Aggregation of Attack Relations: A Social-Choice Theoretical Analysis of Defeasibility Criteria*

Fernando A. Tohmé, Gustavo A. Bodanza, and Guillermo R. Simari

Artificial Intelligence Research and Development Laboratory (LIDIA)
Universidad Nacional del Sur, Av.Alem 1253, (8000) Baha Blanca, Argentina
Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET)
{ftohme,ccbodanza}@criba.edu.ar, grs@cs.uns.edu.ar

Abstract. This paper analyzes the aggregation of different abstract attack relations over a common set of arguments. Each of those attack relations can be considered as the representation of a *criterion* of warrant. It is well known in the field of Social Choice Theory that if some "fairness" conditions are imposed over an aggregation of preferences, it becomes impossible to yield a result. When the criteria lead to *acyclic* attack relations, a positive result may ensue under the same conditions, namely that if the class of winning coalitions in an aggregation process by voting is a *proper prefilter* an outcome will exist. This outcome may preserve some features of the competing attack relations, such as the highly desirable property of acyclicity which can be associated with the existence of a single extension of an argumentation system. The downside of this is that, in fact, the resulting attack relation must be a portion common to the "hidden dictators" in the system, that is, all the attack relations that belong to all the winning coalitions.

1 Introduction

Defeasible reasoning relies on the possibility of comparing conclusions in terms of their support. This support is often given by a set of arguments. Only those arguments (and consequently their conclusions) that remain undefeated in a series of comparisons are deemed warranted. While the literature contains alternative formalisms capturing this intuition [1,2], the groundbreaking work on Abstract Argumentation Frameworks reported in [3] presents a view according to which all the features that are not essential for the study of the attack relation in defeasible argumentation are eliminated. What remains is a system formed by a family of abstract arguments and a relation of *attack* among them. Several alternative semantics have been introduced, but the essential idea is that the set of arguments that survive all possible attacks of other arguments in the system constitute the so-called *extensions* of the system and capture its semantics.

One aspect that has received little attention in the literature[1] is the possibility of considering different relations of attack among the same arguments. In this scenario, the warrant of arguments cannot be established in an unambiguous way without first

* Partially supported by SeCyT - Universidad Nacional del Sur, CONICET, and ANPCYT.

[1] A remarkable exception being [4].

S. Hartmann and G. Kern-Isberner (Eds.): FoIKS 2008, LNCS 4932, pp. 8–23, 2008.

coalescing the multiple attack relation onto a single one acting over the family of arguments. Notice that this ensuing relation does not need to coincide with anyone of those defined over the arguments. On the other hand, nothing precludes this possibility.

In Economics, the process by which a single preference ordering is obtained given a class of individual preferences over the same alternatives, it is known as an *aggregation* of them [5]. Similarly, we can consider that each of the attack relations among arguments represents an individual *criterion* of warrant, since it defines which extensions should obtain. Then, the aggregation process weights up the different criteria and determines which extensions will actually appear, but instead of simply enumerating extensions, it yields an attack relation that supports them.

While there might exist many ways of doing this, a natural form is by means of pairwise voting [6]. That is, each alternative attack relation "votes" over pairs of arguments, and the winning relation over those two arguments is incorporated in the aggregate attack relation.

But such procedure has been shown to have, in certain contexts, serious shortcomings. It is widely known that it may fail to verify some required constraints over the aggregation process [7,8]. These constraints are actually desiderata for a fair aggregation process. Social Choice Theory (SCT) has been studying them for over fifty years and it seems natural to transfer its results to the problem of aggregating attack relations. Arrow's Impossibility Theorem [9] claims that four quite natural constraints, that capture abstractly the properties of a democratic aggregation process, cannot be simultaneously satisfied. That is true for the case of reflexive and transitive preference relations over the alternatives. Once those constraints become incorporated in the framework of argumentation, we could expect something like Arrow's theorem to ensue. But attack relations and preference relations are different in many respects. This point must be emphasized, since it involves the reason why an Arrow-like result may not be a necessary outcome for argumentation systems. This fact makes our purpose non trivial.

The difference between aggregating individual preferences and attack criteria originates from their corresponding order-theoretic characterizations. While in Economics preferences are usually assumed to be *weak orders* (i.e., reflexive, transitive and complete relations), attack relations are free to adopt any configuration. On the other hand, preference relations are expected to have maximal elements, while this is not the case for attack relations. If A attacks B and B attacks C, it is commonly accepted that not only A does not (necessarily) attack C, but that A "defends" C, which implies that A and C can be jointly warranted. So, while a preference relation can lead to the choice of its maximal elements, and attack relation can lead to the choice of a maximal (w.r.t. $\subseteq$) set of "defensible" arguments.

Viewed as criteria of acceptance, the choices should verify at least a minimal degree of rationality. In SCT that requirement is fulfilled by the condition that chosen options should not be transitively better than themselves, i.e. they should not be part of cycles of preference [10]. In the context of attack criteria this condition can be interpreted as that each of the arguments that will be deemed warranted under a criterion should be supported by chains of attacks that do not include themselves. A sufficient condition that ensures this is the *acyclicity* of the attack relations.[2]

[2] An argumentation framework in which the attack relation is acyclic is said *well-founded* [3].

Once we require the acyclicity of the attack relations we look for aggregation processes that have as inputs finite numbers of acyclic attack relations and output also acyclic relations. In that case, as we will show in this paper, under the same conditions of fairness as Arrow's Theorem, we can prove the existence of an aggregate attack relation. In fact, following [7] we show that the class of winning coalitions of attack criteria constitutes an algebraic structure called a *proper prefilter*.

As it has been discussed in the literature on Arrow's Theorem, a prefilter indicates the existence of a *collegium* of attack relations. Each member of the collegium belongs to a winning coalition, while the collegium itself does not need to be one. Each collegium member, by itself, cannot determine the outcome of the aggregation process, but can instead veto the behaviors that run contrary to its prescription. The final outcome can be seen as the agreement of the representatives of the different winning coalitions. In this sense it indicates a very basic consensus among the attack relations.

In a sense this means that even in the case of "equal opportunity" aggregation procedures there will exist some fragment of the individual attack relations that will become imposed on the aggregate one. But while in social context this seems rather undesirable (in the literature the members of the collegium are called *hidden dictators*), in the case of argument systems is far more reassuring, since it indicates that when the attack relations are minimally rational, a consensual outcome may arise.

2 Aggregating Attack Relations

Dung defines an argumentation framework as a pair $AF = \langle AR; \rightarrow \rangle$, where AR is a set of abstract entities called 'arguments' and $\rightarrow \subseteq AR \times AR$ denotes an attack relation among arguments. This relation determines which sets of arguments become "defended" from attacks. Different characterizations of the notion of defense yield alternative sets called *extensions* of AF. These extensions are seen as the semantics of the argumentation framework, i.e. the classes of arguments that can be deemed as the outcomes of the whole process of argumentation. Dung introduces the notions of *preferred*, *stable*, *complete*, and *grounded* extensions, each corresponding to different requirements on the attack relation.

Definition 1. (Dung ([3])). *In any argumentation framework AF an argument σ is said* acceptable *w.r.t. a subset S of arguments of AR, in case that for every argument τ such that $\tau \rightarrow \sigma$, there exists some argument $\rho \in S$ such that $\rho \rightarrow \tau$. A set of arguments S is said* admissible *if each $\sigma \in S$ is acceptable w.r.t. S, and is conflict-free, i.e., the attack relation does not hold for any pair of arguments belonging to S. A* preferred extension *is any maximally admissible set of arguments of AF. A* complete extension *of AF is any conflict-free subset of arguments which is a fixed point of $\Phi(\cdot)$, where $\Phi(S) = \{\sigma : \sigma \text{ is acceptable w.r.t. } S\}$, while the* grounded extension *is the least (w.r.t. $\subseteq$) complete extension. Moreover, a* stable extension *is a conflict-free set S of arguments which attacks every argument not belonging to S.*

Interestingly, if the attack relation is acyclic, the framework has only one extension that is grounded, preferred stable and complete (*cf.* [3], theorem 30, pp. 331). The main application of argumentation frameworks is the field of *defeasible reasoning*. Roughly, arguments are structures that support certain conclusions (claims). The extensions include the arguments, and more importantly their conclusions, that become warranted by a reasoning process that considers the attack relation.

We consider, instead, for a given n an *extended* argumentation framework $AF^n = \langle AR;\ \rightarrow_1, \ldots, \rightarrow_n \rangle$. Each $\rightarrow_i$ is a particular attack relation among the arguments in AR, representing different criteria according to which arguments are evaluated one against another. Such extended frameworks may arise naturally in the context of defeasible reasoning, since there might exist more than one criterion of defeat among arguments.

The determination of *preferred*, *complete* or *grounded* extensions in an argumentation framework is based upon the properties of the single attack relation. There are no equivalent notions for an extended argumentation framework, except for those corresponding to an *aggregate* argumentation framework $AF^* = \langle AR;\ \mathcal{F}(\rightarrow_1, \ldots, \rightarrow_n)\rangle$, where $\mathcal{F}(\rightarrow_1, \ldots, \rightarrow_n) = \rightarrow$, *i.e.*, $\mathcal{F}(\rightarrow_1, \ldots, \rightarrow_n)$ is the aggregated attack relation of AF^*. That is, AF^* is a an argumentation framework in which its attack relation arises as a function of the attack relations of AF^n. Notice that $\mathcal{F}$ may be applied over any extended argumentation framework with n attack relations. It embodies a method that yields a single attack relation up from n alternatives.

To postulate an aggregate relation addresses the problem of managing the diversity of criteria, by yielding a single approach. This is of course analogous to a social system, in which a unified criterion must by reached. While there exist many alternative ways to aggregate different criteria, most of them are based in some form of voting. In fact, the best known case of $\mathcal{F}$ is *majority voting*. Unlike political contests in which for each pair $A, B \in AR$ a majority selects either $A \rightarrow B$ or $B \rightarrow A$, we allow for a third alternative in which the majority votes for the absence of attacks between A and B. Formally:

- $A \rightarrow B$ if $|\{i : A \rightarrow_i B\}| > \max(|\{i : B \rightarrow_i A\}|, |\{i : B \not\rightarrow_i A \wedge A \not\rightarrow_i B\}|)$.
- $B \rightarrow A$ if $|\{i : B \rightarrow_i A\}| > \max(|\{i : A \rightarrow_i B\}|, |\{i : B \not\rightarrow_i A \wedge A \not\rightarrow_i B\}|)$.
- $(A, B) \notin \rightarrow$ (*i.e.*, A does not attack B, nor B does attack A in $\rightarrow$) if $|\{i : B \not\rightarrow_i A \wedge A \not\rightarrow_i B\}| > \max(|\{i : A \rightarrow_i B\}|, |\{i : B \rightarrow_i A\}|)$.

For instance, if out of 100 individual relations, 34 are such that A attacks B, while 33 verify that B attacks A and the rest that there is no attack relation between A and B, majority voting would yield that $A \rightarrow B$. That is, it only matters which alternative is verified by more individual relations than the other two.

Example 1. Consider the following framework in which $AR = \{A, B, C\}$ and the arguments are:

A : "Symptoms x, y and z suggest the presence of disease d_1, so we should apply therapy t_1";
B : "Symptoms x, w and z suggest the presence of disease d_2, so we should apply therapy t_2";
C : "Symptoms x and z suggest the presence of disease d_3, so we should apply therapy t_3".

Assume these are the main arguments discussed in a group of three agents (M.D.s), 1, 2 and 3, having to make a decision on which therapy should be applied to some patient. Suppose that each agent i, $i \in \{1, 2, 3\}$, proposes an attack relation $\rightarrow_i$ over the arguments as follows:

- $\rightarrow_1 = \{(A, B), (B, C)\}$ (agent 1 thinks that it is not convenient to make a joint application of therapies t_1 and t_2 or t_2 and t_3; moreover she thinks that B is more specific than C, hence B defeats C, and that, in the case at stake, symptom y is more clearly present than symptom w. Hence A defeats B),
- $\rightarrow_2 = \{(A, C), (B, C)\}$ (agent 2 thinks that it is not convenient to apply therapies t_1 together with t_3 or t_2 joint with t_3; moreover she thinks that symptoms y and w are equally present in the case at stake. Furthermore, both A and B are more specific than argument C, hence both A and B defeat C),
- $\rightarrow_3 = \{(A, C), (C, B)\}$ (agent 3 thinks that it is not convenient to apply t_1 together with t_3 or t_2 with t_3; moreover she thinks that symptom w is not clearly detectable, hence C defeats B, but A is more specific than C, hence A defeats C).

According to majority voting we obtain $\rightarrow$ over AR:

- $A \rightarrow C$ since A attacks C under $\rightarrow_2$ and $\rightarrow_3$.
- $B \rightarrow C$ since B attacks C under $\rightarrow_1$ and $\rightarrow_2$.
- $(A, B) \notin \rightarrow$ since $(A, B) \notin \rightarrow_2$ and $(A, B) \notin \rightarrow_3$.

In this example, majority voting picks out one of the individual attack relations, showing that $\rightarrow = \rightarrow_2$.[3]

On the other hand, majority voting may yield cycles of attacks up from acyclical individual relations:

Example 2. Consider the following three attack relations over the set $AR = \{A, B, C\}$: $C \rightarrow_1 B \rightarrow_1 A$, $A \rightarrow_2 C \rightarrow_2 B$, and $B \rightarrow_3 A \rightarrow_3 C$. We obtain $\rightarrow$ over AR as follows:

- $A \rightarrow C$ since A attacks C under $\rightarrow_2$ and $\rightarrow_3$.
- $B \rightarrow A$ since B attacks C under $\rightarrow_1$ and $\rightarrow_3$.
- $C \rightarrow B$ since C attacks B under $\rightarrow_1$ and $\rightarrow_2$.

Thus, $\rightarrow$ yields a cycle $A \rightarrow C \rightarrow B \rightarrow A$. This phenomenon is known in the literature on voting systems as the *Condorcet's Paradox* and it shows clearly that even the most natural aggregation procedures may have drawbacks.

Another way of aggregating attack relations is by restricting majority voting to a *qualified voting* aggregation function. It fixes a given class of relations as those that will have more weight in the aggregate. Then, the outcome of majority voting over a pair of arguments is imposed on the aggregate only if the fixed attack relations belong to the majority. Otherwise, in the attack relation none of the arguments attacks the other. That is, given a set $U \subset \{1, \ldots, n\}$:

[3] In any of the extension semantics introduced by [3], arguments A and B become justified under the aggregate attack relation, supporting the decision of applying both therapies t_1 and t_2.

- $A \rightarrow B$ iff $|\{i : A \rightarrow_i B\}| > \max(|\{i : B \rightarrow_i A\}|, |\{i : B \not\rightarrow_i A \wedge A \not\rightarrow_i B\}|)$ and $U \subseteq \{i : A \rightarrow_i B\}$.
- $B \rightarrow A$ iff $|\{i : B \rightarrow_i A\}| > \max(|\{i : A \rightarrow_i B\}|, |\{i : B \not\rightarrow_i A \wedge A \not\rightarrow_i B\}|)$ and $U \subseteq \{i : B \rightarrow_i A\}$.

$(A, B) \notin \rightarrow$ (*i.e.*, A does not attack B, nor B does attack A in $\rightarrow$) can arise as follows:

- either $|\{i : B \not\rightarrow_i A \wedge A \not\rightarrow_i B\}| > \max(|\{i : A \rightarrow_i B\}|, |\{i : B \rightarrow_i A\}|)$ and $U \subseteq \{i : B \not\rightarrow_i A \wedge A \not\rightarrow_i B\}$,
- or if U is not a subset of either $\{i : A \rightarrow_i B\}$, $\{i : B \rightarrow_i A\}$ or $\{i : B \not\rightarrow_i A \wedge A \not\rightarrow_i B\}$.

Example 3. Consider again the individual attack relations in Example 1. If $U = \{2, 3\}$ we have that $A \rightarrow C$, since $A \rightarrow_2 C$ and $A \rightarrow_3 C$. Again $(A, B) \notin \rightarrow$ because $(A, B) \notin \rightarrow_2$ and $(A, B) \notin \rightarrow_3$. But we have also that $(B, C) \notin \rightarrow$ because although there exists a majority for B attacking C ($\{1, 2\}$), $C \rightarrow_3 B$, *i.e.*, there is no consensus among the members of U on B and C.

3 Arrow's Conditions on Aggregation Functions

While different schemes of aggregation of attack relations can be postulated, most of SCT, up from the seminal work of Kenneth Arrow [9] points towards a higher degree of abstraction. Instead of looking for particular functional forms, the goal is to set general constraints over aggregation processes and see if they can be jointly fulfilled. We carry out a similar exercise in the setting of extended argumentation frameworks, in order to investigate the features of aggregation processes that ensure that a few reasonable axioms are satisfied.

Social choice-theoretic analysis can be carried out in terms of an aggregation process that, up from a family of *weak orders* (complete, transitive and reflexive orderings), yields a weak order over the same set of alternatives. This is because both individual and social *preference* relations are represented as weak orders. But attack relations cannot be assimilated to *preference* orderings, since attacks do not verify necessarily any of the conditions that define a weak order.[4] Therefore, the difference of our setting with the usual Arrovian context is quite significant.

Let us begin with a few properties that, very much like in SCT, we would like to be verified in any aggregation function. Below, we will use the alternative notation $\rightarrow_{\mathcal{F}}$ instead of $\mathcal{F}(\rightarrow_1, \ldots, \rightarrow_n)$ when no confusion could arise.

- **Pareto condition.** For all $A, B \in AR$ if for every $i = 1, \ldots, n$, $A \rightarrow_i B$ then $A \rightarrow_{\mathcal{F}} B$.
- **Positive Responsiveness.** For all $A, B \in AR$, and two n-tuples of attack relations, $(\rightarrow_1, \ldots, \rightarrow_n)$, $(\rightarrow'_1, \ldots, \rightarrow'_n)$, if $\{i : A \rightarrow_i B\} \subseteq \{i : A \rightarrow'_i B\}$ and $A \rightarrow_{\mathcal{F}} B$, then $A \rightarrow'_{\mathcal{F}} B$, where $\rightarrow'_{\mathcal{F}} = \mathcal{F}(\rightarrow'_1, \ldots, \rightarrow'_n)$.

[4] So for instance, reflexivity in an attack relation would mean that each argument attacks itself. While isolated cases of self-attack may arise, this is not a general feature of attack relations. The same is true of transitivity that means that if, say $A \rightarrow B$ and $B \rightarrow C$ then $A \rightarrow C$. In fact, in many cases of interest, $A \rightarrow B \rightarrow C$ can be interpreted as indicating that A *defends* B. Finally, completeness is by no means a necessary feature of attacks, since there might exist at least two arguments A and B such that neither $A \rightarrow B$ and $B \rightarrow A$.

- **Independence of Irrelevant Alternatives.** For all $A, B \in AR$, and given two n-tuples of attack relations, $(\rightarrow_1, \ldots, \rightarrow_n)$, $(\rightarrow_1^{'}, \ldots, \rightarrow_n^{'})$, if $\rightarrow_i = \rightarrow_i^{'}$ for each i, over (A, B), then $\rightarrow_{\mathcal{F}} = \rightarrow_{\mathcal{F}}^{'}$ over (A, B).
- **Non-dictatorship.** There does not exist i_0 such that for all $A, B \in AR$ and every $(\rightarrow_1, \ldots, \rightarrow_n)$, if $A \rightarrow_{i_0} B$ then $A \rightarrow_{\mathcal{F}} B$.

All these requirements were intended to represent the abstract features of a democratic collective decision-making system. While in our setting this does no longer apply, we still consider that an aggregation function should yield a fair representative of the whole class of attack relations. Let us see why these conditions imply the fairness of the aggregation process.[5]

The *Pareto condition* indicates that if all the attacks relations coincide over a pair of arguments, the aggregate attack should also agree with them. That is, if all the individual attack relations agree on some arguments, this agreement should translate into the aggregate attack relation.

The *positive responsiveness condition* just asks that the aggregation function should yield the same outcome over a pair of arguments if some attack relation previously dissident over them, now change towards an agreement with the others. It can be better understood in terms of political elections: if a candidate won an election, she should keep winning in an alternative context in which somebody who voted against her now turns to vote for her.

The axiom of *independence of irrelevant alternatives* just states that if there is an agreement over a pair of arguments among alternative n-tuples of attacks, this should be also be true for the aggregation function over both n-tuples. Again, some intuition from political elections may be useful. If the individual preferences over two candidates a and b remain the same when a third candidate c arises, the rank of a and b should be the same in elections with and without c. That is, the third party should be irrelevant to the other two.

Finally, the *non-dictatorship condition* just stipulates that no fixed entry in the n-tuples of attacks should become the outcome in every possible instance. That is, there is no 'dictator' among the individual attack relations. We have the following proposition:

Proposition 1
Both the majority and the qualified voting (with $|U| \geq 2$) aggregation functions verify trivially the four axioms.

PROOF
Majority voting:

- *(Pareto): if for all $A, B \in AR$ if for every $i = 1, \ldots, n$, $A \rightarrow_i B$ then trivially $|\{i : A \rightarrow_i B\}| > \max(|\{i : B \rightarrow_i A\}|, |\{i : B \not\rightarrow_i A \wedge A \not\rightarrow_i B\}|)$ which in turns implies that $A \rightarrow_{\mathcal{F}} B$.*

[5] Whether fairness is *exactly* captured by these requirements is still debated in the philosophy of Social Choice. Nevertheless, there exists a consensus on that they are desirable conditions for an aggregation function.

- *(Positive responsiveness): if for all $A, B \in AR$, and two n-tuples of attack relations, $(\rightarrow_1, \ldots, \rightarrow_n)$, $(\rightarrow_1^{'}, \ldots, \rightarrow_n^{'})$, if $A \rightarrow_{\mathcal{F}} B$, this means that $|\{i : A \rightarrow_i B\}| > \max(|\{i : B \rightarrow_i A\}|, |\{i : B \not\rightarrow_i A \wedge A \not\rightarrow_i B\}|)$ and therefore, if $\{i : A \rightarrow_i B\} \subseteq \{i : A \rightarrow_i^{'} B\}$ it follows that $|\{i : A \rightarrow_i^{'} B\}| > \max(|\{i : B \rightarrow_i^{'} A\}|, |\{i : B \not\rightarrow_i^{'} A \wedge A \not\rightarrow_i^{'} B\}|)$ which in turn implies that $A \rightarrow_{\mathcal{F}}^{'} B$.*
- *(Independence of Irrelevant Alternatives): suppose that for any given $A, B \in AR$, and two n-tuples of attack relations, $(\rightarrow_1, \ldots, \rightarrow_n)$, $(\rightarrow_1^{'}, \ldots, \rightarrow_n^{'})$, $\rightarrow_i = \rightarrow_i^{'}$ for each i, over (A, B). Without loss of generality assume that $|\{i : A \rightarrow_i B\}| > \max(|\{i : B \rightarrow_i A\}|, |\{i : B \not\rightarrow_i A \wedge A \not\rightarrow_i B\}|)$ then, $A \rightarrow_{\mathcal{F}} B$. But then $|\{i : A \rightarrow_i^{'} B\}| > \max(|\{i : B \rightarrow_i^{'} A\}|, |\{i : B \not\rightarrow_i^{'} A \wedge A \not\rightarrow_i^{'} B\}|)$, which implies that $A \rightarrow_{\mathcal{F}}^{'} B$. That is, $\rightarrow_{\mathcal{F}} = \rightarrow_{\mathcal{F}}^{'}$ over (A, B).*
- *(Non-dictatorship): suppose there where a i_0 such that for all $A, B \in AR$ and every $(\rightarrow_1, \ldots, \rightarrow_n)$, if $A \rightarrow_{i_0} B$ then $A \rightarrow_{\mathcal{F}} B$. Consider in particular that $A \rightarrow_{i_0} B$ while $|\{i : B \rightarrow_i A\}| = n - 1$, i.e. except i_0 all other attack relations have B attacking A. But then $B \rightarrow_{\mathcal{F}} A$. Contradiction.*

The proof for qualified voting, when $|U| \geq 2$, is quite similar:

- *(Pareto): if for all $A, B \in AR$ if for every $i = 1, \ldots, n$, $A \rightarrow_i B$ then trivially $|\{i : A \rightarrow_i B\}| > \max(|\{i : B \rightarrow_i A\}|, |\{i : B \not\rightarrow_i A \wedge A \not\rightarrow_i B\}|)$ and $U \subseteq \{i : A \rightarrow_i B\}$ which implies that $A \rightarrow_{\mathcal{F}} B$.*
- *(Positive responsiveness): if for all $A, B \in AR$, and two n-tuples of attack relations, $(\rightarrow_1, \ldots, \rightarrow_n)$, $(\rightarrow_1^{'}, \ldots, \rightarrow_n^{'})$, if $A \rightarrow_{\mathcal{F}} B$, this means that $|\{i : A \rightarrow_i B\}| > \max(|\{i : B \rightarrow_i A\}|, |\{i : B \not\rightarrow_i A \wedge A \not\rightarrow_i B\}|)$ and $U \subseteq \{i : A \rightarrow_i B\}$. Therefore, if $\{i : A \rightarrow_i B\} \subseteq \{i : A \rightarrow_i^{'} B\}$ it follows that $|\{i : A \rightarrow_i^{'} B\}| > \max(|\{i : B \rightarrow_i^{'} A\}|, |\{i : B \not\rightarrow_i^{'} A \wedge A \not\rightarrow_i^{'} B\}|)$ and $U \subseteq \{i : A \rightarrow_i^{i} B\}$ which in turn implies that $A \rightarrow_{\mathcal{F}}^{'} B$.*
- *(Independence of Irrelevant Alternatives): suppose that for any given $A, B \in AR$, and two n-tuples of attack relations, $(\rightarrow_1, \ldots, \rightarrow_n)$, $(\rightarrow_1^{'}, \ldots, \rightarrow_n^{'})$, $\rightarrow_i = \rightarrow_i^{'}$ for each i, over (A, B). Without loss of generality assume that $|\{i : A \rightarrow_i B\}| > \max(|\{i : B \rightarrow_i A\}|, |\{i : B \not\rightarrow_i A \wedge A \not\rightarrow_i B\}|)$ and $U \subseteq \{i : A \rightarrow_i B\}$ then, $A \rightarrow_{\mathcal{F}} B$. But then $|\{i : A \rightarrow_i^{'} B\}| > \max(|\{i : B \rightarrow_i^{'} A\}|, |\{i : B \not\rightarrow_i^{'} A \wedge A \not\rightarrow_i^{'} B\}|)$ and also $U \subseteq \{i : A \rightarrow_i^{'} B\}$ which implies that $A \rightarrow_{\mathcal{F}}^{'} B$. That is, $\rightarrow_{\mathcal{F}} = \rightarrow_{\mathcal{F}}^{'}$ over (A, B).*
- *(Non-dictatorship): suppose there where a i_0 such that for all $A, B \in AR$ and every $(\rightarrow_1, \ldots, \rightarrow_n)$, if $A \rightarrow_{i_0} B$ then $A \rightarrow_{\mathcal{F}} B$. Consider in particular that $A \rightarrow_{i_0} B$ while $|\{i : B \rightarrow_i A\}| = n - 1$, i.e. except i_0 all other attack relations have B attacking A. If $i_0 \notin U$, $B \rightarrow_{\mathcal{F}} A$, while if $i_0 \in U$, $(A, B) \notin \rightarrow_{\mathcal{F}}$. In either case we have a contradiction.* □

While qualified voting seems in certain sense less fair than majority voting it can be shown that it is not prone to phenomena like Condorcet's Paradox:

Proposition 2
If F is a qualified voting aggregation function and each $\rightarrow_i$ is acyclic, then $\rightarrow_{\mathcal{F}}$ is acyclic.

PROOF
Suppose that $\rightarrow_{\mathcal{F}}$ *has a cycle of attacks, say* $A^0 \rightarrow_{\mathcal{F}} A^1 \rightarrow_{\mathcal{F}} \ldots \rightarrow_{\mathcal{F}} A^k \rightarrow_{\mathcal{F}} A^0$*. By definition of qualified voting, for* $j = 0 \ldots k-1$ $A^j \rightarrow_{\mathcal{F}} A^{j+1}$ *if and only if* $U \subseteq \{i : A^j \rightarrow_i A^{j+1}\}$*. By the same token,* $A^k \rightarrow_{\mathcal{F}} A^0$ *iff*

$$U \subseteq \{i : A^k \rightarrow_i A^0\}.$$

That is, for each $i \in U$, $A^0 \rightarrow_i A^1 \rightarrow_i \ldots \rightarrow_i A^k \rightarrow_i A^0$*. But this contradicts that each individual attack relation is acyclic.* □

4 Decisive Sets of Attack Relations

The aggregation function determines a class of *decisive* sets (*i.e.*, winning coalitions) of attack relations. Interestingly, the structure of this class exhibits (in relevant cases) clear algebraic features that shed light on the behavior of the aggregation function. Formally: $\Omega \subset \{1, \ldots, n\}$ be a *decisive* set if for every possible n-tuple $(\rightarrow_1, \ldots, \rightarrow_n)$ and every $A, B \in AR$, if $A \rightarrow_i B$, for *every* $i \in \Omega$, then $A \rightarrow_{\mathcal{F}} B$ (*i.e.*, $A\ \mathcal{F}(\rightarrow_1, \ldots, \rightarrow_n)\ B$). As we have already seen in the case of qualified voting aggregation functions, if not every member of a decisive set agrees with the others over a pair of arguments, the aggregate attack relation should not include the pair. But this is so unless any other decisive set can force the pair of arguments into the aggregate attack relation.[6]

Example 4. In Example 1, each of $\{1, 2\}, \{2, 3\}$ is a decisive set, since they include more than half of the agents that coincide with pairs of attacks in the aggregate attack relation. On the other hand, for the qualified voting function of Example 3, $U = \{2, 3\}$ is decisive, but not $\{1, 2\}$ or $\{1, 3\}$.

If we recall that the U is a decisive set for qualified voting, we can conjecture that there might exist a close relation between the characterization of an aggregation function and the class of its decision sets. Furthermore, if a function verifies Arrow's axioms and yields an acyclic attack relation up from acyclic individual attack relations, it can be completely characterized in terms of the class of its decision sets:

Proposition 3
Consider an aggregate attack relation $\mathcal{F}$ *that for every n-tuple* $(\rightarrow_1, \ldots, \rightarrow_n)$ *of acyclic attack relations yields an acyclic* $\rightarrow_{\mathcal{F}}$*. It verifies the Pareto condition, Positive Responsiveness, Independence of Irrelevant Alternatives, and Non-Dictatorship if and only if its class of decisive sets* $\bar{\Omega} = \{\Omega^j\}_{j \in J}$ *verifies the following properties:*

- $\{1, \ldots, n\} \in \bar{\Omega}$.
- *If* $O \in \bar{\Omega}$ *and* $O \subseteq O'$ *then* $O' \in \bar{\Omega}$.
- *Given* $\bar{\Omega} = \{\Omega^j\}_{j \in J}$*, where* $J = |\bar{\Omega}|$, $\cap \bar{\Omega} = \bigcap_{j=1}^{J} \Omega^j \neq \emptyset$.
- *No* $O \in \bar{\Omega}$ *is such that* $|O| = 1$.

[6] Of course, in a qualified voting function U is always a decisive set.

PROOF

($\Rightarrow$)

We will begin our proof noticing that $|\bar{\Omega}| \leq 2^n$. *That is, it includes only a finite number of decisive sets. Then, by Pareto, the grand coalition* $\{1, \ldots, n\}$ *must be decisive. On the other hand, by Positive Responsiveness, if a set O is decisive and* $O \subseteq O'$, *if the attack relations in* $O' \setminus O$ *agree with those in O, the result will be the same, and therefore* O' *becomes decisive too.*

By Independence of Irrelevant Alternatives, if over a pair of arguments A, B, *the attack relations remain the same then the aggregate attack relation will be the same over* A, B. *We will prove that this implies that* $\cap\bar{\Omega} \neq \emptyset$. *First, consider the case where at least two decisive sets* $O, W \in \bar{\Omega}$ *are such that* $O \cap W = \emptyset$. *Suppose furthermore that O determines* $\rightarrow_{\mathcal{F}}$ *up from* $\{\rightarrow_i\}_{i=1}^n$ *while W defines* $\rightarrow_{\mathcal{F}}{}'$ *up from* $\{\rightarrow_i'\}_{i=1}^n$. *Then, if over* A, B $\rightarrow_i = \rightarrow_i'$ *then* $\rightarrow_{\mathcal{F}} = \rightarrow_{\mathcal{F}}{}'$ *over* A, B. *But then, since there is no element common to O and W, the choice over* A, B *will differ from* $\rightarrow_{\mathcal{F}}$ *to* $\rightarrow_{\mathcal{F}}{}'$. *Contradiction. Furthermore, if* $\cap\bar{\Omega} = \emptyset$, *then there is no* $\bar{i}$ *such that* $\rightarrow_{\mathcal{F}} \subseteq \rightarrow_{\bar{i}}$. *But then,* $\rightarrow_{\mathcal{F}}$ *includes other attacks than those in each individual attack relation. Without loss of generality, consider an extended argument framework over* n *arguments and a profile in which the attack relations over them is such that each of them constitutes a linear chain of attacks:*

- $A^1 \rightarrow_1 A^2 \ldots \rightarrow_1 A^n$,
- $A^2 \rightarrow_2 \ldots A^n \rightarrow_2 A^1$,
- $\ldots$,
- $A^n \rightarrow_n A^1 \ldots \rightarrow_n A^{n-1}$.

Then, over each pair A^j, A^k, $\rightarrow_{\mathcal{F}}$ *has to coincide with some of the individual attack relations. In particular for each pair of arguments* A^j, A^{j+1}. *But also on* A^n, A^1. *But then,* $\rightarrow_{\mathcal{F}}$ *yields a cycle (see Example 2):* $A^1 \rightarrow_{\mathcal{F}} A^2 \ldots \rightarrow_{\mathcal{F}} A^n \rightarrow_{\mathcal{F}} A^1$ *But this contradicts the assumption that* $\rightarrow_{\mathcal{F}}$ *is acyclic. Then,* $\cap\bar{\Omega} \neq \emptyset$.

Finally, a dictator i_0 *is such that* $\{i_0\} \in \bar{\Omega}$. *Therefore, non-dictatorship implies that there is no* $O \in \bar{\Omega}$ *such that* $|O| = 1$.

($\Leftarrow$)

The Pareto condition *follows from the fact that* $\{1, \ldots, n\} \in \bar{\Omega}$. *That is, if for a given pair* $A, B \in AR$, $A \rightarrow_i B$ *for every* $i = 1, \ldots, n$, *since* $\{1, \ldots, n\}$ *is decisive, it follows that* $A \rightarrow_{\mathcal{F}} B$.

Positive Responsiveness *follows from the fact that if O is decisive and* $O \subseteq O'$, O' *is also decisive. This is so since, given any* $A, B \in AR$, *and two* n*-tuples of attack relations,* $(\rightarrow_1, \ldots, \rightarrow_n)$, $(\rightarrow_1', \ldots, \rightarrow_n')$, *if* $\{i : A \rightarrow_i B\} \subseteq \{i : A \rightarrow_i' B\}$ *and* $A \rightarrow_{\mathcal{F}} B$, *then* $\{i : A \rightarrow_i B\}$ *is decisive, and therefore* $\{i : A \rightarrow_i' B\}$ *is also decisive, and then* $A \rightarrow_{\mathcal{F}}' B$.

Independence of Irrelevant Alternatives *obtains from the fact that* $\cap\bar{\Omega} \neq \emptyset$. *Suppose this were not the case. That is, there exists a pair* $A, B \in AR$, *and two* n*-tuples of attack relations,* $(\rightarrow_1, \ldots, \rightarrow_n)$, $(\rightarrow_1', \ldots, \rightarrow_n')$, *such that* $\rightarrow_i = \rightarrow_i'$ *over* (A, B), *but* $\rightarrow_{\mathcal{F}} \neq \rightarrow_{\mathcal{F}}'$ *over* (A, B). *Consider* $\bar{i} \in \cap\bar{\Omega} \neq \emptyset$. *That is* $\bar{i}$ *belongs to every decisive set. Then if, without loss of generality,* $A \rightarrow_{\bar{i}} B$ *then* $A \rightarrow_{\mathcal{F}} B$, *but also, since* $A \rightarrow_{\bar{i}}' B$, *we have that* $A \rightarrow_{\mathcal{F}}' B$. *Contradiction.*

Non-dictatorship *follows from the fact that no set with a single criterion is decisive and therefore, no single attack relation can be imposed over the aggregate for every profile of attack relations.*

Finally, notice that since there exists $\bar{i} \in \cap\bar{\Omega}$ *over each pair of arguments* A, B, $\rightarrow_{\mathcal{F}}$ *either coincides with* $\rightarrow_{\bar{i}}$ *or* $(A, B) \notin \rightarrow_{\mathcal{F}}$. *Since* $\rightarrow_{\bar{i}}$ *has no cycles of attacks,* $\rightarrow_{\mathcal{F}}$ *will also be acyclic.* □

When $\bar{\Omega}$ satisfies the properties described in Proposition 3, we say that $\bar{\Omega}$ is a *proper prefilter* over $\{1, \ldots, n\}$ [7]. Moreover, if the class of decision sets for an aggregation function has this structure, it aggregates acyclic attack relations into an acyclic relation, verifying Arrow's conditions.

Example 5. Over $\{\rightarrow_1, \rightarrow_2, \rightarrow_3\}$ (or $\{1, 2, 3\}$, for short), the only possible proper prefilters are:

- $\bar{\Omega}^I = \{\{1, 2\}, \{1, 2, 3\}\}$.
- $\bar{\Omega}^{II} = \{\{1, 3\}, \{1, 2, 3\}\}$.
- $\bar{\Omega}^{III} = \{\{2, 3\}, \{1, 2, 3\}\}$.
- $\bar{\Omega}^{IV} = \{\{1, 2\}, \{2, 3\}, \{1, 2, 3\}\}$.
- $\bar{\Omega}^{V} = \{\{1, 3\}, \{2, 3\}, \{1, 2, 3\}\}$.
- $\bar{\Omega}^{VI} = \{\{1, 2\}, \{1, 3\}, \{1, 2, 3\}\}$.

Notice that the corresponding aggregation functions $\mathcal{F}^I$, $\mathcal{F}^{II}$ and $\mathcal{F}^{III}$ are qualified voting functions. To see how the other three functions act, just consider $\mathcal{F}^{IV}$ over $A \rightarrow_1 B \rightarrow_1 C$, $A \rightarrow_2 C$, $B \rightarrow_2 C$, and $A \rightarrow_3 C \rightarrow_3 B$. Then, $\rightarrow_{\mathcal{F}} = \mathcal{F}^{IV}(\rightarrow_1, \rightarrow_2, \rightarrow_3)$ is defined as follows:

- $A \rightarrow_{\mathcal{F}} C$ since while there is no agreement in $\{1, 2\}$, $\{2, 3\}$ agree in that A attacks C.
- $B \rightarrow_{\mathcal{F}} C$ since B attacks C in $\rightarrow_1$ and $\rightarrow_2$, but there is no agreement in $\{2, 3\}$.
- $(A, B) \notin \rightarrow_{\mathcal{F}}$ since $(A, B) \notin \rightarrow_2$ and $(A, B) \notin \rightarrow_{\mathcal{F}\,3}$, but there is no agreement in $\{1, 2\}$.

That means that $\mathcal{F}^{IV}$ behaves in the same manner as a majority function over the following profile: $(\rightarrow_1, \rightarrow_2, \rightarrow_3)$. The same conclusion can be drawn for $\mathcal{F}^V$ and $\mathcal{F}^{VI}$.

Notice that, $\mathcal{F}^{IV}$ is actually a majority function only in the case that $\rightarrow_{\mathcal{F}}$ is acyclic. That is, it behaves like the majority function in well-behaved cases. Instead, for the individual attack relations in Example 2 it yields an acyclic order $A \rightarrow_{\mathcal{F}} C \rightarrow_{\mathcal{F}} B$, which is *not* the outcome of the majority function. Therefore, we should actually say that $\mathcal{F}^{IV}$, $\mathcal{F}^V$ and $\mathcal{F}^{VI}$ are *acyclic majority* functions. Notice that any $\bar{i} \in \cap\bar{\Omega}$ is kind of a "hidden dictator", in the sense made precise in the following result:

Proposition 4
If $\bar{i} \in \cap\bar{\Omega}$, *and* $\mathcal{F}_{\bar{\Omega}}$ *is the aggregation function characterized by the prefilter then* $\bar{\Omega}$, $\rightarrow_{\mathcal{F}_{\bar{\Omega}}} \subseteq \rightarrow_{\bar{i}}$.

PROOF

Suppose that given $A, B \in AR$, we have, without loss of generality, that $A \rightarrow_{\bar{i}} B$. Let us consider two cases:

- *There exists a decisive set $O \in \bar{\Omega}$ such that for every $i \in O$ (by definition $\bar{i} \in O$), $A \rightarrow_i B$. Then $A \rightarrow_{\mathcal{F}_{\bar{\Omega}}} B$, and therefore $\rightarrow_{\mathcal{F}_{\bar{\Omega}}}$ coincides with $\rightarrow_{\bar{i}}$ over (A, B).*
- *There does not exist any decisive set O in which for every $i \in O$, $A \rightarrow_i B$. Then, neither $A \rightarrow_{\mathcal{F}_{\bar{\Omega}}} B$ nor $B \rightarrow_{\mathcal{F}_{\bar{\Omega}}} A$ can obtain. Therefore $\bar{i}$* vetoes *$B \rightarrow_i A$, although it cannot imposes $A \rightarrow B$. In this case $\rightarrow_{\mathcal{F}_{\bar{\Omega}}} \subset \rightarrow_{\bar{i}}$ over (A, B).*

□

We will consider now the question of *Aggregation and Cycles of Attack*. The analysis of argumentation systems is usually carried out in terms of their *extensions*. The existence and properties of the extensions can be ascertained according to the properties of the attack relation. In the case that several alternative attack relations compete over the same class of arguments, the class of extensions may vary from one to another. The structure of extensions of such an argumentation system should not be seen as just the enumeration of the classes corresponding to each attack relation but should arise from the same aggregation process we have discussed previously. That is, it should follow from the properties of the aggregate attack relation.

In particular, since our main results concern the aggregation of acyclic attack relations into an acyclic aggregate one, we will focus on the case of well-founded argumentation frameworks (*cf.* [3], p. 10). We can say, roughly, that their main feature is the absence of cycles of attack among their arguments. For them, all the types of extensions described by Dung coincide. Furthermore, they all yield a single set of arguments (*cf.* [3], theorem 30, p. 331). To see how such a single extension of an argument system over a family of individual attack relations may obtain, let us recall that if for each i, $\rightarrow_i$ has no cycles of attack, an aggregate relation $\rightarrow_{\mathcal{F}_{\bar{\Omega}}}$, obtained through an aggregation function F with a prefilter of decisive sets $\bar{\Omega}$, is acyclic as well. The following result is an immediate consequence of this claim.

Proposition 5

Consider an aggregate argument framework $AF^* = \langle AR;\ \mathcal{F}(\rightarrow_1, \ldots, \rightarrow_n)\rangle$. If each $\rightarrow_i$ $(i = 1, \ldots, n)$ is acyclic and $\mathcal{F}$ is such that its corresponding class of decisive sets $\bar{\Omega}$ is a prefilter, then $\rightarrow_{\mathcal{F}} = \mathcal{F}(\rightarrow_1, \ldots, \rightarrow_n)$ is acyclic and AF^* has a single extension which is grounded, preferred and stable.

Furthermore:

Corollary 1. *If $AF^* = \langle AR;\ \mathcal{F}(\rightarrow_1, \ldots, \rightarrow_n)\rangle$ has a single extension when each $\rightarrow_i$ $(i = 1, \ldots, n)$ is acyclic, then if $\mathcal{F}$ is such that its corresponding class of decisive sets $\bar{\Omega}$ is a prefilter, it also verifies the Pareto condition, Positive Responsiveness, Independence of Irrelevant Alternatives, and Non-Dictatorship.*

PROOF

Immediate. If AF^* has a single extension and $\mathcal{F}$ is such that its corresponding class of decisive sets $\bar{\Omega}$ is a prefilter, then by Proposition 5 the aggregate attack $\rightarrow_{\mathcal{F}}$ is acyclic. Then, the claim follows from Proposition 3. □

5 Discussion

As indicated by Brown in [11], the fact that the class of decisive sets constitutes a prefilter is an indication of the existence of a *collegium*. In SCT this means a kind of "shadow" decisive set, being its members interspersed among all the actual decisive sets. Their actual power comes not from being able to enforce outcomes but from their ability to veto alternatives that are not desirable for them. In the current application, the existence of a collegium means that there exists a class of attack relations that by themselves cannot determine the resulting attack relation, but can instead block (veto) alternatives.

This feature still leaves many possibilities open, but as our examples intended to show, there are few aggregation functions that may adopt this form, while at the same time verifying the conditions postulated by Arrow for fair aggregation functions. The main instance is constituted by the acyclic majority function, but qualified voting functions yield also fair outcomes. The difference is that with the acyclic majority functions one of the several attack relations in AF^n is selected, while with qualified voting functions, new attack relations may arise. But these new attack relations just combine those of the winning coalitions, and therefore can be seen as resulting from the application of generalized variants of the majority function. That is, if two or more rules belong to all the decisive sets, their common fragments plus the non-conflicting ones add up to constitute the aggregate attack relation. In a way or another, the attack relations that are always decisive end up acting as hidden dictators in the aggregation process.

Explicit dictators arise in aggregation processes in other branches of non-monotonic reasoning. Doyle and Wellman [12], in particular, suggested to translate Reiter's defaults into total preorders of autoepistemic formulas, representing preferences over worlds. Reasoning with different defaults implies to find, first, the aggregation of the different preorders. These authors show that it is an immediate consequence of Arrow's theorem that no aggregation function can fulfill all the properties that characterize fairness (*i.e.*, the equivalents of the Pareto condition, Positive Responsiveness, Independence of Irrelevant Alternatives, and Non-Dictatorship). In terms of decisive sets, it means that $\hat{\Omega}$ constitutes a *principal ultrafilter*.[7] Since the number of defaults is assumed to be finite, it follows that there exists one of these default rules, say R^* that belongs to each $U \in \hat{\Omega}$. Of course, the existence of R^* violates the Non-Dictatorship condition, and consequently the actual class of formulas that arise in the aggregation are determined by R^*.

It can be said that Doyle and Wellman's analysis is concerned with the *generation* of a class of arguments arising from different default rules while we, instead, concentrate on the *comparison* among arguments in an abstract argumentation framework. But Dung [3] has shown that Reiter's system can be rewritten as an argumentation framework, and therefore both approaches can be made compatible. In this sense, Doyle and Wellman's result can be now interpreted as indicating that $\bar{\Omega}$ (over attack relations) is **not** a *proper prefilter* over $\{1, \ldots, n\}$ (where these indexes range over the attack relations determined each by a corresponding default rule). Instead, as said, it constitutes a

[7] Notice that $\hat{\Omega}$ denotes the decisive set over *default rules* and therefore should not be confused with $\bar{\Omega}$, the class of decisive sets over *attack relations*.

principal ultrafilter and therefore it implies the existence of a "dictatorial" attack relation, that is imposed over the framework.

Finally, the approach most related to ours is Coste-Marquis et al. [4], which already presented some early ideas on how to merge Dung's argumentation frameworks. The authors' aim is to find a set of arguments collectively warranted, addressing two main problems. One is the individual problem faced by each agent while considering a set of arguments different to that of other agents. The other is the aggregate problem of getting the collectively supported extension. The second one is the most clearly related to our approach, but differs in that the authors postulate a specific way of merging the individual frameworks.

The approach is based on a notion of *distance* between partial argumentation frameworks (PAFs, each one representing one agent's criterion) over a common set of arguments A. Each partial argumentation framework is defined by three binary relations R, I and N over A: R is the attack relation sanctioned by the agent, I includes the pairs of arguments about which the agent cannot establish any attacks, and $N = (A \times A) \setminus R \cup I$. A pseudo-distance d between PAFs over A is a mapping that associates a real number to each pair of PAFs over A and satisfies the properties of symmetry $(d(x, y) = d(y, x))$ and minimality $(d(x, y) = 0$ iff $x = y)$. d is a distance if it satisfies also the triangular inequality $(d(x, y) \leq d(x, y) + d(y, z))$. These mappings give a way of measuring how "close" is a collective framework from a given profile.

In a further step the authors define an aggregation function as a mapping from $(R+)^n$ to $(R+)$ that satisfies non-decreasingness (if $x_i \geq x'_i$), then $\otimes(x_1, \ldots, x_i, \ldots, x_n) \geq \otimes(x_1, \ldots, x'_i, \ldots, x_n)$, minimality ($\otimes(x_1, \ldots, x_n) = 0$ if $\forall i \; x_i = 0$), and identity ($\otimes(x) = x$). The idea is that merging a profile of AFs is a two-step process: first, to compute an *expansion* of each AF_i over the profile; and second, a *fusion* in which the AFs over A that are selected as result of the merging are the ones that are the "closest" to the profile. We are currently trying to establish a formal relation between this approach and our results about decisive sets of agents, in particular to determine whether Coste-Marquis et al.'s aggregation procedure satisfies the Arrovian properties.

6 Further Work

A relevant question arises from our analysis of the semantics of aggregate frameworks when the attack relations are acyclic. Namely, whether there exist a sensible notion of aggregation of extensions that could correspond to the aggregation of attack criteria. In this paper we have focused on the path that goes from several attack relations to a single aggregate one and from it to its corresponding extension.

An open question is whether it is possible to go through the alternative path from several attack relations to their corresponding extensions and from there on to a single family of extensions. If, furthermore, these two alternative paths commute (in category-theoretic terms), the aggregation of attack relations would be preferable in applications since it is simpler to aggregate orderings than families of sets.

Nevertheless, there are reasons to be pessimistic, due to the similarities between this problem and the aggregation of judgments for which List and Pettit have found negative results [13]. They consider the so-called "discursive dilemma" in which several judgments (*i.e.*, pairs of the form ⟨*premises*, *conclusion*⟩), are aggregated componentwise, that is, a pair formed by an aggregated premises set and an aggregated conclusion is obtained, but it does not constitute an acceptable judgment. It could happen that similar problems may arise while trying to match aggregated attack criteria and aggregated extensions.

Pigozzi [14] postulates a solution to the discursive dilemma based on the use of operators for merging belief bases in AI [15]. To pose a *merge* operation as an *aggregation* one involves to incorporate a series of trade-offs among the several alternatives that hardly will respect Arrow's conditions, as it is well known in the literature on political systems (see [16]). In relation to this issue, a wider point that we plan to address is to systematize the *non*-fair aggregation procedures that could be applied to the aggregation of attack relations in argument frameworks. The idea would be to lesser the demands on the aggregation function and to see which features arise in the aggregate. It seems sensible to think that depending on the goals of the aggregation process, one or another function should be chosen.

Another question to investigate is the connections between the correspondence of the aggregate attack relation among arguments and the relation of *dominance* among alternatives ([17]). From a SCT view, alternative A *dominates* alternative B iff the number of individuals for which A is preferred to B is larger than the number of individuals for which B is preferred to A. This implies that the dominance relation is asymmetric. Although it is not commonly assumed in the literature that attack relations are asymmetric, it follows from our definition of majority voting over pairs of arguments that the resulting attack relations will have this property (even when cycles of order > 2 may occur). Dominance relations lead to the choice of stable sets. A stable set is such that none of its elements dominate another, and every alternative outside the set is dominated by some of its elements. The correspondence between stable semantics in argumentation frameworks and stable sets was previously studied by Dung ([3]). It is natural, so, to inquire about the relationship between our majoritarian voting aggregation mechanism on attack relations and stable sets in argumentation frameworks.

References

1. Chesñevar, C., Maguitman, A., Loui, R.: Logical Models of Argument. ACM Computing Surveys 32, 337–383 (2000)
2. Prakken, H., Vreeswijk, G.: Logical systems for defeasible argumentation. In: Gabbay, D. (ed.) Handbook of Philosophical Logic, 2nd edn. Kluwer Academic Pub., Dordrecht (2002)
3. Dung, P.M.: On the acceptability of arguments and its fundamental role in nonmonotonic reasoning, logic programming and n-person games. Artificial Intelligence 77, 321–358 (1995)
4. Coste-Marquis, S., Devred, C., Konieczny, S., Lagasquie-Schiex, M.C., Marquis, P.: Merging argumentation systems. In: Veloso, M.M., Kambhampati, S. (eds.) Proc. of the 20th National Conference on Artificial Intelligence and the 17th Innovative Applications of Artificial Intelligence Conference, pp. 614–619. AAAI Press, USA (2005)
5. Moulin, H.: Social choice. In: Aumann, R., Hart, S. (eds.) Handbook of Game Theory, vol. 2. North-Holland, Amsterdam (1994)
6. Black, D.: The Theory of Committees and Elections. Cambridge University Press, Cambridge (1958)

7. Brown, D.J.: Aggregation of preferences. Quarterly Journal of Economics 89, 456–469 (1975)
8. Blair, D., Pollack, R.: Acyclic collective choice rules. Econometrica 50, 931–943 (1982)
9. Arrow, K.J.: Social Choice and Individual Values, 2nd edn. Yale University Press, London (1970)
10. Sen, A.: Maximization and the act of choice. Econometrica 65, 745–779 (1997)
11. Brown, D.J.: An approximate solution to arrow's problem. Journal of Economic Theory 9, 375–383 (1974)
12. Doyle, J., Wellman, M.: Impediments to universal preference-based default theories. Artificial Intelligence 49, 97–128 (1991)
13. List, C., Pettit, P.: Aggregating sets of judgments. Two impossibility results compared. Synthese 140, 207–235
14. Pigozzi, G.: Belief merging and the discursive dilemma: An argument-based account to paradoxes of judgment aggregation. Synthese 152, 285–298 (2006)
15. Konieczny, S., Pino-Pérez: Propositional belief base merging or how to merge beliefs/goals coming from several sources and some links with social choice theory. European Journal of Operational Research 160, 785–802 (2005)
16. Austen-Smith, D., Banks, J.S.: Positive Political Theory I: Collective Preference. In: Michigan Studies in Political Analysis. University of Michigan Press, Ann Arbor, Michigan (2000)
17. Brandt, F., Fischer, F., Harrenstein, P.: The Computational Complexity of Choice Sets. In: Samet, D. (ed.) TARK. Proceedings of the 11th Conference on Theoretical Aspects of Rationality and Knowledge, pp. 82–91. Presses Universitaires de Louvain (2007)