation methods themselves.

Let us first argue that some of the properties are too strong (this means that not satisfying them is not that a disqualifying feature for an aggregation operator). To show that, we will use a simple and intuitive example which contradicts nine properties.

**Example 1** Let  $AF_1$ ,  $AF_2$  and  $AF_3$  be the three argumentation frameworks represented in Figure 1. They all have the same (unique) complete/preferred/stable/grounded extension:  $\{a, c, d, e\}$ .



Target intuitive aggregated outcome:

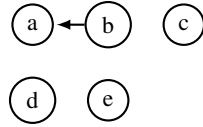


Figure 1: Counterexample for majority, unanimity, closure

This proposed result can seem, at first sight, illogical if we look at the obtained extensions (and consequently the accepted arguments) since each  $AF_i$  has the same extension  $\{a, c, d, e\}$ , so we could expect this extension to be the outcome. But if we look more closely at this example, we can see that this same extension is obtained for very different reasons. Each agent has a reason (argument) to reject  $b$ , but this attack is challenged by all the other agents, and thus could be interpreted as an error of this agent. So it is quite natural to refuse all attacks on  $b$  for the outcome aggregation framework, and that means that  $\{a, c, d, e\}$  should be rejected as the extension of the outcome (and that  $\{b, d, c, e\}$  is much more natural). In fact, it is the outcome that is obtained by the majority method (majority vote on the attack relation): all the  $AF_i$  agree on the attack  $(b, a)$ , whereas all other attacks have a maximum of one  $AF_i$  supporting it. This is also the result obtained with  $FUS_{MAJNT}$  operators. However, that goes against the properties of unanimity, majority and closure related to extensions, credulous inference and skeptical inference<sup>5</sup>. So if one wants to obtain the expected outcome of this example, only the properties about the attack relations seem not problematic.

<sup>5</sup> $\sigma$ -MAJ,  $ca_\sigma$ -MAJ,  $sa_\sigma$ -MAJ,  $\sigma$ -U,  $ca_\sigma$ -U,  $sa_\sigma$ -U, CLO,  $\sigma$ -C,  $ca_\sigma$ -C, and  $sa_\sigma$ -C

We want to insist on the fact that we give here a simple example with only three bases, but this example can be generalized with 100 agents (and 102 arguments), such that  $(b, a)$  is supported by all, and each agent  $i$  supports only an additional attack between argument  $a_i$  and  $b$  (and he is the only one to support it). In this case the quasi-unanimity situation (all agents except one are against the other attacks) is much more striking.

The properties of decisiveness seem also much too strong requirements for most semantics that accept several extensions (and are trivial for the ones that accept at most one extension), so we propose to remove them from necessary properties also.

Basically our opinion is that the proposed properties were more or less direct translations of properties coming from social choice theory. This was certainly an important first step. However, argumentation frameworks have more structure than sets of candidates in voting problems, so the specificities of this structure of AFs have to be taken into account. We argue that these structural specificities invalidate some of the properties from social choice theory as being required for aggregation of AFs. This does not mean that they are not of interest, since they can be used to characterize some aggregation methods (they should be some methods that satisfy them), but they can not be considered as absolutely necessary requirements.

The shaded rows in Table 1 contain, in our opinion, the most desirable properties if one concentrates on the attack relation. And one can see that there is no existing aggregation method that fully satisfies all of these important properties. So this means that there is still work needed to define good aggregation operators. Indeed, Example 1 illustrates that there seem to be some incompatibilities between the rationality properties for aggregation of argumentation frameworks that deal with extensions and the ones that deal with attacks. Both approaches seems sensible, so this means that they should be two different sets of postulates, depending on the chosen priority, i.e. one that focuses on extensions and one that concentrates on the attack relation.

## 9 Conclusion

In this paper we put together the works from the literature on aggregation methods for Dung’s abstract argumentation frameworks. We focus on the methods that take as input a profile of abstract argumentation frameworks, and give as result an argumentation framework, set of argumentation frameworks, or a set of extensions.

We also investigate the use of WAFs in order to aggregate profiles of AFs, and we end up with three possible definitions,  $FUS_{MAJNT}$  being certainly the most convincing.

We show that few of the proposed properties are satisfied by existing aggregation operators. The explanation seems to incriminate both suspects: the properties and the methods. At one hand, some of the properties seem to be too demanding in the general case. At the other hand, the existing operators do not satisfy even the most desirable properties.

Our results seem to suggest that a lot of work is still needed on the two fronts. A more careful study of the rationality

properties for aggregation methods for abstract argumentation is required. And there is clearly room for definition of other (better?) aggregation methods.

We plan to study if using properties from propositional belief merging [Konieczny and Pino Pérez, 2002] could be more appropriate. Indeed, these properties were defined for a framework (propositional logic) that also has some structure. So we are working on the translation of these properties for argumentation framework merging, similarly to what have been done recently for belief revision [Coste-Marquis *et al.*, 2014a], in order to define maybe more adequate rationality properties, and get some ideas to define new aggregation methods.

## Acknowledgments

This work benefited from the support of the project AMANDE ANR-13-BS02-0004 of the French National Research Agency (ANR).

## References

- [Amgoud and Vesic, 2009] Leila Amgoud and Srdjan Vesic. On revising argumentation-based decision systems. In *10th European Conference on Symbolic and Quantitative Approaches to Reasoning with Uncertainty (EC-SQARU'09)*, pages 71–82. Springer, 2009.
- [Baumann, 2012] Ringo Baumann. Normal and strong expansion equivalence for argumentation frameworks. *Artificial Intelligence*, 193:18–44, 2012.
- [Boella *et al.*, 2009] Guido Boella, Souhila Kaci, and Leendert van der Torre. Dynamics in argumentation with single extensions: Abstraction principles and the grounded extension. In *10th European Conference on Symbolic and Quantitative Approaches to Reasoning with Uncertainty (EC-SQARU'09)*, pages 107–118. Springer, 2009.
- [Bonzon and Maudet, 2011] Elise Bonzon and Nicolas Maudet. On the outcomes of multiparty persuasion. In *10th International Conference on Autonomous Agents and Multiagent Systems (AAMAS'11)*, pages 47–54, 2011.
- [Cayrol and Lagasquie-Schiex, 2011] Claudette Cayrol and Marie-Christine Lagasquie-Schiex. Weighted argumentation systems: A tool for merging argumentation systems. In *IEEE 23rd International Conference on Tools with Artificial Intelligence (ICTAI'11)*, pages 629–632, 2011.
- [Cayrol *et al.*, 2008] Claudette Cayrol, Florence Dupin de Saint-Cyr, and Marie-Christine Lagasquie-Schiex. Revision of an argumentation system. In *11th International Conference on the Principles of Knowledge Representation and Reasoning (KR'08)*, pages 124–134, 2008.
- [Coste-Marquis *et al.*, 2007] Sylvie Coste-Marquis, Caroline Devred, Sébastien Konieczny, Marie-Christine Lagasquie-Schiex, and Pierre Marquis. On the merging of Dung's argumentation systems. *Artificial Intelligence*, 171(10-15):730–753, 2007.
- [Coste-Marquis *et al.*, 2012a] Sylvie Coste-Marquis, Sébastien Konieczny, Pierre Marquis, and Mohand-Akli Ouali. Selecting extensions in weighted argumentation frameworks. In *4th International Conference on Computational Models of Argument (COMMA'12)*, pages 342–349, 2012.
- [Coste-Marquis *et al.*, 2012b] Sylvie Coste-Marquis, Sébastien Konieczny, Pierre Marquis, and Mohand-Akli Ouali. Weighted attacks in argumentation frameworks. In *13th International Conference on the Principles of Knowledge Representation and Reasoning (KR'12)*, pages 593–597, 2012.
- [Coste-Marquis *et al.*, 2014a] Sylvie Coste-Marquis, Sébastien Konieczny, Jean-Guy Mailly, and Pierre Marquis. On the revision of argumentation systems: Minimal change of arguments statuses. In *14th International Conference on the Principles of Knowledge Representation and Reasoning (KR'14)*, pages 52–61, 2014.
- [Coste-Marquis *et al.*, 2014b] Sylvie Coste-Marquis, Sébastien Konieczny, Jean-Guy Mailly, and Pierre Marquis. A translation-based approach for revision of argumentation frameworks. In *14th European Conference on Logics in Artificial Intelligence (JELIA'14)*, pages 77–85, 2014.
- [da Costa Pereira *et al.*, 2011] Célia da Costa Pereira, Andrea Tettamanzi, and Serena Villata. Changing one's mind: Erase or rewind? In *22nd International Joint Conference on Artificial Intelligence (IJCAI'11)*, pages 164–171, 2011.
- [Dung, 1995] Phan Minh Dung. On the acceptability of arguments and its fundamental role in nonmonotonic reasoning, logic programming and n-person games. *Artificial Intelligence*, 77(2):321–358, 1995.
- [Dunne *et al.*, 2011] Paul E. Dunne, Anthony Hunter, Peter McBurney, Simon Parsons, and Michael Wooldridge. Weighted argument systems: Basic definitions, algorithms, and complexity results. *Artificial Intelligence*, 175(2):457–486, 2011.
- [Dunne *et al.*, 2012] Paul E. Dunne, Pierre Marquis, and Michael Wooldridge. Argument aggregation: Basic axioms and complexity results. In *4th International Conference on Computational Models of Argument (COMMA'12)*, pages 129–140, 2012.
- [Gabbay and Rodrigues, 2012] Dov M. Gabbay and Odinaldo Rodrigues. A numerical approach to the merging of argumentation networks. In *13th International Workshop on Computational Logic in Multi-Agent Systems (CLIMA'12)*, pages 195–212, 2012.
- [Konieczny and Pino Pérez, 2002] Sébastien Konieczny and Ramón Pino Pérez. Merging information under constraints: a logical framework. *Journal of Logic and Computation*, 12(5):773–808, 2002.
- [Tohmé *et al.*, 2008] Fernando A. Tohmé, Gustavo Adrian Bodanza, and Guillermo Ricardo Simari. Aggregation of attack relations: A social-choice theoretical analysis of defeasibility criteria. In *5th International Symposium on Foundations of Information and Knowledge Systems (FoIKS'08)*, pages 8–23, 2008.