Preservation of Semantic Properties in Collective Argumentation: The Case of Aggregating Abstract Argumentation Frameworks

Chen, W.; Endriss, U.

DOI
10.1016/j.artint.2018.10.003

Publication date
2019

Document Version
Final published version

Published in
Artificial Intelligence

Citation for published version (APA):
Chen, W., & Endriss, U. (2019). Preservation of Semantic Properties in Collective Argumentation: The Case of Aggregating Abstract Argumentation Frameworks. *Artificial Intelligence*, *269*, 27-48. https://doi.org/10.1016/j.artint.2018.10.003

Artificial Intelligence 269 (2019) 27–48

Contents lists available at ScienceDirect

Artificial Intelligence

www.elsevier.com/locate/artint

Preservation of semantic properties in collective argumentation: The case of aggregating abstract argumentation frameworks ☆

Weiwei Chen [a,b,*], Ulle Endriss [b]

[a] *Institute of Logic and Cognition and Department of Philosophy, Sun Yat-sen University, China*
[b] *ILLC, University of Amsterdam, the Netherlands*

ARTICLE INFO

Article history:
Received 12 December 2017
Received in revised form 20 July 2018
Accepted 10 October 2018
Available online 21 December 2018

Keywords:
Argumentation theory
Social choice theory

ABSTRACT

An abstract argumentation framework can be used to model the argumentative stance of an agent at a high level of abstraction, by indicating for every pair of arguments that is being considered in a debate whether the first attacks the second. When modelling a group of agents engaged in such a debate, we may wish to aggregate their individual argumentation frameworks to obtain a single such framework that reflects the consensus of the group. While agents typically will not agree on every single attack, there may well be high-level agreement on semantic properties, such as whether a given argument should be accepted or whether there are any acceptable arguments at all. Using techniques from social choice theory, we analyse the circumstances under which such semantic properties agreed upon by the individual agents will be preserved under aggregation. Our results cover semantic properties formulated in terms of six of the most widely used extension-based semantics for abstract argumentation and range from positive results that show that certain aggregation rules can provide the desired preservation guarantees to impossibility theorems that show that certain combinations of requirements cannot be met by any reasonable aggregation rule.

1. Introduction

Formal argumentation theory provides tools for modelling both the arguments an agent may wish to employ in a debate and the relationships that hold between such arguments [2–6]. This applies both to human agents and to intelligent software agents. In the widely used model of *abstract argumentation*, introduced in the seminal work of Dung [7], we abstract away from the internal structure of arguments and only model whether or not one argument *attacks* another argument. Thus, arguments are vertices in a directed graph, with edges representing attacks. This is a useful perspective when we require a high-level understanding of how different arguments relate to each other. When several agents engage in a debate, they may differ on their assessment of some of the arguments and their relationships. How best to model such scenarios of

☆ This is an extended version of a paper first presented at TARK-2017 [1]. We would like to thank Sirin Botan, Umberto Grandi, Ronald de Haan, Zoi Terzopoulou, and Shengyang Zhong for numerous enlightening discussions on this material and the anonymous TARK and *Artificial Intelligence* reviewers for their constructive feedback.

* Corresponding author.
E-mail address: chenww26@mail2.sysu.edu.cn (W. Chen).

https://doi.org/10.1016/j.artint.2018.10.003
0004-3702/

collective argumentation is a question of considerable interest, not only in AI. Over the past decade or so, several authors have started to contribute to its resolution [see, e.g., 8–17].

Specifically, when agents differ on their assessment of which attacks between the arguments are in fact justified, i.e., when they put forward different *attack-relations*, we may wish to *aggregate* these individual pieces of information to obtain a global view. In this paper we analyse the circumstances under which a given *aggregation rule* will *preserve* relevant properties of the individual attack-relations, particularly properties that relate to the various *semantics* that have been proposed for abstract argumentation. For example, if all agents agree that argument A is *acceptable*, either because it is not attacked by any other argument or because it can be successfully defended against any such attack, then we would like A to also be considered acceptable relative to the attack-relation returned by our aggregation rule. Thus, *argument acceptability* is an example for a property that, ideally, should be preserved under aggregation. Our objective is to analyse what kind of simple aggregation rules can guarantee that this will be the case. Our approach is grounded in *social choice theory*, the formal study of collective decision making [18–20]. In particular, we apply the so-called *axiomatic method* [21–23] and make use of recent results on *graph aggregation* [24].

Related work. Coste-Marquis et al. [8] were the first to address the question of how best to aggregate several abstract argumentation frameworks, but without making explicit reference to social choice theory. Instead, they focus on a family of sophisticated aggregation rules that minimise the *distance* between the input argumentation frameworks and the output argumentation framework.

Tohmé et al. [9] were the first to explicitly use social choice theory to analyse the aggregation of argumentation frameworks. Their focus is on the preservation of the *acyclicity* of attack-relations under aggregation. Acyclicity is an important property in the context of abstract argumentation, because it greatly simplifies the evaluation of arguments (in an acyclic argumentation framework, it is unambiguous which arguments to accept and which to reject). Tohmé et al. show that *qualified majority rules* will always preserve this property.[1]

Bodanza and Auday [25] were the first to give a completely general definition of an aggregation rule mapping any set of individual argumentation frameworks into a collective argumentation framework. In contrast to this, all earlier authors restrict attention to specific aggregation rules or specific classes of aggregation rules. Bodanza and Auday compare two different scenarios of collective argumentation that both combine abstract argumentation with social choice theory. In the first scenario, we assume that every agent reports an argumentation framework and we need to find a good way of aggregating this input into a single collective argumentation framework. This is the scenario we study in this paper. In the other scenario, every agent is presented with the same argumentation framework but reports a (possibly) different set of arguments she considers acceptable given that argumentation framework, i.e., every agent reports her own *extension* of the common argumentation framework. This latter scenario—which is equally interesting but technically very different from the one we investigate here—has later been studied in more depth by a number of authors, including Caminada and Pigozzi [10] and Rahwan and Tohmé [11], as well as the present authors [17]. Other authors, such as Gabbay and Rodrigues [26] and Delobelle et al. [27], work with models that are hybrids between these two scenarios. In these hybrid models each agent comes equipped an individual argumentation framework, but we also have access to some additional information to exploit during aggregation. For instance, for Delobelle et al. this additional information specifies the argumentation semantics adopted by all agents, which allows the mechanism performing the aggregation to compute extensions for all agents, which can then feed into the aggregation process.

While Tohmé et al. [9] had made the important step of introducing the methodology of social choice theory into the study of collective argumentation, their work in fact was largely a study *within* social choice theory that made little reference to the specifics of the domain of argumentation, the only exception being the focus on the property of acyclicity. A few years later, Dunne et al. [13] succeeded in bringing abstract argumentation and social choice theory closer together by defining several preservation requirements on aggregation rules that directly refer to the semantics of the argumentation frameworks concerned. This includes the requirement that an argument that is acceptable to all individual agents should also be acceptable in the argumentation framework returned by the aggregation rule ("credulous σ-acceptance unanimity") and the requirement specifying that, when all agents agree on what the acceptable arguments are, then this agreement should be preserved under aggregation ("σ-unanimity").[2] The focus of their technical contribution, however, is on analysing the computational complexity of deciding whether a given aggregation rule has a given property, rather than on the axiomatic method. In follow-up work, Delobelle et al. [14] establish for several concrete rules whether or not they satisfy the preservation requirements introduced by Dunne et al. [13].

In further related work, employing a similar model and making more explicit use of the axiomatic method of social choice theory to analyse scenarios of collective argumentation, Li [28] discusses a variant of Sen's famous Paradox of the Paretian Liberal [29] in the context of aggregating abstract argumentation frameworks. He shows that granting some agents "expert rights", i.e., the right to autonomously decide on certain attack relations that are within the scope of their special

[1] A qualified majority rule accepts a given attack between two arguments if (i) a majority of the agents accept that attack and (ii) none of a subset of the agents with a special status (the right to "veto") reject it.

[2] Dunne et al. [13] refer to these requirements as "axioms", while we prefer to distinguish axioms (normative properties of aggregation rules that typically encode some notion of fairness) from the "collective rationality" requirement that certain properties of argumentation frameworks should be preserved under aggregation.

expertise, is incompatible with basic efficiency requirements when we require aggregation rules that always return an argumentation framework that has at least one stable extension.

While Endriss and Grandi [24] explicitly mention abstract argumentation as a possible domain of application for the model of graph aggregation they develop, they do not present any technical results related to argumentation.

Let us also briefly mention a small number of contributions a little further afield. Bonzon and Maudet [30] define an *interaction protocol* for agents intent on persuading each other about the acceptability of a given argument and compare the outcomes of such persuasion dialogues with the results obtained by applying an aggregation rule to the argumentation frameworks initially held by these agents. Leite and Martins [12] introduce *social abstract argumentation frameworks*, which are abstract argumentation frameworks enriched with a function mapping each argument to the numbers of agents voting in favour and against it. In particular the enriched model proposed by Eğilmez et al. [31], which also allows for votes in favour of and against attacks, bears some relation to the scenario we study here. For instance, if an aggregation rule aggregates individual positions separately on each individual attack (i.e., if it satisfies the axiom of *independence* to be introduced in the sequel), then we can use such an enriched social abstract argumentation framework as an intermediate form of representation. Airiau et al. [15] introduce the concept of the *rationalisability* of a profile of argumentation frameworks. A profile is rationalisable if the diversity of views it contains can be explained in terms of (i) an underlying factual argumentation framework shared by all agents and (ii) everyone's individual preferences. Thus, their work is concerned with understanding what kind of profiles a good aggregation rule should be able to deal with, rather than with aggregation itself.

For a detailed review of research on collective argumentation beyond this small selection we refer to the recent survey by Bodanza et al. [16].

Contribution. Our first contribution is the formulation of a clear and simple model for the axiomatic study of the preservation of semantic properties during the aggregation of attack-relations over a common set of arguments. Our technical results delineate how fundamental axiomatic properties of aggregation rules interact with such preservation requirements. These results range from characterisation theorems that indicate what kind of aggregation rule can satisfy certain combinations of desiderata, to impossibility theorems that show that only aggregation rules that are clearly unacceptable from an axiomatic point of view (namely, so-called dictatorships) can preserve the most demanding semantic properties.

In terms of methodology, we show how techniques originally developed for the more general domain of graph aggregation [24] can be applied in the context of abstract argumentation. At the same time, we identify a number of properties of graphs that are of interest in the specific context of abstract argumentation that cannot be analysed with existing techniques and the novel technique we develop to handle these properties—which we call *k-exclusive properties*—likely will find future application in other domains where graphs need to be aggregated.

Finally, while we restrict attention to the aggregation of argumentative positions that can be modelled using Dung's system of abstract argumentation, the idea of collective argumentation is more general than that and we believe that our approach can, at least in principle, be extended to richer models of argumentation that also account for the internal structure of arguments. This is important, given that—despite the enormous popularity and widespread use of Dung's model in the literature on argumentation in AI and other disciplines [2,3,5,6]—there is broad consensus that the expressive power of this model is limited and can only account for certain high-level aspects of argumentation [32–34].

Paper overview. The remainder of this paper is organised as follows. Section 2 is a brief review of relevant concepts from the theory of abstract argumentation. Section 3 introduces our model and Section 4 presents our technical results on the preservation of semantic properties of argumentation frameworks under aggregation. We conclude in Section 5 with a brief summary of the insights obtained as well as suggestions for future work. This includes opportunities for applying our approach to richer models of argumentation than Dung's classical model of abstract argumentation as well as applying some of the methodological tools we develop, particularly regarding the analysis of k-exclusive graph properties, in other domains of aggregation—possibly unrelated to the study of argumentation.

2. Abstract argumentation

In this section we recall some of the fundamentals of the model of abstract argumentation as originally introduced by Dung [7]. An *argumentation framework* is a pair $AF = \langle Arg, \rightharpoonup \rangle$, where Arg is a finite set of *arguments* and $\rightharpoonup$ is an irreflexive binary relation on Arg.[3] If $A \rightharpoonup B$ holds for two arguments $A, B \in Arg$, we say that A *attacks* B. The internal structure of individual arguments and thus the reasons for why one argument attacks or does not attack another are explicitly left unspecified in this model of argumentation. This approach has both advantages and disadvantages, which have been discussed at length in the literature [see, e.g., 32–34].

Attacking and defending arguments. Let us introduce some further notation and terminology. We use the term *attack* to refer to an element $att \in Arg \times Arg$ of an *attack-relation* $(\rightharpoonup) \subseteq Arg \times Arg$. That is, these are the "individual arrows" in $\langle Arg, \rightharpoonup \rangle$.

[3] Neither the finiteness nor the irreflexivity assumption are crucial for our results, but they simplify exposition and clearly are natural for most applications.

For a set of arguments $\Delta \subseteq Arg$ and an argument $B \in Arg$, we say that Δ *attacks* B, denoted as $\Delta \rightharpoonup B$, if $A \rightharpoonup B$ holds for some argument $A \in \Delta$. We write $\Delta^+ = \{B \in Arg \mid \Delta \rightharpoonup B\}$ for the set of arguments attacked by Δ. We further say that Δ *defends* the argument $B \in Arg$, if $\Delta \rightharpoonup A$ holds for all arguments $A \in Arg$ such that $A \rightharpoonup B$. The *characteristic function* of AF is defined as the function $f_{AF} : 2^{Arg} \to 2^{Arg}$ that maps any given set of arguments $\Delta \subseteq Arg$ to the set of arguments defended by Δ:

$$f_{AF}(\Delta) = \{B \in Arg \mid \Delta \textit{ defends } B\}$$

Semantics. Given an argumentation framework AF, the question arises which arguments to accept. For example, we may not want to accept two arguments that attack each other. A *semantics* specifies which sets of arguments can be accepted together for a given argumentation framework. Any such set of arguments is called an *extension* of AF under the semantics in question. For all the definitions of specific choices of semantics that follow, consider an arbitrary but fixed argumentation framework $AF = \langle Arg, \rightharpoonup \rangle$ and a set of arguments $\Delta \subseteq Arg$. The following notions of conflict-freeness and admissibility play a central role in these definitions. We say that Δ is *conflict-free*, if there exist no arguments $A, B \in \Delta$ such that $A \rightharpoonup B$; and Δ is called *admissible* if it is conflict-free and defends every single one of its members. We are going to work with six different types of semantics, which are amongst the most widely studied abstract argumentation semantics in the literature. The first four of them were introduced by Dung [7] in his original paper. The remaining two semantics were introduced more recently, by Caminada [35] and Dung et al. [36], respectively.[4]

Definition 1. A **stable extension** of AF is a conflict-free set Δ of arguments in Arg that attacks all other arguments $B \in Arg \setminus \Delta$, i.e., $\Delta \cup \Delta^+ = Arg$.

Definition 2. A **preferred extension** of AF is an admissible set of arguments in Arg that is maximal with respect to set inclusion.

Definition 3. A **complete extension** of AF is an admissible set of arguments in Arg that includes all of the arguments it defends.

Definition 4. A **grounded extension** of AF is a least fixed point of its characteristic function f_{AF}.

Definition 5. A **semi-stable extension** of AF is a complete extension Δ of AF for which $\Delta \cup \Delta^+$ is maximal with respect to set inclusion.

Definition 6. An **ideal extension** of AF is an admissible subset of the intersection of all preferred extensions that is maximal with respect to set inclusion.

The research community has produced several tools to automatically compute the extensions for a given argumentation framework under these semantics. One such tool is ConArg [38].[5]

Relationships between different semantics. Let us briefly recall some well-known facts about these different semantics and how they relate to each other [37]. The set of stable extensions may be empty for a given argumentation framework, while there always exists at least one extension under each of the other five semantics. Unlike for the other four semantics, there always are exactly one grounded extension and exactly one ideal extension. However, these extensions may be empty. We can compute the grounded extension Δ by initialising Δ with the empty set $\emptyset$ and then repeatedly executing the program $\Delta := f_{AF}(\Delta)$, until no more changes occur. Thus, the grounded extension is nonempty if and only if there is at least one argument that is not attacked at all. The ideal extension is always a (not necessarily proper) superset of the grounded extension. All stable extensions are also semi-stable extensions, and all semi-stable extensions are also preferred extensions. As the ideal extension is a subset of every preferred extension, it also is a subset of every semi-stable and every stable extension. The grounded, the ideal, and all preferred extensions are also complete extensions. Indeed, the grounded extension is the (unique) complete extension that is minimal with respect to set inclusion, while every preferred extension is a (not necessarily unique) complete extension that is maximal with respect to set inclusion. Furthermore, the grounded extension is a subset of every complete extension, and thereby also a subset of every preferred, every semi-stable, and every stable extension. Most of these relationships are summarised in Fig. 1. Finally, any extension under any of the six semantics considered here is admissible and thus also conflict-free.

---

[4] Note that what we call the *ideal extension* is called the *maximal ideal set* in the original paper by Dung et al. [36]. This change in terminology is in line with the more recent literature [37].

[5] As several of the proofs in this paper rely on constructions involving certain argumentation frameworks having certain extensions, to provide an additional level of verification of the correctness of these proofs, we have used ConArg to recompute the relevant extensions.

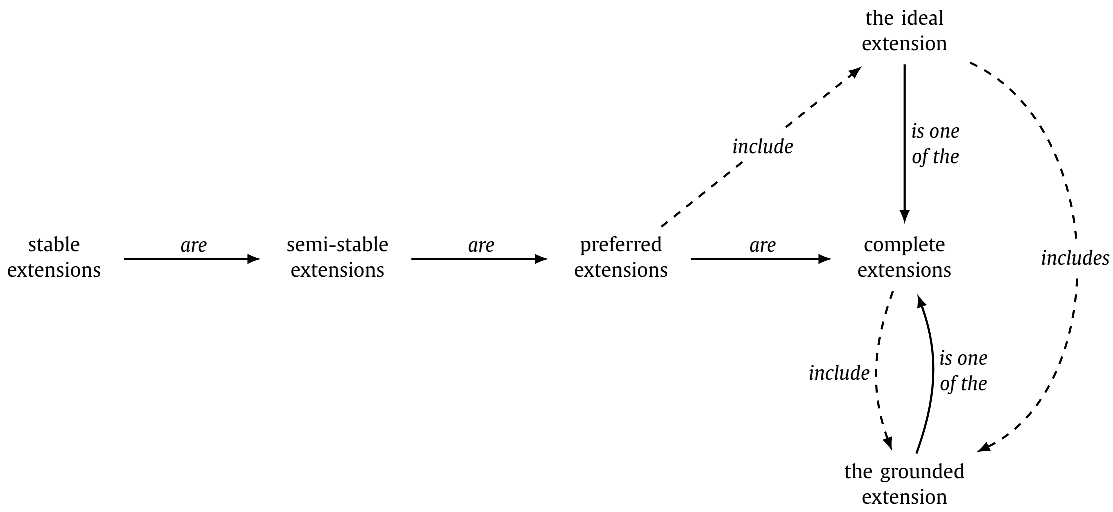

Fig. 1. Relationships between argumentation semantics.

Example 1. Consider an argumentation framework $AF = \langle Arg, \rightharpoonup \rangle$ with an isolated cycle of length 3. Thus, $(\rightharpoonup) \supseteq \{A \rightharpoonup B, B \rightharpoonup C, C \rightharpoonup A\}$ for three arguments $A, B, C \in Arg$ and there are no further attacks on either A, B, or C. Then none of A, B, and C can be part of a complete extension Δ: if we include two or more of them in Δ, then Δ is not conflict-free; if we include just one of them in Δ, then Δ does not defend itself. Hence, by the relationships between the different semantics shown in Fig. 1, none of A, B, and C can be part of an extension under any of the other five semantics either. We are going to make use of this kind of reasoning—which also works for longer isolated odd-length cycles—several times in this paper. △

An interesting question is under what circumstances the extensions determined by different semantics coincide. Probably the clearest example for when they do coincide is the case of argumentation frameworks with an *acyclic* attack-relation: if $\rightharpoonup$ does not include any cycles, then the grounded extension coincides with the ideal extension and it is the only stable extension, the only semi-stable extension, the only preferred extension, and the only complete extension. Indeed, for an acyclic $\rightharpoonup$ it is entirely uncontroversial which arguments to accept. A condition of this kind that is weaker than acyclicity is what is known as *coherence* in the literature [7]: the argumentation framework AF is called *coherent* if every preferred extension of AF is stable, i.e., if the two semantics coincide.

3. The model

Fix a finite set Arg of arguments and a set $N = \{1, \dots, n\}$ of n *agents*. Suppose each agent $i \in N$ supplies us with an argumentation framework $AF_i = \langle Arg, \rightharpoonup_i \rangle$, reflecting her individual views on the status of possible attacks between arguments. Thus, we are given a *profile* of attack-relations $\boldsymbol{\rightharpoonup} = (\rightharpoonup_1, \dots, \rightharpoonup_n)$.[6] What would be a good method of aggregating these individual argumentation frameworks to arrive at a single argumentation framework that appropriately reflects the views of the group as a whole? This is the central question we address in this paper. An *aggregation rule* is a function $F : (2^{Arg \times Arg})^n \to 2^{Arg \times Arg}$ mapping any given profile of attack-relations into a single attack-relation. Thus, we are interested in understanding what makes a good aggregation rule.

Example 2. The first aggregation rule that comes to mind is the majority rule: include attack $A \rightharpoonup B$ in the outcome if and only if a (weak) majority of the individual agents do. If we apply this rule to the profile shown in Fig. 2, then we obtain the argumentation framework consisting of the three attacks $A \rightharpoonup B$, $B \rightharpoonup C$, and $C \rightharpoonup A$. △

Section overview. In the remainder of this section we first define a number of *specific aggregation rules*. We then define several *properties of aggregation rules*, usually referred to as *axioms*, and briefly discuss which of the rules defined satisfy which of these properties. Finally, we review several *properties of argumentation frameworks*, particularly properties relating to their semantics, and formulate the question of whether a given rule will *preserve* such a property when all individual agents report argumentation frameworks that satisfy the property in question.

---

[6] Note that we assume that all agents report an attack-relation over *the same* set of arguments Arg. As argued by Coste-Marquis et al. [8], generalisations, where different agents may be aware of different subsets of Arg, are possible and interesting, but—in line with most existing work in the area—we shall not explore them here.

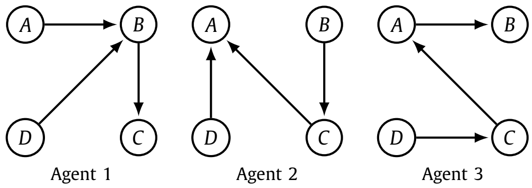

Fig. 2. Example for a profile with $Arg = \{A, B, C, D\}$.

3.1. Specific aggregation rules

Recall that an aggregation rule is a function F, mapping any given profile $\boldsymbol{\rightharpoonup} = (\rightharpoonup_1, \dots, \rightharpoonup_n) \in (2^{Arg \times Arg})^n$ of attack-relations on Arg to a single attack-relation $F(\boldsymbol{\rightharpoonup}) \subseteq Arg \times Arg$. We sometimes write $(A \rightharpoonup B) \in F(\boldsymbol{\rightharpoonup})$ for $(A, B) \in F(\boldsymbol{\rightharpoonup})$. We use $N^{\boldsymbol{\rightharpoonup}}_{att} := \{i \in N \mid att \in (\rightharpoonup_i)\}$ to denote the set of *supporters* of the attack att in profile $\boldsymbol{\rightharpoonup}$.

We now introduce two families of aggregation rules, the *quota rules* and the *oligarchic rules*. These are simple rules that are adaptations of well-known rules used in the social choice literature, particularly in judgment aggregation [39] and graph aggregation [24].

Definition 7. Let $q \in \{1, \dots, n\}$. The **quota rule** F_q with quota q accepts all those attacks that are supported by at least q agents:

$$F_q(\boldsymbol{\rightharpoonup}) = \left\{att \in Arg \times Arg \mid \#N^{\boldsymbol{\rightharpoonup}}_{att} \geqslant q\right\}$$

The *weak majority rule* is the quota rule F_q with $q = \lfloor \frac{n+1}{2} \rfloor$ and the *strict majority rule* is the quota rule F_q with $q = \lceil \frac{n+1}{2} \rceil$. Two further quota rules are also of special interest. The *unanimity rule* only accepts attacks that are supported by everyone, i.e., this is F_q with $q = n$. The *nomination rule* is the quota rule F_q with $q = 1$. Despite being a somewhat extreme choice, the nomination rule has some intuitive appeal in the context of argumentation, as it reflects the idea that we should take seriously any conflict between arguments raised by at least one member of the group.

Definition 8. Let $\mathfrak{C} \in 2^N \setminus \{\emptyset\}$ be a nonempty coalition of agents. The **oligarchic rule** $F_{\mathfrak{C}}$ accepts all those attacks that are accepted by all members of $\mathfrak{C}$:

$$F_{\mathfrak{C}}(\boldsymbol{\rightharpoonup}) = \left\{att \in Arg \times Arg \mid \mathfrak{C} \subseteq N^{\boldsymbol{\rightharpoonup}}_{att}\right\}$$

Thus, any member of the oligarchy $\mathfrak{C}$ can *veto* an attack from being accepted. Observe that the unanimity rule can also be characterised as the oligarchic rule $F_{\mathfrak{C}}$ with $\mathfrak{C} = N$. A subclass of the oligarchic rules are the *dictatorships*. The dictatorship of dictator $i \in N$ is the oligarchic rule $F_{\mathfrak{C}}$ with $\mathfrak{C} = \{i\}$. Thus, under a dictatorship, to compute the outcome, we simply copy the attack-relation of the dictator. Intuitively speaking, oligarchic rules, and dictatorships in particular, are unattractive rules, as they unfairly exclude everyone not in $\mathfrak{C}$ from the decision process.

Some rules combine features of the quota rules and the oligarchic rules. For example, we may choose to accept an attack only if it is accepted by (*i*) a weak majority of all agents and (*ii*) a small number of distinguished agents to which we want to give the right to veto the acceptance of attacks. Such rules (sometimes called *qualified majority rules*) are certainly more attractive than the oligarchic rules, but they are still unfair in the sense of granting some agents more influence than others.

Definition 9. Agent $i \in N$ has **veto powers** under aggregation rule F, if $F(\boldsymbol{\rightharpoonup}) \subseteq (\rightharpoonup_i)$ for every profile $\boldsymbol{\rightharpoonup}$.

Thus, under an oligarchic rule $F_{\mathfrak{C}}$ the agents in $\mathfrak{C}$, and only those, have veto powers. With the exception of the unanimity rule, under which all agents have veto powers, a quota rule does not grant veto powers to any agent.

3.2. Axioms: properties of aggregation rules

Next, we introduce several basic *axioms*, each of which encodes an intuitively desirable property of an aggregation rule F. All of these axioms are direct adaptations of axioms formulated in the literature on graph aggregation [24], which in turn are very similar to axioms used in both the literature on judgment aggregation [39] and that on preference aggregation [19].

Definition 10. An aggregation rule F is said to be **anonymous**, if $F(\boldsymbol{\rightharpoonup}) = F(\rightharpoonup_{\pi(1)}, \dots, \rightharpoonup_{\pi(n)})$ holds for all profiles $\boldsymbol{\rightharpoonup} = (\rightharpoonup_1, \dots, \rightharpoonup_n)$ and all permutations $\pi : N \to N$.

Definition 11. An aggregation rule F is said to be **neutral**, if $N^{\boldsymbol{\rightharpoonup}}_{att} = N^{\boldsymbol{\rightharpoonup}}_{att'}$ implies $att \in F(\boldsymbol{\rightharpoonup}) \Leftrightarrow att' \in F(\boldsymbol{\rightharpoonup})$ for all profiles $\boldsymbol{\rightharpoonup}$ and all attacks att, att'.

Definition 12. An aggregation rule F is said to be **independent**, if $N^{\rightharpoonup}_{att} = N^{\rightharpoonup'}_{att}$ implies $att \in F(\rightharpoonup) \Leftrightarrow att \in F(\rightharpoonup')$ for all profiles $\rightharpoonup$, $\rightharpoonup'$ and all attacks att.

Definition 13. An aggregation rule F is said to be **monotonic**, if $N^{\rightharpoonup}_{att} \subseteq N^{\rightharpoonup'}_{att}$ (together with $N^{\rightharpoonup}_{att'} = N^{\rightharpoonup'}_{att'}$ for all attacks $att' \neq att$) implies $att \in F(\rightharpoonup) \Rightarrow att \in F(\rightharpoonup')$ for all profiles $\rightharpoonup$, $\rightharpoonup'$ and all attacks att.

Definition 14. An aggregation rule F is said to be **unanimous**, if $F(\rightharpoonup) \supseteq (\rightharpoonup_1) \cap \cdots \cap (\rightharpoonup_n)$ holds for all profiles $\rightharpoonup = (\rightharpoonup_1, \ldots, \rightharpoonup_n)$.

Definition 15. An aggregation rule F is said to be **grounded**, if $F(\rightharpoonup) \subseteq (\rightharpoonup_1) \cup \cdots \cup (\rightharpoonup_n)$ holds for all profiles $\rightharpoonup = (\rightharpoonup_1, \ldots, \rightharpoonup_n)$.

Anonymity is a symmetry (and thus fairness) requirement regarding agents, and neutrality is a symmetry requirement regarding attacks. Independence expresses that whether an attack is accepted should only depend on its supporters. Monotonicity says that additional support for an accepted attack should never cause it to be rejected. Unanimity postulates that an attack supported by everyone must be accepted, while groundedness means that only attacks with at least one supporter can be collectively accepted.[7]

Observe that all quota rules and all oligarchic rules are easily seen to be unanimous, grounded, neutral, independent, and monotonic.[8] The quota rules furthermore are also anonymous. In fact, it is not difficult to adapt a well-known result from judgment aggregation due to Dietrich and List [40] to our setting, so as to see that the quota rules are *the only* aggregation rules that satisfy all of these six axioms (refer to Endriss and Grandi [24] for a formulation of this result in the context of graph aggregation).

Note that, if an aggregation rule F is independent, then we can represent it by listing for every potential attack $att = (A \rightharpoonup B)$ the coalitions of agents that would be sufficient to get that attack accepted if exactly the members of that coalitions were to support it. Formally, if F is independent, then for every attack $att \in Arg \times Arg$ there exist a family of coalitions $\mathcal{W}_{att} \subseteq 2^N$ such that for every profile $\rightharpoonup$ it is the case that $att \in F(\rightharpoonup)$ if and only if $N^{\rightharpoonup}_{att} \in \mathcal{W}_{att}$. The elements of $\mathcal{W}_{att}$ are called *winning coalitions*. If F is both independent and neutral, then the family of winning coalitions must be the same for all attacks, i.e., in that case there exists a single family $\mathcal{W} \subseteq 2^N$ such that for every profile $\rightharpoonup$ and every attack att it is the case that $att \in F(\rightharpoonup)$ if and only if $N^{\rightharpoonup}_{att} \in \mathcal{W}$.

3.3. Preservation of semantic properties of argumentation frameworks under aggregation

Typically, agents will disagree on whether certain attacks between arguments in Arg are in fact justified (if not, aggregation becomes trivial). But even when they disagree on the details, there may be high-level agreement on certain features. For example, maybe all agents agree that, under a particular semantics, argument A is acceptable. Whenever we observe such agreement on semantic features in a profile, we would like those features to be preserved under aggregation. Thus, for our example, under the same semantics, we would like A to be acceptable also in the argumentation framework computed by our aggregation rule. In other words, we are interested in the preservation of properties of argumentation frameworks (i.e., of the attack-relations that define them) under aggregation.

An example for a property is antisymmetry (i.e., the absence of mutual attacks between arguments). Another example is the existence of an argument that is not attacked by any other argument. But in some cases, what we are really interested in is the preservation of entire collections of properties. For example, for every argument A, we may want the acceptability of A under a certain semantics to be preserved under aggregation.

Formally, an *AF-property* $P \subseteq 2^{Arg \times Arg}$ is simply the set of all attack-relations on Arg that satisfy P. But in the interest of readability, we write $P(\rightharpoonup)$ rather than $(\rightharpoonup) \in P$. A *collection of AF-properties* $\mathcal{P}$ is a set of such AF-properties. Typically, the elements of $\mathcal{P}$ will be indexed by either the arguments $A \in Arg$ or the sets $\Delta \subseteq Arg$. Technically, every single AF-property P can also be thought of as a collection of AF-properties, namely $\mathcal{P} = \{P\}$.

Definition 16. Let F be an aggregation rule and let $\mathcal{P}$ be a collection of AF-properties. We say that F **preserves** $\mathcal{P}$, if for every profile $\rightharpoonup$ and every AF-property $P \in \mathcal{P}$ we have that $P(\rightharpoonup_i)$ being the case for all agents $i \in N$ implies $P(F(\rightharpoonup))$.

Thus, F preserves the single AF-property P if for every profile $\rightharpoonup$ we have that $P(\rightharpoonup_i)$ being the case for all agents $i \in N$ implies $P(F(\rightharpoonup))$. This notion of preservation (of a single property) is known under the name of *collective rationality* in other parts of social choice theory [41,42,24].

[7] Note that, in line with the existing literature in argumentation theory on the one hand and social choice theory on the other, we use the term "grounded" in two unrelated ways (grounded extensions *vs.* grounded aggregation rules).

[8] Thus, in particular the unanimity rule satisfies the unanimity axiom, but so do all other quota and oligarchic rules. Note that, while the unanimity axiom requires that all unanimously accepted attacks need to be returned by an aggregation rule, the unanimity rule returns *only* those unanimously accepted attacks.

Properties of interest. We now review the specific AF-properties for which we study preservation in this paper. Two of them we have already introduced in Section 2, namely *acyclicity* and *coherence*. Recall that these are attractive properties, because—if satisfied by an argumentation framework—they ensure that several different semantics will coincide and result in the same recommendations about which arguments to accept, thereby making decisions less controversial.

Both the grounded and the ideal semantics are attractive for two reasons. First, they encode a notion of scepticism in the sense of only accepting arguments we can have high confidence in [37]. Second, unlike the other extensions we have defined, the grounded and the ideal extension are always unique. On the downside, that unique extension may be empty, i.e., these semantics will sometimes not suggest any arguments to be accepted. Thus, argumentation frameworks that satisfy the AF-property of *nonemptiness of the grounded extension* or *nonemptiness of the ideal extension* are of particular interest.

Collections of properties of interest. Next, we turn to collections of AF-properties we may wish to preserve. Let $A \in Arg$ be one of the arguments under consideration. Then, for any given argumentation framework, A may or may not belong to the grounded extension. Thus, every $A \in Arg$ defines an AF-property, namely the property of membership of A in the grounded extension, i.e., of acceptance of A under the grounded semantics. This in itself would be too narrow a property to be of much interest for our purposes. However, what is of interest is whether membership is preserved for *all* arguments. We say that F *preserves argument acceptability under the grounded semantics*, if it is the case that, for all arguments $A \in Arg$, whenever A belongs to the grounded extension of $\langle Arg, \rightharpoonup_i\rangle$ for all agents $i \in N$, then A also belongs to the grounded extension of $\langle Arg, F(\rightharpoonup)\rangle$. Thus, argument acceptability under the grounded semantics is a collection of AF-properties, consisting of one AF-property for every argument $A \in Arg$. The collection of AF-properties of *argument acceptability under the ideal semantics* is defined accordingly.

For the stable, semi-stable, preferred, and complete semantics, we require a more refined definition, given that extensions under these semantics need not be unique. We say that F *preserves credulous argument acceptability under the stable semantics*, if it is the case that, for all arguments $A \in Arg$, whenever A belongs to *some* stable extension of $\langle Arg, \rightharpoonup_i\rangle$ for all agents $i \in N$, then A also belongs to *some* stable extension of $\langle Arg, F(\rightharpoonup)\rangle$. Analogously, F *preserves sceptical argument acceptability under the stable semantics*, if it is the case that, for all arguments $A \in Arg$, whenever A belongs to *all* stable extensions of $\langle Arg, \rightharpoonup_i\rangle$ for all agents $i \in N$, then A also belongs to *all* stable extensions of $\langle Arg, F(\rightharpoonup)\rangle$.[9] The corresponding concepts for the semi-stable, the preferred, and the complete semantics are defined accordingly. All of these are also collections of AF-properties, one for every argument $A \in Arg$.

Rather than just preserving the acceptability status of a single argument, we may also be interested in preserving entire extensions. For example, we say that F *preserves extensions under the stable semantics*, if it is the case that, for all sets $\Delta \subseteq Arg$, whenever Δ is a stable extension of $\langle Arg, \rightharpoonup_i\rangle$ for all agents $i \in N$, then Δ is also a stable extension of $\langle Arg, F(\rightharpoonup)\rangle$. So this again concerns the preservation of a collection of AF-properties, one for every set $\Delta \subseteq Arg$. The corresponding concept can be defined analogously for the other five semantics.

Similarly, we say that F *preserves conflict-freeness*, if it is the case that, for all sets $\Delta \subseteq Arg$, whenever Δ is conflict-free in $\langle Arg, \rightharpoonup_i\rangle$ for all agents $i \in N$, then Δ is also conflict-free in $\langle Arg, F(\rightharpoonup)\rangle$. Finally, *preservation of admissibility* is defined accordingly.

Summary. To summarise, we have identified the following AF-properties that, in case all agents agree on one of them being satisfied, we would like to see preserved under aggregation:

- *acyclicity* and *coherence* (reducing semantic ambiguity),
- *nonemptiness* of the grounded extension and the ideal extension (enabling a sceptical approach to argument evaluation),
- *argument acceptability* under different semantics (allowing for agreement on arguments even in the face of disagreement on the attacks between them), and
- the property of a set *being an extension* under different semantics or of being either *conflict-free* or *admissible* (also allowing for semantic agreement despite disagreement on attacks).

The latter two items concern collections of AF-properties (rather than single AF-properties), one for every argument $A \in Arg$ and every set $\Delta \subseteq Arg$, respectively.

Example 3. Consider again the profile of Fig. 2 and recall that the (weak or strict) majority rule will return the argumentation framework with $A \rightharpoonup B$, $B \rightharpoonup C$, and $C \rightharpoonup A$. Thus, the majority rule does not preserve acyclicity.[10] What about some of the other AF-properties? The grounded extension of AF_1 is $\{A, C, D\}$, that of AF_2 is $\{B, D\}$, that of AF_3 is $\{A, D\}$, and that of the majority outcome is $\{D\}$. Thus, preservation of the property of nonemptiness of the grounded extension is not violated by this particular example, given that the grounded extension of the majority outcome is nonempty. Preservation of argument acceptability under the grounded semantics also is not violated: the only argument contained in the grounded

[9] Recall that a stable extension need not exist, so sometimes this will hold vacuously.

[10] This observation is closely related to the famous *Condorcet Paradox* in the theory of preference aggregation [43].

extension of all three individual argumentation frameworks is D, and D is also included in the extension of the argumentation framework returned by the majority rule. Of course, this is not to say that these two properties might not be violated for other profiles. Finally, observe that also preservation of being the grounded extension is not violated by this example, given that the three agents do not agree on the grounded extension to begin with. △

While we believe that this is the first time that the notion of preservation of a semantic property has been developed systematically in the literature on collective argumentation, there are—as already noted in the introduction—some specific instances of this idea that have been discussed in earlier work, notably by Dunne et al. [13]. These authors define the notions of an aggregation rule preserving the nonemptiness of certain extensions (termed "σ-weak nontriviality" by Dunne et al.), of extension preservation ("σ-unanimity"), and of preserving either credulous or sceptical argument acceptability ("ca_σ-unanimity" and "sa_σ-unanimity", respectively). A few years earlier, Tohmé et al. [9] furthermore studied the preservation of acyclicity during the aggregation of argumentation frameworks.

4. Preservation results

In this section we present our results on the preservation of semantic properties under aggregation. This includes both positive and negative results: some properties can be preserved by intuitively appealing aggregation rules, while others require us to use rules that give veto powers or even dictatorial powers to some of the agents. Most of our results have the following form: if we look for an aggregation rule F that satisfies a certain combination of axioms and if we would like F to preserve a certain AF-property P (or a certain collection $\mathcal{P}$ of AF-properties), then F must belong to a certain family of aggregation rules.

Section overview. We begin with two very simple properties, namely conflict-freeness and admissibility. As we are going to see, the requirement of preserving admissibility is closely related to the neutrality axiom and this connection allows us to derive neutrality (rather than having to assume it) for several subsequent results. We then cover, in turn, results pertaining to the preservation of argument acceptability, extension preservation, preservation of the nonemptiness of uniquely determined extensions, and acyclicity and coherence.

4.1. Conflict-freeness, admissibility, and the neutrality axiom

Recall that a set of arguments is called conflict-free if it does not contain two arguments for which it is the case that the first attacks the second. Our first result demonstrates that this most basic property of sets of arguments is preserved under essentially all reasonable aggregation rules.

Theorem 1. *Every aggregation rule F that is grounded preserves conflict-freeness.*

Proof. Let F be an aggregation rule that is grounded. Consider any set $\Delta \subseteq Arg$ and any profile $\boldsymbol{\rightharpoonup} = (\rightharpoonup_1, \ldots, \rightharpoonup_n)$ such that Δ is conflict-free in $\langle Arg, \rightharpoonup_i \rangle$ for all $i \in N$. For the sake of contradiction, assume Δ is *not* conflict-free in $\langle Arg, F(\boldsymbol{\rightharpoonup})\rangle$, i.e., there exist two arguments $A, B \in \Delta$ such that $(A \rightharpoonup B) \in F(\boldsymbol{\rightharpoonup})$. Due to the groundness of F, there then must be at least one agent $i \in N$ such that also $A \rightharpoonup_i B$, i.e., Δ is not conflict-free in $\langle Arg, \rightharpoonup_i \rangle$ either. But this contradicts our assumption. □

Next, we turn to admissibility. Recall that a set of arguments is admissible if it is conflict-free and defends all of its members. Before we establish our main result regarding the preservation of admissibility, we are going to prove a lemma that unveils an interesting connection between admissibility and the neutrality axiom. It shows that every unanimous, grounded, and independent aggregation rule that preserves admissibility must be neutral. This lemma is similar in spirit to the *Contagion Lemma* in the literature on preference aggregation, which shows that any independent and Pareto efficient preference aggregation rule that preserves the transitivity of the input must be neutral [21]. In fact, we are first going to prove a more general lemma that shows that this implication holds not only for admissibility but for every collection of AF-properties of a certain type. This more general lemma is similar to a recent result by Endriss and Grandi [24, Lemma 12], which—when adapted to our terminology—states that any unanimous, grounded, and independent aggregation rule F that preserves some AF-property P must be neutral whenever P belongs to what they call the family of *contagious* properties. We now adapt this notion of contagiousness to collections of properties.[11]

[11] We stress that contagiousness is a meta-property—a property of (collections of) properties—that serves as a purely technical device we use in some of our proofs. It is of interest to the study of argumentation only in so far as one can show that semantic properties of argumentation frameworks (such as admissibility) that are of direct demonstrable interest to argumentation theory turn out to be properties that are contagious. As we are going to see, this indeed is the case. Analogous considerations apply to the meta-properties of implicativeness, disjunctiveness, and k-exclusiveness to be introduced in the sequel.

Definition 17. A collection $\mathcal{P}$ of AF-properties is called **contagious** if, for every distinct arguments $A, B, C \in Arg$, there exist a property $P \in \mathcal{P}$ and a set $Att \subseteq Arg \times Arg$ of attacks such that $\langle Arg, Att \cup S\rangle$ with $S \subseteq \{A \rightharpoonup B, B \rightharpoonup C\}$ satisfies P if and only if $S \neq \{B \rightharpoonup C\}$.

Thus, $\mathcal{P}$ is contagious if for every triple of arguments $A, B, C \in Arg$ we can find an AF-property P in the collection $\mathcal{P}$ such that satisfaction of P requires that $B \rightharpoonup C$ implies $A \rightharpoonup B$ (at least if the rest of the argumentation framework looks as specified by Att). A single AF-property P is contagious if $\mathcal{P} = \{P\}$ is.[12] This choice of terminology is intended to convey the idea that—in the context of Att and assuming you would like to satisfy P—accepting arguments is "contagious", given that accepting $B \rightharpoonup C$ forces you to also accept $A \rightharpoonup B$.

Lemma 2. *For $|Arg| \geqslant 3$, any unanimous, grounded, and independent aggregation rule F that preserves some contagious collection $\mathcal{P}$ of AF-properties must be neutral.*

Proof. Suppose $|Arg| \geqslant 3$, let $\mathcal{P}$ be a contagious collection of AF-properties, and let F be an aggregation rule that is unanimous, grounded, and independent and that preserves $\mathcal{P}$. Due to being independent, F can be described in terms of one family of winning coalitions $\mathcal{W}_{att}$ for every potential attack $att \in Arg \times Arg$. To show that F is neutral, we must prove that $\mathcal{W}_{att} = \mathcal{W}_{att'}$ for any two attacks $att, att' \in Arg \times Arg$. Now consider any three (distinct) arguments $A, B, C \in Arg$. We are going to prove $\mathcal{W}_{B \rightharpoonup C} \subseteq \mathcal{W}_{A \rightharpoonup B}$. As A, B, and C have been chosen arbitrarily, it then follows that $\mathcal{W}_{att} = \mathcal{W}_{att'}$ for all $att, att' \in Arg \times Arg$. To see this, suppose that $att = (\alpha \rightharpoonup \beta)$ and $att' = (\alpha' \rightharpoonup \beta')$. Then repeated application of the reasoning pattern we are about to establish yields $\mathcal{W}_{\alpha \rightharpoonup \beta} \subseteq \mathcal{W}_{\beta' \rightharpoonup \alpha} \subseteq \mathcal{W}_{\alpha' \rightharpoonup \beta'}$ as well as $\mathcal{W}_{\alpha' \rightharpoonup \beta'} \subseteq \mathcal{W}_{\beta \rightharpoonup \alpha'} \subseteq \mathcal{W}_{\alpha \rightharpoonup \beta}$, and thus $\mathcal{W}_{\alpha \rightharpoonup \beta} = \mathcal{W}_{\alpha' \rightharpoonup \beta'}$.

So pick an arbitrary coalition $\mathfrak{C} \in \mathcal{W}_{B \rightharpoonup C}$. Due to $\mathcal{P}$ being contagious, for our choice of A, B, and C there exist a property $P \in \mathcal{P}$ and a set $Att \subseteq Arg \times Arg$ such that $\langle Arg, Att \cup S\rangle$ with $S \subseteq \{A \rightharpoonup B, B \rightharpoonup C\}$ satisfies P if and only if $S \neq \{B \rightharpoonup C\}$. Construct a profile $\boldsymbol{\rightharpoonup}$ in which exactly the agents in $\mathfrak{C}$ report the attack-relation $Att \cup \{A \rightharpoonup B, B \rightharpoonup C\}$ and all others report Att. Thus, in this profile all individual attack-relations satisfy P. As F preserves $\mathcal{P}$, the outcome $F(\boldsymbol{\rightharpoonup})$ must satisfy P as well. Due to the unanimity and groundedness of F, $F(\boldsymbol{\rightharpoonup})$ must be of the form $Att \cup S$ with $S \subseteq \{A \rightharpoonup B, B \rightharpoonup C\}$. Due to $\mathfrak{C}$ being a winning coalition, we furthermore must have $(B \rightharpoonup C) \in F(\boldsymbol{\rightharpoonup})$. This, together with the fact that $F(\boldsymbol{\rightharpoonup})$ must satisfy P means that we must have $(A \rightharpoonup B) \in F(\boldsymbol{\rightharpoonup})$ as well. But this means that coalition $\mathfrak{C}$ succeeded in getting attack $A \rightharpoonup B$ accepted, i.e., $\mathfrak{C}$ must be a winning coalition also for this attack. Thus, we have succeeded in deriving $\mathfrak{C} \in \mathcal{W}_{A \rightharpoonup B}$ and are done. $\square$

Lemma 3. *For $|Arg| \geqslant 3$, any unanimous, grounded, and independent aggregation rule F that preserves admissibility must be neutral.*

Proof. The claim follows from Lemma 2, provided we can show that admissibility is a contagious collection of AF-properties. So consider any distinct $A, B, C \in Arg$. With reference to Definition 17, let P be the property of the set $\{A, C\}$ being admissible and let Att be the empty set of attacks. Now consider the four argumentation frameworks of the form $\langle Arg, Att \cup S\rangle$ with $S \subseteq \{A \rightharpoonup B, B \rightharpoonup C\}$. By the definition of admissibility, the only argumentation framework of this kind that does not satisfy P is the one we get for $S = \{B \rightharpoonup C\}$. This concludes the proof. $\square$

We are now ready to state our main result regarding the preservation of admissibility. It is significantly less broad than Theorem 1, our result for the preservation of conflict-freeness, but it still clearly is a positive result. It shows that there exists a reasonable rule that preserves the admissibility of arbitrary sets of arguments.

Theorem 4. *For $|Arg| \geqslant 4$, the only unanimous, grounded, anonymous, independent, and monotonic aggregation rule F that preserves admissibility is the nomination rule.*

Proof. We first show that the nomination rule indeed preserves admissibility. So let F be the nomination rule. Consider any set $\Delta \subseteq Arg$ and any profile $\boldsymbol{\rightharpoonup} = (\rightharpoonup_1, \ldots, \rightharpoonup_n)$ such that Δ is admissible in $AF_i = \langle Arg, \rightharpoonup_i\rangle$ for all $i \in N$. For the sake of contradiction, assume Δ is *not* admissible in $\langle Arg, F(\boldsymbol{\rightharpoonup})\rangle$, i.e., there is an argument $A \in \Delta$ that, in $F(\boldsymbol{\rightharpoonup})$, is attacked by an argument $B \in Arg \setminus \Delta$ and there does not exist a $C \in \Delta$ such that $(C \rightharpoonup B) \in F(\boldsymbol{\rightharpoonup})$. As $(B \rightharpoonup A) \in F(\boldsymbol{\rightharpoonup})$ and as F is grounded, we must have $B \rightharpoonup_i A$ for some $i \in N$. And as there does not exist a $C \in \Delta$ such that $(C \rightharpoonup A) \in F(\boldsymbol{\rightharpoonup})$, given the definition of the nomination rule, there cannot exist an argument $C \in \Delta$ such that $C \rightharpoonup_i A$ for that same agent i. Hence, Δ is not admissible in AF_i, in contradiction to our original assumption.

We still need to show that there can be no other aggregation rule than the nomination rule that preserves admissibility and that satisfies all of the axioms mentioned in the statement of Theorem 4. Let F be a unanimous, grounded, independent, and monotonic aggregation rule that preserves admissibility. By Lemma 3, we know that F is also neutral. So, the claim is

[12] This definition of contagiousness of a single property P is a special case of the more complex definition given by Endriss and Grandi [24]. We do not require the greater generality of the original definition for our purposes here.

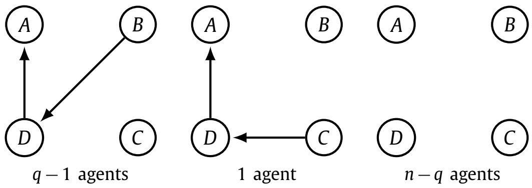

Fig. 3. Profile used in the proof of Theorem 4.

equivalent to the claim that for $|Arg| \geqslant 4$, the only unanimous, grounded, anonymous, neutral, independent, and monotonic aggregation rule F that preserves admissibility is the nomination rule. By the characterisation result for quota rules due to Dietrich and List [40] in the context of judgment aggregation, which has been adapted to graph aggregation by Endriss and Grandi [24] and which we have briefly recalled near the end of Section 3.2, this claim is equivalent to the claim that *no quota rule* F_q with a quota $q > 1$ always preserves admissibility. So let us prove this.

Consider the generic profile shown in Fig. 3 (and note that $q > 1$ ensures $q - 1 > 0$, i.e., there is at least one agent of the first kind). The set $\{A, B, C\}$ is admissible in all argumentation frameworks in such a profile. But when we aggregate using a quota rule F_q with a quota $q > 1$, we obtain an argumentation framework with a single attack $D \rightharpoonup A$, which means that A cannot be part of any admissible set. Hence, no such rule can preserve admissibility. □

4.2. Credulous and sceptical argument acceptability

Recall that an argument is credulously accepted under a given semantics if it is a member of at least one extension under that semantics. It is sceptically accepted if it is a member of every extension. In the case of the grounded and the ideal semantics, the notions of credulous and sceptical acceptability coincide. We are going to demonstrate that preserving credulous or sceptical acceptability of an argument when using a "simple" aggregation rule is impossible, unless we are willing to use a dictatorship. This is true under any of the six semantics. To prove this result—and some of those that follow—we are going to use a technique developed by Endriss and Grandi [24] for the more general framework of graph aggregation, which in turn was inspired by the seminal work on preference aggregation of Arrow [41]. It amounts to showing that, under certain assumptions, the families of winning coalitions defining an aggregation rule must form an ultrafilter. The technique developed by Endriss and Grandi, however, greatly simplifies the process of deriving such results. We only need to show that the properties preserved under aggregation are of a certain type.

Using our present terminology, Endriss and Grandi [24, Theorem 18] show that, for argumentation frameworks with three or more arguments, any aggregation rule that satisfies certain basic axioms and that is supposed to preserve some AF-property P must be a dictatorship—at least in case P belongs to what they call the family of *implicative* and *disjunctive* properties. Let us first adapt these two concepts to our needs.

Definition 18. An AF-property P is called **implicative** if there exist a set $Att \subseteq Arg \times Arg$ of attacks and three individual attacks $att_1, att_2, att_3 \in Arg \times Arg \setminus Att$ such that $\langle Arg, Att \cup S \rangle$ with $S \subseteq \{att_1, att_2, att_3\}$ satisfies P if and only if $S \neq \{att_1, att_2\}$.

Definition 19. An AF-property P is called **disjunctive** if there exist a set $Att \subseteq Arg \times Arg$ of attacks and two individual attacks $att_1, att_2 \in Arg \times Arg \setminus Att$ such that $\langle Arg, Att \cup S \rangle$ with $S \subseteq \{att_1, att_2\}$ satisfies P if and only if $S \neq \emptyset$.

Thus, an implicative property P requires that, in the context of Att, accepting att_1 and att_2 implies accepting att_3 (and all seven patterns of acceptance consistent with that requirement are possible). A disjunction AF-property P requires that, given Att, we must accept at least one of att_1 and att_2 (and all three patterns of acceptance consistent with that requirement are possible).[13] We call a collection $\mathcal{P}$ of AF-properties implicative if it includes at least one implicative property. Disjunctive collections of properties are defined analogously. Note that these definitions are different in nature from the definition of contagiousness. Contagiousness requires *every* $P \in \mathcal{P}$ to satisfy certain requirements, and those requirements concern *all* triples $A, B, C \in Arg$, while for both implicativeness and disjunctiveness we merely have to find a single pattern of the relevant kind. For this reason, we are able to directly reuse the results of Endriss and Grandi [24] regarding implicativeness and disjunctiveness here, while we had to prove Lemma 2 from scratch.

With all the relevant definitions now in place, we can formally restate the result of Endriss and Grandi [24, Theorem 18] using our present terminology as follows:

Let $\mathcal{P}$ be a collection of AF-properties that is both implicative and disjunctive. Then, for $|Arg| \geqslant 3$, any unanimous, grounded, neutral, and independent aggregation rule F that preserves $\mathcal{P}$ must be a dictatorship.

[13] Our definitions of implicativeness and disjunctiveness are special cases of the more general definitions given by Endriss and Grandi [24]. They simplify exposition and are sufficient for our purposes here.

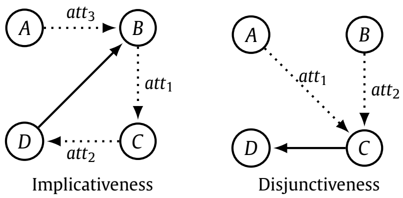

Fig. 4. Scenarios used in the proof of Theorem 6.

We are now ready to state and prove our result on the preservation of argument acceptability.[14] It relies on the following lemma, the proof of which can be found in the appendix. Like the proof of Lemma 3, it is a simple application of Lemma 2.

Lemma 5. *Let $\mathcal{P}$ be the collection of AF-properties representing either credulous or sceptical argument acceptability under either the grounded, the ideal, the complete, the preferred, the semi-stable, or the stable semantics. Then, for $|Arg| \geqslant 4$, any unanimous, grounded, and independent aggregation rule F that preserves $\mathcal{P}$ must be neutral.*

Theorem 6. *Let $\mathcal{P}$ be the collection of AF-properties representing either credulous or sceptical argument acceptability under either the grounded, the ideal, the complete, the preferred, the semi-stable, or the stable semantics. Then, for $|Arg| \geqslant 4$, any unanimous, grounded, and independent aggregation rule F that preserves $\mathcal{P}$ must be a dictatorship.*

Proof. Let $Arg = \{A, B, C, D, \ldots\}$, let $\mathcal{P}$ be one of the twelve collections of AF-properties of interest (credulous or sceptical acceptability under one of the six semantics), and let F be defined as in the statement of the theorem. By Lemma 5, F must also be neutral. Thus, by the aforementioned result of Endriss and Grandi [24, Theorem 18], we are done if we can show that each of the twelve instances of $\mathcal{P}$ is both implicative and disjunctive.

Let us first prove implicativeness. Suppose we are interested in the acceptability of argument C. Let $Att = \{D \rightharpoonup B\}$, $att_1 = (B \rightharpoonup C)$, $att_2 = (C \rightharpoonup D)$, and $att_3 = (A \rightharpoonup B)$. This scenario is sketched in the lefthand part of Fig. 4. Now consider the argumentation frameworks of the form $\langle Arg, Att \cup S\rangle$ with $S \subseteq \{att_1, att_2, att_3\}$. If $S \subseteq \{att_2, att_3\}$, then C is not attacked by any other argument. If $S = \{att_1\}$ or $S = \{att_1, att_3\}$, then C is defended by D, which is not attacked by any other argument. If $S = \{att_1, att_2, att_3\}$, then C is defended by A, which is not attacked by any other argument. Thus, in all of these seven cases, either C is not attacked by any other argument or it is defended by an argument that is not attacked by any other argument. This implies that C must be part of the grounded extension. Hence, C is both credulously and sceptically accepted under each of the six semantics. On the other hand, if $S = \{att_1, att_2\}$, then $\{B, C, D\}$ forms an isolated odd-length cycle. This means that C is neither credulously nor sceptically acceptable under any of the six semantics. We have thus found a set of attacks Att and three individual attacks att_1, att_2, and att_3 such that $P(Att \cup S)$ if and only if $S \neq \{att_1, att_2\}$, where P is either the property of credulous or of sceptical acceptability of C under either one of the six semantics. Hence, the collection $\mathcal{P}$, which includes P, is implicative.

Next, we show disjunctiveness. Suppose we are interested in the acceptability of D. Let $Att = \{C \rightharpoonup D\}$, $att_1 = (A \rightharpoonup C)$, and $att_2 = (B \rightharpoonup C)$. This scenario is depicted on the righthand side of Fig. 4. Consider all argumentation frameworks $\langle Arg, Att \cup S\rangle$ with $S \subseteq \{att_1, att_2\}$. If $S = \{att_1\}$, then D is defended by A. If $S = \{att_2\}$, then D is defended by B. If $S = \{att_1, att_2\}$, then D is defended by both A and B. In all three cases, D is defended by some argument that is not attacked by any other argument. This implies that D must be part of the grounded extension and thus both credulously and sceptically accepted under all six semantics. However, if $S = \emptyset$, then D is attacked by C and not defended by any other argument, which means that D is neither credulously nor sceptically acceptable under any of the six semantics. To summarise, we have seen that $P(Att \cup S)$ if and only if $S \neq \emptyset$, where P is the property of credulous or of sceptical acceptability of D under either one of the six semantics. Hence, $\mathcal{P}$ is a disjunctive collection of AF-properties. □

Recall that Theorem 6 applies only when $|Arg| \geqslant 4$. This covers all cases of practical interest, but from a purely technical point of view one might still wonder whether the theorem could maybe be strengthened to $|Arg| \geqslant 3$. We conjecture that the bound on the cardinality of Arg used in Theorem 6 and all similar bounds in later theorems are sharp, but we have no been able to verify this conjecture in all cases. Only in some cases are there obvious counterexamples. For example, we know that for $|Arg| = 3$ argument acceptability under the grounded semantics is preserved by the strict majority rule,[15] which of course satisfies the three axioms mentioned in Theorem 6.

[14] Recall that for the grounded and the ideal semantics, the notions of credulous and sceptical acceptability coincide, i.e., for them the formulation of the theorem could be simplified, by simply speaking of "argument acceptability".

[15] To see this, suppose $Arg = \{A, B, C\}$ and focus on the acceptability of C. If C is part of the grounded extension for all individual argumentation frameworks, then in each of them either C is not attacked at all or it is defended by a third argument that is itself not attacked. In the latter case, w.l.o.g., suppose C is attacked by B, which is attacked by A, which is not attacked by any argument. Thus, if there is a strict majority for $B \rightharpoonup C$, then there also must be a strict majority for $A \rightharpoonup B$. Hence, either C is not attacked in the outcome, or it is defended successfully.

4.3. The property of being an extension

Next, we turn to the property of a given set of arguments being an extension under one of the six semantics. We obtain impossibility results for five out of the six semantics and—somewhat surprisingly—a positive result for the stable semantics. Our impossibility results differ very subtly for the grounded and the ideal semantics on the one hand, and the complete, the preferred, and the semi-stable semantics on the other. In all five cases, they show that the preservation of extensions is impossible by means of a "simple" aggregation rule, unless we are willing to use a dictatorship.

Theorem 7. *For $|Arg| \geqslant 4$, any unanimous, grounded, and independent aggregation rule F that preserves either grounded or ideal extensions must be a dictatorship.*

Theorem 8. *For $|Arg| \geqslant 5$, any unanimous, grounded, and independent aggregation rule F that preserves either complete, preferred, or semi-stable extensions must be a dictatorship.*

Proofs of both theorems can be found in the appendix. They employ the same technique as for the proof of Theorem 6: we first show that the relevant collections of AF-properties are contagious (to be able to apply Lemma 2) and then that they are both implicative and disjunctive (to be able to apply the result of Endriss and Grandi [24]).

Note that Theorem 7 and Theorem 8 differ with respect to the number of arguments required for the result to apply. Again, we do not know whether all of these bounds are sharp, but we do know that the same proof technique cannot be used to lower the bounds stated in the theorems. For instance, for Theorem 8 the same kind of proof does not go through for $|Arg| \geqslant 4$. To verify this, we have written a computer program that enumerates all relevant scenarios involving four arguments and found that we cannot establish the conditions for implicativeness in this manner. On the other hand, we also know that neither the strict majority rule nor any other quota rule can be used to construct a counterexample to Theorem 7 for $|Arg| = 3$.[16] So, while we still conjecture the bound stated in that theorem to be sharp, proving that it is by finding a counterexample is more difficult in this case than it is for Theorem 6.

Interestingly, for the preservation of stable extensions we obtain a much more positive result:

Proposition 9. *The nomination rule preserves stable extensions.*

Proof. Let F be the nomination rule. Consider any set $\Delta \subseteq Arg$ and any profile $\rightharpoonup = (\rightharpoonup_1, \ldots, \rightharpoonup_n)$ such that Δ is stable in $\langle Arg, \rightharpoonup_i\rangle$ for all $i \in N$. According to Theorem 1, given that F is grounded, F preserves conflict-freeness. Thus, Δ is conflict-free in $\langle Arg, F(\rightharpoonup)\rangle$.

What remains to be shown is that Δ attacks every argument $B \in Arg \setminus \Delta$. In case $\Delta = Arg$, the claim holds vacuously. Otherwise, consider an arbitrary argument $B \in Arg \setminus \Delta$. We need to show that B is attacked by some argument in Δ in $F(\rightharpoonup)$. Take the argumentation framework $AF_i = \langle Arg, \rightharpoonup_i\rangle$ for some $i \in N$. As Δ is stable in AF_i by assumption, there exists an argument $A \in \Delta$ such that $A \rightharpoonup_i B$. As F is the nomination rule, we also get $(A \rightharpoonup B) \in F(\rightharpoonup)$ as claimed. $\square$

4.4. Nonemptiness of the grounded and the ideal extension

We have seen that preserving the grounded extension and the ideal extension is impossible for simple yet reasonable aggregation rules (see Theorem 7). What about the seemingly less demanding requirement of at least preserving nonemptiness of the unique extension defined by each one of these two semantics? First, the bad news is that for the ideal semantics this intuition fails and the same kind of impossibility prevails.

Theorem 10. *For $|Arg| \geqslant 4$, any unanimous, grounded, and independent aggregation rule F that preserves nonemptiness of the ideal extension must be a dictatorship.*

We prove Theorem 10 in the appendix by showing that nonemptiness of the ideal extension is an AF-property that is contagious, implicative, and disjunctive.

For the grounded semantics, however, we can do better. For instance, it is easy to check that the unanimity rule preserves nonemptiness of the grounded extension. Still, as we shall see next, we cannot do *much* better: only rules that grant veto powers to some agents will work. Recall that the grounded extension is nonempty if an only if at least one argument is not attacked by any other argument. Thus, this AF-property is about the *absence* of attacks, while the technique we employed to prove Theorem 7 (and all other impossibility results we have encountered so far) exploits the *presence* of certain attacks (to

[16] For example, to see that the strict majority rule does not preserve grounded extensions in all cases when $|Arg| = 3$ (even though, as we have seen, it does preserve acceptability of individual arguments under the grounded semantics), consider the profile with two agents where the first agent reports $\{A \rightharpoonup B, B \rightharpoonup C\}$ and the second reports $\{C \rightharpoonup B, B \rightharpoonup A\}$. Both individual grounded extensions are equal to $\{A, C\}$, but the strict majority outcome is the argumentation framework without any attacks and thus has the grounded extension $\{A, B, C\}$.

see this, recall the definitions of implicativeness and disjunctiveness). We are now going to present our preservation result regarding the nonemptiness of the grounded extension as a corollary to a more general theorem about the preservation of AF-properties that require the absence of certain attacks. We first define a suitable meta-property.

Definition 20. Let $k \in \mathbb{N}$. An AF-property P is called ***k*-exclusive** if there exist k distinct attacks $att_1, \ldots, att_k \in Arg \times Arg$ such that (i) $\{att_1, \ldots, att_k\} \subseteq (\rightharpoonup)$ for no attack-relation $\rightharpoonup$ with $P(\rightharpoonup)$, and (ii) for every $S \subsetneq \{att_1, \ldots, att_k\}$ there exists an attack-relation $\rightharpoonup$ such that $S \subseteq (\rightharpoonup)$ and $P(\rightharpoonup)$.

Thus, you cannot accept all k attacks, but you should be able to accept any proper subset of them. We call of collection $\mathcal{P}$ of AF-properties k-exclusive if at least one property $P \in \mathcal{P}$ is k-exclusive.[17] We are able to prove the following powerful theorem (recall that n is the number of agents in N).

Theorem 11. *Let $k \geqslant n$ and let P be an AF-property that is k-exclusive. Then under any neutral and independent aggregation rule F that preserves P at least one agent must have veto powers.*

Proof. Let $k \geqslant n$, let P be an AF-property that is k-exclusive, and let F be an aggregation rule that is neutral and independent. We need to show that, if F preserves P, then F must give some agents the power to veto the collective acceptance of attacks.

First, observe that, if an aggregation rule F is both neutral and independent, then there exists a (single) family of winning coalitions $\mathcal{W} \subseteq 2^N$ such that, for all profiles $\boldsymbol{\rightharpoonup}$ and all potential attacks $att \in Arg \times Arg$, the following relationship holds:

$$att \in F(\boldsymbol{\rightharpoonup}) \text{ if and only if } N^{\boldsymbol{\rightharpoonup}}_{att} \in \mathcal{W}$$

Recall that $i \in N$ having veto powers under F means that $F(\boldsymbol{\rightharpoonup}) \subseteq (\rightharpoonup_i)$ for every profile $\boldsymbol{\rightharpoonup}$. Let us show that an agent $i \in N$ has veto powers, if she is a member of all winning coalitions:

$$i \in \bigcap_{\mathfrak{C} \in \mathcal{W}} \mathfrak{C} \text{ implies } F(\boldsymbol{\rightharpoonup}) \subseteq (\rightharpoonup_i) \text{ for all profiles } \boldsymbol{\rightharpoonup}$$

If $\bigcap_{\mathfrak{C} \in \mathcal{W}} \mathfrak{C} = \emptyset$, then the above claim holds vacuously. Otherwise, take any attack $att \in F(\boldsymbol{\rightharpoonup})$. As att got accepted, $N^{\boldsymbol{\rightharpoonup}}_{att}$ must be a winning coalition, i.e., $N^{\boldsymbol{\rightharpoonup}}_{att} \in \mathcal{W}$ and therefore $i \in N^{\boldsymbol{\rightharpoonup}}_{att}$. But this is just another way of saying $att \in (\rightharpoonup_i)$, so we are done.

Next, we are going to show that the fact that F preserves the k-exclusive AF-property P implies that the intersection of any k winning coalitions must be nonempty:

$$\mathfrak{C}_1 \cap \cdots \cap \mathfrak{C}_k \neq \emptyset \text{ for all } \mathfrak{C}_1, \ldots, \mathfrak{C}_k \in \mathcal{W}$$

For the sake of contradiction, assume there do exist winning coalitions $\mathfrak{C}_1, \ldots, \mathfrak{C}_k \in \mathcal{W}$ such that $\mathfrak{C}_1 \cap \cdots \cap \mathfrak{C}_k = \emptyset$. We construct a profile $\boldsymbol{\rightharpoonup} = (\rightharpoonup_1, \ldots, \rightharpoonup_n)$ with $P(\rightharpoonup_i)$ for all $i \in N$ as follows: for every $j \in \{1, \ldots, k\}$, exactly the agents in $\mathfrak{C}_j$ accept attack att_j (for all other attacks, it is irrelevant which agents accept them). As, by our assumption, no agent is a member of all k winning coalitions, no agent accepts all k attacks, so this construction indeed is possible for a k-exclusive property such as P. However, as each of the k attacks is supported by a winning coalition, they all get accepted, i.e., $\{att_1, \ldots, att_k\} \subseteq F(\boldsymbol{\rightharpoonup})$, meaning that the outcome does *not* satisfy P. Thus, we have found a contradiction to our assumption of F preserving P and are done.

Let us briefly recap where we are at this point. We know that F is characterised by a family of winning coalitions $\mathcal{W}$. We also know that $\mathfrak{C}_1 \cap \cdots \cap \mathfrak{C}_k \neq \emptyset$ for all $\mathfrak{C}_1, \ldots, \mathfrak{C}_k \in \mathcal{W}$. We need to show that some agents have veto powers, and we know that this is the case if we can prove that $\mathfrak{C}^{(1)} \cap \cdots \cap \mathfrak{C}^{(\ell)} \neq \emptyset$, where $\{\mathfrak{C}^{(1)}, \ldots, \mathfrak{C}^{(\ell)}\}$ is some enumeration of the coalitions in $\mathcal{W}$. Thus, we are done, if we can show that $\mathfrak{C}_1 \cap \cdots \cap \mathfrak{C}_k \neq \emptyset$ for *all* $\mathfrak{C}_1, \ldots, \mathfrak{C}_k \in \mathcal{W}$ implies $\mathfrak{C}^{(1)} \cap \cdots \cap \mathfrak{C}^{(\ell)} \neq \emptyset$. We are going to prove the contrapositive, namely that the following holds for *some* $\mathfrak{C}_1, \ldots, \mathfrak{C}_k \in \mathcal{W}$:

$$\mathfrak{C}^{(1)} \cap \cdots \cap \mathfrak{C}^{(\ell)} = \emptyset \text{ implies } \mathfrak{C}_1 \cap \cdots \cap \mathfrak{C}_k = \emptyset$$

In words, we need to show that in case the intersection of *all* winning coalitions is empty, then so is at least one intersection of just k winning coalitions.

Recall that we have assumed $k \geqslant n$. We construct a set $\mathcal{W}' \subseteq \mathcal{W}$ of k (or fewer) winning coalitions as follows. Initially, set $\mathcal{W}' := \emptyset$. Then, for every j from 1 to ℓ in turn, add $\mathfrak{C}^{(j)}$ to $\mathcal{W}'$ if and only if the following condition is satisfied:[18]

[17] We include this generalisation of the definition of k-exclusiveness for the sake of completeness, even though, in this paper, we only apply the concept of k-exclusiveness to single AF-properties.

[18] By convention, let $\bigcap_{\mathfrak{C} \in \emptyset} \mathfrak{C} = N$, i.e., the intersection of *no* winning coalitions is defined as the universe N of *all* agents.

$$\left(\mathfrak{C}^{(j)} \cap \bigcap_{\mathfrak{C}\in\mathcal{W}'} \mathfrak{C}\right) \subsetneq \left(\bigcap_{\mathfrak{C}\in\mathcal{W}'} \mathfrak{C}\right)$$

Thus, every additional $\mathfrak{C}^{(j)}$ is selected only if it causes the removal of at least one further agent from the intersection. As there are only n agents, we therefore will pick at most n coalitions. Hence, we will indeed arrive at a family $\mathcal{W}'$ of n or fewer—and thus certainly at most k—winning coalitions, the intersection of which is empty. This completes the proof. □

We note that, unlike for impossibility theorems such as Theorem 6, for Theorem 11 (and the results we are going to prove in the sequel by reference to this theorem), it is not possible to remove the neutrality axiom from the set of assumptions and to instead derive neutrality using independence and the requirement of preserving P. Indeed, it is easy to construct counterexamples. One such counterexample is the rule that always rejects all attacks except for $A \rightharpoonup B$, on which it decides by majority. This rule is independent and does not grant veto powers to any of the agents, yet it guarantees preservation of any k-exclusive AF-property.[19]

Let us now return to the issue of the preservation of the nonemptiness of the grounded extension. It suffices to show that this AF-property is an $|Arg|$-exclusive property to obtain the following result.

Theorem 12. *If $|Arg| \geqslant n$, then under any neutral and independent aggregation rule F that preserves nonemptiness of the grounded extension at least one agent must have veto powers.*

Proof. To obtain the claim as a corollary to Theorem 11, we need to show that the property of an argumentation framework having a nonempty grounded extension is a k-exclusive AF-property for $k = |Arg|$. Recall that having a nonempty grounded extension is equivalent to the property of having at least one argument that is not attacked by any other argument. We are going to show that the latter property is k-exclusive for $k = |Arg|$.

So let $k = |Arg|$. If $k = 1$, then the claim holds vacuously. So, w.l.o.g., assume that $k > 1$. Take an arbitrary enumeration $\{A^{(1)}, \ldots, A^{(k)}\}$ of Arg and consider the set of attacks $\{att_1, \ldots, att_k\}$ with $att_i = (A^{(i)} \rightharpoonup A^{(i+1)})$ for $i < k$ and $att_k = (A^{(k)} \rightharpoonup A^{(1)})$. Clearly, this set of attacks meets our requirements: (i) if $\{att_1, \ldots, att_k\} \subseteq (\rightharpoonup)$, then $\rightharpoonup$ does not have the property of leaving at least one argument without an attacker and (ii) for every $S \subsetneq \{att_1, \ldots, att_k\}$ there exists an attack-relation $\rightharpoonup$ with $S \subseteq (\rightharpoonup)$, namely S itself, that does leave one or more arguments without an attacker. □

We note that it is not difficult to prove that the converse of Theorem 12 holds as well: all rules that grant veto powers to at least one agent preserve nonemptiness of the grounded extension. To see this, observe that, if we start with an argumentation framework with a nonempty grounded extension (and thus at least one unattacked argument) and remove some of the attacks, then the grounded extension will remain nonempty (as that same argument remains unattacked). Therefore, as long as at least one agent with veto powers submits an argumentation framework in which at least one argument is unattacked, the same will be true for the outcome.

4.5. Acyclicity and coherence

The final group of AF-properties for which we wish to analyse the conditions under which they can be preserved under aggregation are properties that guarantee that several of the argumentation semantics agree on what arguments are (credulously or sceptically) acceptable. Recall that acyclicity guarantees that all six semantics agree with the grounded semantics and thus unambiguously define which arguments to accept. Also recall that coherence is a weaker property that ensures that the stable, the semi-stable, and the preferred semantics coincide. It is defined as the AF-property of every preferred extension being a stable extension (the rest follows from the known relationships between these three semantics).

Acyclicity is a prime example for a k-exclusive property, so we immediately obtain the following result as another simple corollary to Theorem 11.[20]

Theorem 13. *If $|Arg| \geqslant n$, then under any neutral and independent aggregation rule F that preserves acyclicity at least one agent must have veto powers.*

For the sake of completeness, the straightforward proof is given in the appendix. We note that, just as for Theorem 12, the converse of Theorem 13 is immediately seen to hold as well, i.e., all aggregation rules that grant veto powers to some agents clearly preserve acyclicity. This includes the qualified majority rules studied by Tohmé et al. [9].

[19] For $k > 1$, this holds vacuously, as the outcome will at most include the single attack $A \rightharpoonup B$. For $k = 1$ and if the attack of interest is $att_1 = (A \rightharpoonup B)$, then no individual agent is allowed to accept att_1, so it will not be collectively accepted either. Note that this construction crucially depends on the aggregation rule violating the neutrality axiom.

[20] Theorem 13 was anticipated in the work of Tohmé et al. [9], who make a similar claim, but without appealing to the neutrality axiom. We stress that Theorem 13 cannot be strengthened by dropping neutrality from the set of assumptions. Indeed, there are rules that preserve acyclicity, that are independent (but not neutral), and that do not give veto powers (regarding all potential attacks) to any of the agents. An example, for $N = \{1, 2\}$ and $Arg = \{A, B\}$, is the rule that accepts $A \rightharpoonup B$ if at least one agent does and that accepts $B \rightharpoonup A$ if both agents do.

Finally, regarding the preservation of the coherence of argumentation frameworks, we obtain the following impossibility result.

Theorem 14. *For $|Arg| \geqslant 4$, any unanimous, grounded, and independent aggregation rule F that preserves coherence must be a dictatorship.*

Thus, somewhat surprisingly, even though acyclicity is a stronger property than coherence, it is easier to preserve under aggregation. The proof of Theorem 14 can be found in the appendix. It amounts to showing that coherence is an AF-property that is contagious, implicative, and disjunctive, i.e., this is yet another application of the technique of Endriss and Grandi [24].

5. Conclusion

Using a variety of techniques, we have attempted to paint a clear picture of the capabilities and limitations of simple aggregation rules regarding the preservation of properties related to the semantics of abstract argumentation frameworks. While the significance of this issue and the promise of social choice theory for its resolution have previously been emphasised in the work of several authors [9,13,14,16], this is the first systematic analysis of its kind. Our results show that only the most basic of properties, namely conflict-freeness, is preserved by essentially all rules. More demanding properties require either the nomination rule, a rule granting some agents veto powers, or a rule that is dictatorial. Thus, the rules imposed on us by these results range from the positive, to the highly restrictive, to the clearly unacceptable.

We stress that these results only apply to *simple* rules, in particular, to rules that satisfy the axiom of independence. Using aggregation rules that are independent has the advantage that we can focus on one attack at a time when we determine the outcome, thereby simplifying the process of aggregation in both conceptual and computational terms. On the other hand, requiring independence clearly limits the design space for aggregation rules we can explore and prevents us from incorporating complex dependencies in the aggregation process. An alternative route, the one chosen by Coste-Marquis et al. [8], is to use distance-based rules (which violate independence). Such rules can be designed so as to guarantee specific properties of the outcome, so the question of preservation does not arise. On the downside, distance-based rules are computationally intractable [44–46]. We also stress that our results are based on the assumption that all agents report attack-relations over a single *common* set of arguments. Richer models, where different agents may choose to put forward of different sets of arguments [see, e.g., 8,15], are clearly of great interest as well and should be studied in future work. Finally, we stress that our results apply to one specific and highly *abstract* model of argumentation only [7], albeit one that has been exceptionally well received by the scientific community. Future work should also be directed at understanding to what extent our approach can be applied to models other than Dung's classical model of abstract argumentation. Natural candidates for such models are, first and foremost, those that are relatively close to Dung's model, such as bipolar abstract argumentation systems [47] or Bench-Capon's value-based argumentation frameworks [48]. In the long run, we recommend to attempt also going beyond such abstract models of argumentation and to consider structured forms of argumentation that account for the internal logical structure of individual arguments and thereby come closer to accurately modelling features of argumentation found in debates occurring in the real world [see, e.g., 4,33].

At the methodological level, we believe that Theorem 11, which shows that simple rules that preserve k-exclusive properties must give some agents veto powers, is of particular interest as it likely will find application also beyond the confines of abstract argumentation, in the same way as the results of Endriss and Grandi [24] on contagious, implicative, and disjunctive properties can be applied to a range of domains of graph aggregation. Also of some methodological interest are our simplified forms of the meta-properties of contagiousness, implicativeness, and disjunctiveness originally due to Endriss and Grandi and our observation that these meta-properties can not only be applied to single properties of graphs but also to collections of properties. For contagiousness, in particular, this makes a significant difference, as a collection of properties can be contagious even if no single property in that collection is contagious.

Staying with the theme of methodology for a moment, the techniques we have employed to prove impossibility theorems reduce the task of finding a proof to the task of identifying suitable scenarios that show that a given property (or collection of properties) is contagious, implicative, disjunctive, or k-exclusive. Once such a scenario is found, presenting and verifying the proof is routine, but finding such a scenario can be difficult. In some cases we have found these scenarios with the help of a computer program (the same program we used to verify that our techniques cannot be used to lower the bounds on the number of arguments for the theorems reported in Section 4.3) and in some cases we have verified the correctness of the constructions on which our proofs rely using an existing tool for computing extensions of abstract argumentation frameworks [38]. This suggests that there is room for applying automated reasoning tools in this domain, an approach that recently has been used very successfully in several other areas of computational social choice [49]. Investigating this point further constitutes another promising direction for future work.

Finally, there are natural opportunities for investigating application scenarios of our work. One such scenario has been presented in recent work by Shi et al. [50], who propose an approach for modelling an agent's beliefs in which *beliefs are grounded in arguments*. If one were to attempt to extend this model to also allow for the representation of groups of agents as well as their beliefs and pieces of evidence for those beliefs, then it may be possible to use our techniques to analyse aggregation rules for selecting the beliefs and the supporting evidence for the group. As a second example for a

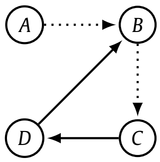

Fig. 5. Scenario used in the proof of Lemma 5.

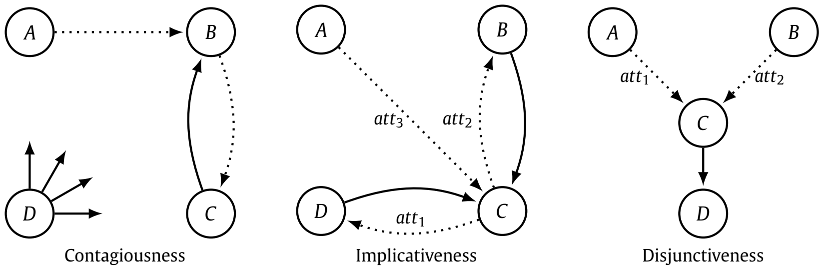

Fig. 6. Scenarios used in the proof of Theorem 7.

promising scenario of application, it would be interesting to investigate the *strategic incentives* of agents who are reporting an argumentation framework to an aggregation rule and whose objective might be to get a certain argument accepted.[21] This problem is similar to the problem of strategic manipulation in voting [52]. Recall that it is well-known that strategyproofness is closely linked to the independence axiom in voting [23], meaning that the insights regarding independent aggregation rules collected in this paper may well be of direct relevance to such an investigation.

Appendix. Remaining proofs

In this appendix we present the proofs omitted from the body of the paper. They all have the same structure: they show that a given collection of semantic AF-properties of interest has certain meta-properties, which allows us to apply certain more general results. Only the first proof for each meta-property is included in the body of the paper.

Proof of Lemma 5 (neutrality lemma for argument acceptability)

Suppose $|Arg| \geqslant 4$. Let $\mathcal{P}$ be the collection of AF-properties representing either credulous or sceptical argument acceptability under either the grounded, the ideal, the complete, the preferred, the semi-stable, or the stable semantics. The claim follows from Lemma 2 if we can show that $\mathcal{P}$ is contagious. So consider any three arguments $A, B, C \in Arg$. As $|Arg| \geqslant 4$, we can pick a fourth argument $D \in Arg$. Let $Att = \{C \rightharpoonup D, D \rightharpoonup B\}$ and consider the four argumentation frameworks of the form $\langle Arg, Att \cup S\rangle$ with $S \subseteq \{A \rightharpoonup B, B \rightharpoonup C\}$. They are indicated in Fig. 5. Now focus on the acceptability of argument C. If $S = \emptyset$ or $S = \{A \rightharpoonup B\}$, then C is not attacked by any other argument and thus a member of the grounded extension and thereby credulously and sceptically accepted under all six semantics. If $S = \{A \rightharpoonup B, B \rightharpoonup C\}$, then C is defended by A, which is not attacked by any other argument. Hence, C is a member of the grounded extension and thereby credulously and sceptically accepted under all six semantics also in this case. If $S = \{B \rightharpoonup C\}$, on the other hand, we obtain an isolated odd-length cycle including C, which means that C is not part of any extension under any of the six semantics. Hence, both credulous and sceptical acceptability of C is an AF-property of the kind we require, under all of the six semantics. □

Proof of Theorem 7 (extension preservation for the grounded and ideal semantics)

Suppose $|Arg| \geqslant 4$. Let $\mathcal{P}$ be either the collection of AF-properties representing a given set of arguments being the grounded extension or that of a given set of arguments being the ideal extension. We need to show that $\mathcal{P}$ is contagious, implicative, and disjunctive in both cases.

Contagiousness. Consider any four arguments $A, B, C, D \in Arg$. Suppose we are interested in the property of $\{A, C, D\}$ being the grounded or the ideal extension. We define Att as follows:

$$Att = \{C \rightharpoonup B\} \cup \{D \rightharpoonup X \mid X \in Arg \setminus \{A, B, C, D\}\}$$

[21] Prior work has dealt with a related but different question connecting abstract argumentation and game theory [51]: We are given a fixed argumentation framework. Every agent can report a subset of the available arguments. We then restrict the given argumentation framework to the union of the sets of arguments reported by the agents and apply an argumentation semantics to select an extension. What are the incentives of an individual agent to not report an argument she in fact is aware of?

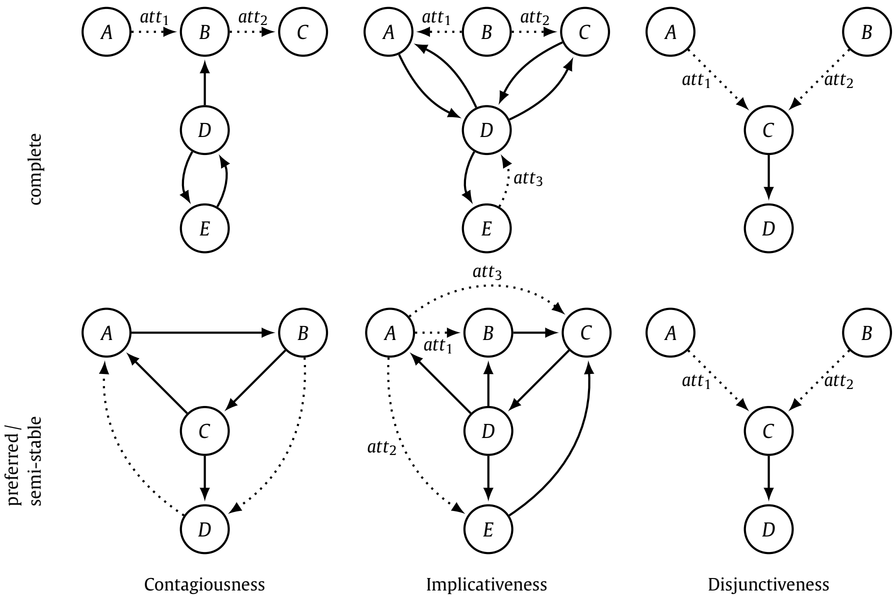

Fig. 7. Scenarios used in the proof of Theorem 8.

Consider the argumentation frameworks of the form $\langle Arg, Att \cup S\rangle$ with $S \subseteq \{A \rightharpoonup B, B \rightharpoonup C\}$. They are indicated in the leftmost part of Fig. 6. The reader may verify that for $S \neq \{B \rightharpoonup C\}$, both the grounded and the ideal extension indeed are $\{A, C, D\}$: these are the arguments that are either successfully defended or not attacked at all, while B is not. On the other hand, for $S = \{B \rightharpoonup C\}$, both the grounded and the ideal extension are $\{A, D\}$. Thus, $\mathcal{P}$ is contagious.

Implicativeness. Let $Arg = \{A, B, C, D, \ldots\}$. We focus on $Arg \setminus \{C\}$ as the subset of arguments that may (or may not) form the grounded or the ideal extension. We define $Att = \{B \rightharpoonup C, D \rightharpoonup C\}$, $att_1 = (C \rightharpoonup D)$, $att_2 = (C \rightharpoonup B)$, and $att_3 = (A \rightharpoonup C)$. This scenario is depicted in the middle part of Fig. 6. Consider all argumentation frameworks of the form $AF = \langle Arg, Att \cup S\rangle$ with $S \subseteq \{att_1, att_2, att_3\}$. The reader may verify that, indeed, for $S \neq \{att_1, att_2\}$ both the grounded and the ideal extension are equal to $Arg \setminus \{C\}$. On the other hand, for $S = \{att_1, att_2\}$ both of them are equal to $Arg \setminus \{B, C, D\}$. Thus, $\mathcal{P}$ is implicative.

Disjunctiveness. Let $Arg = \{A, B, C, D, \ldots\}$. We again focus on $Arg \setminus \{C\}$ as the subset of arguments that may (or may not) form the grounded or the ideal extension. We define $Att = \{C \rightharpoonup D\}$, $att_1 = (A \rightharpoonup C)$, and $att_2 = (B \rightharpoonup C)$. This is shown on the righthand side of Fig. 6. Consider argumentation frameworks $AF = \langle Arg, Att \cup S\rangle$ with $S \subseteq \{att_1, att_2\}$. As the reader can easily verify, if $S \neq \emptyset$, then $Arg \setminus \{C\}$ is both the grounded and ideal extension of AF. But if $S = \emptyset$, then $Arg \setminus \{D\}$ is both the grounded and the ideal extension. Thus, $\mathcal{P}$ is disjunctive. $\square$

Proof of Theorem 8 (extension preservation for the complete, preferred, and semi-stable semantics)

Suppose $|Arg| \geqslant 5$. Let $\mathcal{P}$ be the collection of AF-properties representing a given set of arguments being an extension under either the complete, the preferred, or the semi-stable semantics. We need to show that $\mathcal{P}$ is contagious, implicative, and disjunctive in all three cases.

Contagiousness. We first consider the case of complete extensions. Consider any five arguments $A, B, C, D, E \in Arg$. Let $Arg \setminus \{B, D, E\}$ be the set of arguments of interest. We define $Att = \{D \rightharpoonup B, D \rightharpoonup E, E \rightharpoonup D\}$. Now consider the argumentation frameworks of the form $\langle Arg, Att \cup S\rangle$ with $S \subseteq \{A \rightharpoonup B, B \rightharpoonup C\}$. This scenario is depicted in the upper lefthand corner of Fig. 7. If $S \neq \{B \rightharpoonup C\}$, then $Arg \setminus \{B, D, E\}$ is a complete extension. But if $S = \{B \rightharpoonup C\}$, then it is not, because C is attacked but not defended by A or any other argument in the set. Thus, being a complete extension is contagious.

Next, we consider the case of preferred and semi-stable extensions, where contagiousness can be established even for $|Arg| \geqslant 4$. So consider four arguments $A, B, C, D, \in Arg$. Let $Arg \setminus \{A, B, C, D\}$ be the set of interest, i.e., we ask whether it is possible that none of our four distinguished arguments will get accepted. We define $Att = \{A \rightharpoonup B, B \rightharpoonup C, C \rightharpoonup A, C \rightharpoonup D\}$. Now consider the argumentation frameworks of the form $\langle Arg, Att \cup S\rangle$ with $S \subseteq \{B \rightharpoonup D, D \rightharpoonup A\}$. This scenario is shown in the lower lefthand corner of Fig. 7. If $S \neq \{D \rightharpoonup A\}$, then no subset of $\{A, B, C, D\}$ other than $\emptyset$ is admissible. Hence, in these three cases, $Arg \setminus \{A, B, C, D\}$ is the only preferred extension, and thus also the only semi-stable extension. However, for $S = \{D \rightharpoonup A\}$ the set $Arg \setminus \{A, B, C, D\}$ is not a preferred extension (and thus also not a semi-stable extension), because its superset $Arg \setminus \{A, C\}$ is admissible as well. Thus, being either a preferred or a semi-stable extension is contagious.

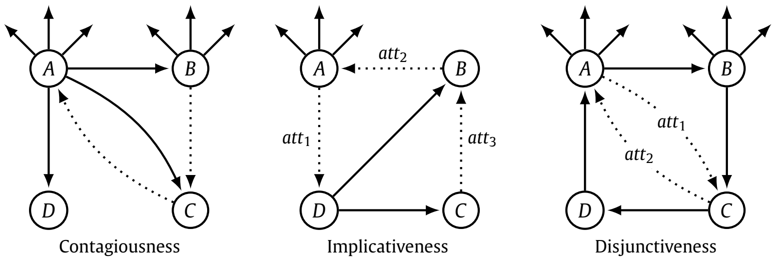

Fig. 8. Scenarios used in the proof of Theorem 10.

Implicativeness. Let $Arg = \{A, B, C, D, E, \ldots\}$. We again start with the case of complete extensions. We focus on the set $Arg \setminus \{A, C, D, E\}$ as a possible complete extension. Define $Att = \{A \rightharpoonup D, D \rightharpoonup A, C \rightharpoonup D, D \rightharpoonup C, D \rightharpoonup E\}$, $att_1 = (B \rightharpoonup A)$, $att_2 = (B \rightharpoonup C)$, and $att_3 = (E \rightharpoonup D)$. This scenario is shown in the middle of the top row in Fig. 7. Now consider the eight argumentation frameworks of the form $\langle Arg, Att \cup S\rangle$ with $S \subseteq \{att_1, att_2, att_3\}$. In all eight cases, B is part of every complete extension, because it is not attacked. If $S \neq \{att_1, att_2\}$, then (as far as our five distinguished arguments are concerned) it is possible to accept only B, i.e., $Arg \setminus \{A, C, D, E\}$ is a complete extension. But for $S = \{att_1, att_2\}$, we also must accept D, because it is successfully defended by B. Hence, being a complete extension is implicative.

Next, we turn to the preferred and the semi-stable semantics. Let $Arg \setminus \{A, B, C, D, E\}$ be the set of arguments under consideration. Define $Att = \{B \rightharpoonup C, D \rightharpoonup A, D \rightharpoonup B, D \rightharpoonup E, C \rightharpoonup D, E \rightharpoonup C\}$, $att_1 = (A \rightharpoonup B)$, $att_2 = (A \rightharpoonup E)$, and $att_3 = (A \rightharpoonup C)$. This situation is sketched in the middle of the bottom row in Fig. 7. We again consider the eight argumentation frameworks $\langle Arg, Att \cup S\rangle$ with $S \subseteq \{att_1, att_2, att_3\}$. First, consider the seven argumentation frameworks with $S \neq \{att_1, att_2\}$. The reader may verify that none of the nonempty and conflict-free subsets of $\{A, B, C, D, E\}$ is admissible. Hence, $Arg \setminus \{A, B, C, D, E\}$ is the only preferred extension (and thus also the only semi-stable extension) for any of these seven argumentation frameworks. On the other hand, if $S = \{att_1, att_2\}$, then $\{A, C\}$ is admissible and thus $Arg \setminus \{A, B, C, D, E\}$ cannot be either preferred or semi-stable. Hence, both being a preferred extension and being a semi-stable extension are implicative collections of AF-properties.

Disjunctiveness. To prove disjunctiveness of $\mathcal{P}$ we can use the same construction as in the proof of Theorem 7 in all three cases. To see that this is possible, it is sufficient to observe that the argumentation frameworks used in the proof are all acyclic, i.e., all our semantics coincide with the grounded semantics. Observe that this means that disjunctiveness holds even for $|Arg| \geqslant 4$. $\square$

Proof of Theorem 10 (preservation of nonemptiness of the ideal extension)

Suppose $|Arg| \geqslant 4$. We need to show that nonemptiness of the ideal extension is an AF-property that is contagious, implicative, and disjunctive.[22]

Contagiousness. Fix four arguments $A, B, C, D \in Arg$. We define the set Att of attacks as follows:

$$\begin{aligned} Att = {} & \{A \rightharpoonup B, A \rightharpoonup C, A \rightharpoonup D\} \cup \\ & \{A \rightharpoonup X \mid X \in Arg \setminus \{A, B, C, D\}\} \cup \\ & \{B \rightharpoonup X \mid X \in Arg \setminus \{A, B, C, D\}\} \end{aligned}$$

Now consider the four argumentation frameworks $\langle Arg, Att \cup S\rangle$ with $S \subseteq \{B \rightharpoonup C, C \rightharpoonup A\}$. This scenario is depicted in the leftmost part of Fig. 8. If $S \neq \{C \rightharpoonup A\}$, then the only preferred extension is $\{A\}$. Hence, in these cases, the ideal extension is $\{A\}$ as well and thus nonempty. But if $S = \{C \rightharpoonup A\}$, then there are two preferred extensions, namely $\{A\}$ and $\{B, C, D\}$. As their intersection is empty, the ideal extension must be empty as well. Thus, nonemptiness of the ideal extension is a contagious AF-property.

Implicativeness. Let $Arg = \{A, B, C, D, \ldots\}$. We define Att as follows:

$$Att = \{D \rightharpoonup B, D \rightharpoonup C\} \cup \{A \rightharpoonup X \mid X \in Arg \setminus \{A, B, C, D\}\}$$

Furthermore, let $att_1 = (A \rightharpoonup D)$, $att_2 = (B \rightharpoonup A)$, and $att_3 = (C \rightharpoonup B)$. This scenario is shown in the middle part of Fig. 8. Now consider the eight argumentation frameworks of the form $\langle Arg, Att \cup S\rangle$ with $S \subseteq \{att_1, att_2, att_3\}$. Whenever $S \neq \{att_1, att_2\}$, there is only a single preferred extension and A is part of it. Hence, the ideal extension is a superset of $\{A\}$

[22] Note that this is the first time we are proving this for a single AF-property rather than a collection of properties.

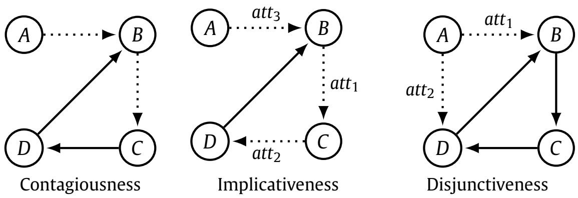

Fig. 9. Scenarios used in the proof of Theorem 14.

and thus nonempty. But if $S = \{att_1, att_2\}$, then the only preferred extension is the empty set and thus the ideal extension is empty as well. Hence, nonemptiness of the ideal extension is an implicative AF-property.

Disjunctiveness. Let $Arg = \{A, B, C, D, \ldots\}$. We define Att as follows:

$$\begin{aligned} Att = \{&A \rightharpoonup B, B \rightharpoonup C, C \rightharpoonup D, D \rightharpoonup A\} \;\cup \\ &\{A \rightharpoonup X \mid X \in Arg \setminus \{A, B, C, D\}\} \;\cup \\ &\{B \rightharpoonup X \mid X \in Arg \setminus \{A, B, C, D\}\} \end{aligned}$$

Furthermore, let $att_1 = (A \rightharpoonup C)$ and $att_2 = (C \rightharpoonup A)$. This scenario is shown on the righthand side of Fig. 8. Now consider the four argumentation frameworks of the form $\langle Arg, Att \cup S\rangle$ with $S \subseteq \{att_1, att_2\}$. If $S \neq \emptyset$, then the only preferred extension is $\{B, D\}$, which is also the ideal extension. On the other hand, for $S = \emptyset$ the preferred extensions are $\{A, C\}$ and $\{B, D\}$, meaning that the ideal extension is empty. Thus, nonemptiness of the ideal extension is a disjunctive AF-property. □

Proof of Theorem 13 (preservation of acyclicity)

The claim holds vacuously for $|Arg| = 1$. So, w.l.o.g., let us assume that $|Arg| > 1$. Acyclicity is an AF-property that is k-exclusive for every $k \in \{2, \ldots, |Arg|\}$. To see this, consider the case where the attack relations $\{att_1, \ldots, att_k\}$ form a cycle, and observe that the shortest (proper) cycle has length 2, while the longest cycle visits every vertex exactly once and thus has length $|Arg|$. The claim now follows from Theorem 11. □

Proof of Theorem 14 (preservation of coherence)

Recall that an argumentation framework is coherent if all its preferred extensions are also stable extensions. Suppose $|Arg| \geqslant 4$. We need to show that coherence is an AF-property that is contagious, implicative, and disjunctive.

Contagiousness. Fix four arguments $A, B, C, D \in Arg$. Let $Att = \{C \rightharpoonup D, D \rightharpoonup B\}$. Now consider the argumentation frameworks of the form $\langle Arg, Att \cup S\rangle$ with $S \subseteq \{A \rightharpoonup B, B \rightharpoonup C\}$. This scenario is depicted on the lefthand side of Fig. 9. If $S = \emptyset$, then $Arg \setminus \{D\}$ is the only preferred and the only stable extension. If $S = \{A \rightharpoonup B\}$ or $S = \{A \rightharpoonup B, B \rightharpoonup C\}$, then $Arg \setminus \{B, D\}$ is the only preferred and the only stable extension. Hence, each of these three argumentation frameworks is coherent. On the other hand, if $S = \{B \rightharpoonup C\}$, then $\{B, C, D\}$ forms an isolated odd-length cycle. Then $Arg \setminus \{B, C, D\}$ is the only preferred extension, which however is not stable. Hence, this argumentation framework is not coherent. In conclusion, we have shown that coherence is a contagious AF-property.

Implicativeness. Let $Arg = \{A, B, C, D, \ldots\}$. We define $Att = \{D \rightharpoonup B\}$, $att_1 = (B \rightharpoonup C)$, $att_2 = (C \rightharpoonup D)$, and $att_3 = (A \rightharpoonup B)$. This scenario is shown in the middle of Fig. 9 and is identical to the scenario used in the proof of Theorem 6. Now consider the eight argumentation frameworks $\langle Arg, Att \cup S\rangle$ with $S \subseteq \{att_1, att_2, att_3\}$. If $S \subseteq \{att_1, att_3\}$, then the only preferred extension is $Arg \setminus \{B\}$, which is also stable. If $S = \{att_2\}$, then the only preferred extension is $Arg \setminus \{D\}$, which again is also stable. If $S = \{att_2, att_3\}$ or $S = \{att_1, att_2, att_3\}$, then the only preferred extension is $Arg \setminus \{B, D\}$, which once again also is stable. Thus, in all seven cases we obtain coherent argumentation frameworks. However, if $S = \{att_1, att_2\}$, then the only preferred extension is $Arg \setminus \{B, C, D\}$, which is not stable. So in this case, coherence is violated. Hence, coherence is an implicative AF-property.

Disjunctiveness. Let $Arg = \{A, B, C, D, \ldots\}$. We define $Att = \{B \rightharpoonup C, C \rightharpoonup D, D \rightharpoonup B\}$, $att_1 = (A \rightharpoonup B)$, and $att_2 = (A \rightharpoonup D)$. This scenario is shown on the righthand side of Fig. 9. Now consider the four argumentation frameworks $\langle Arg, Att \cup S\rangle$ with $S \subseteq \{att_1, att_2\}$. If $S = \{att_1\}$ or $S = \{att_1, att_2\}$, then the only preferred extension is $Arg \setminus \{B, D\}$, which is also stable. If $S = \{att_2\}$, then the only preferred extension is $Arg \setminus \{C, D\}$, which again is also stable. Thus, in all three cases every preferred extension is stable. On the other hand, if $S = \emptyset$, then the only preferred extension is $Arg \setminus \{B, C, D\}$, which is not stable. Hence, coherence is a disjunctive AF-property. □

References

[1] W. Chen, U. Endriss, Preservation of semantic properties during the aggregation of abstract argumentation frameworks, in: Proceedings of the 16th Conference on Theoretical Aspects of Rationality and Knowledge, TARK, 2017, pp. 118–133.
[2] P.E. Dunne, T.J.M. Bench-Capon (Eds.), Proceedings of the 1st International Conference on Computational Models of Argument, COMMA, IOS Press, 2006.
[3] T.J.M. Bench-Capon, P.E. Dunne, Argumentation in artificial intelligence, Artif. Intell. 171 (2007) 619–641.
[4] P. Besnard, A. Hunter, Elements of Argumentation, MIT Press, 2008.
[5] I. Rahwan, G.R. Simari (Eds.), Argumentation in Artificial Intelligence, Springer, 2009.
[6] S. Modgil, N. Oren, F. Toni (Eds.), Proceedings of the 1st International Workshop on Theory and Applications of Formal Argumentation, TAFA, Springer, 2012.
[7] P.M. Dung, On the acceptability of arguments and its fundamental role in nonmonotonic reasoning, logic programming and n-person games, Artif. Intell. 77 (1995) 321–358.
[8] S. Coste-Marquis, C. Devred, S. Konieczny, M.-C. Lagasquie-Schiex, P. Marquis, On the merging of Dung's argumentation systems, Artif. Intell. 171 (2007) 730–753.
[9] F.A. Tohmé, G.A. Bodanza, G.R. Simari, Aggregation of attack relations: a social-choice theoretical analysis of defeasibility criteria, in: Proceedings of the 5th International Symposium on Foundations of Information and Knowledge Systems, FoIKS, Springer, 2008, pp. 8–23.
[10] M. Caminada, G. Pigozzi, On judgment aggregation in abstract argumentation, J. Auton. Agents Multiagent Syst. 22 (2011) 64–102.
[11] I. Rahwan, F.A. Tohmé, Collective argument evaluation as judgement aggregation, in: Proceedings of the 9th International Conference on Autonomous Agents and Multiagent Systems, AAMAS, IFAAMAS, 2010, pp. 417–424.
[12] J. Leite, J. Martins, Social abstract argumentation, in: Proceedings of the 22nd International Joint Conference on Artificial Intelligence, IJCAI, 2011, pp. 2287–2292.
[13] P.E. Dunne, P. Marquis, M. Wooldridge, Argument aggregation: basic axioms and complexity results, in: Proceedings of the 4th International Conference on Computational Models of Argument, COMMA, IOS Press, 2012, pp. 129–140.
[14] J. Delobelle, S. Konieczny, S. Vesic, On the aggregation of argumentation frameworks, in: Proceedings of the 24th International Joint Conference on Artificial Intelligence, IJCAI, 2015, pp. 2911–2917.
[15] S. Airiau, E. Bonzon, U. Endriss, N. Maudet, J. Rossit, Rationalisation of profiles of abstract argumentation frameworks: characterisation and complexity, J. Artif. Intell. Res. 60 (2017) 149–177.
[16] G.A. Bodanza, F.A. Tohmé, M.R. Auday, Collective argumentation: a survey of aggregation issues around argumentation frameworks, Argum. Comput. 8 (2017) 1–34.
[17] W. Chen, U. Endriss, Aggregating alternative extensions of abstract argumentation frameworks: preservation results for quota rules, in: Proceedings of the 7th International Conference on Computational Models of Argument, COMMA, IOS Press, 2018, pp. 425–436.
[18] K.J. Arrow, A.K. Sen, K. Suzumura (Eds.), Handbook of Social Choice and Welfare, North-Holland, 2002.
[19] W. Gaertner, A Primer in Social Choice Theory, LSE Perspectives in Economic Analysis, Oxford University Press, 2006.
[20] F. Brandt, V. Conitzer, U. Endriss, J. Lang, A.D. Procaccia (Eds.), Handbook of Computational Social Choice, Cambridge University Press, 2016.
[21] A.K. Sen, Social choice theory, in: K.J. Arrow, M.D. Intriligator (Eds.), Handbook of Mathematical Economics, vol. 3, North-Holland, 1986, pp. 1073–1181.
[22] W. Thomson, On the axiomatic method and its recent applications to game theory and resource allocation, Soc. Choice Welf. 18 (2001) 327–386.
[23] U. Endriss, Logic and social choice theory, in: A. Gupta, J. van Benthem (Eds.), Logic and Philosophy Today, vol. 2, College Publications, 2011, pp. 333–377.
[24] U. Endriss, U. Grandi, Graph aggregation, Artif. Intell. 245 (2017) 86–114.
[25] G.A. Bodanza, M.R. Auday, Social argument justification: some mechanisms and conditions for their coincidence, in: Proceedings of the 10th European Conference on Symbolic and Quantitative Approaches to Reasoning with Uncertainty, ECSQARU, Springer, 2009, pp. 95–106.
[26] D.M. Gabbay, O. Rodrigues, An equational approach to the merging of argumentation networks, J. Log. Comput. 24 (2013) 1253–1277.
[27] J. Delobelle, A. Haret, S. Konieczny, J.-G. Mailly, J. Rossit, S. Woltran, Merging of abstract argumentation frameworks, in: Proceedings of the 15th International Conference on Principles of Knowledge Representation and Reasoning, KR, 2016, pp. 33–42.
[28] N. Li, A paradox of expert rights in abstract argumentation, Soc. Choice Welf. 51 (2018) 737–752.
[29] A.K. Sen, The impossibility of a Paretian liberal, J. Polit. Econ. 78 (1970) 152–157.
[30] E. Bonzon, N. Maudet, On the outcomes of multiparty persuasion, in: Proceedings of the 10th International Conference on Autonomous Agents and Multiagent Systems, AAMAS, IFAAMAS, 2011, pp. 47–54.
[31] S. Eğilmez, J. Martins, J. Leite, Extending social abstract argumentation with votes on attacks, in: Proceedings of the 2nd International Workshop on Theory and Applications of Formal Argumentation, TAFA, Springer, 2014, pp. 16–31.
[32] M. Caminada, L. Amgoud, On the evaluation of argumentation formalisms, Artif. Intell. 171 (2007) 286–310.
[33] H. Prakken, An abstract framework for argumentation with structured arguments, Argum. Comput. 1 (2010) 93–124.
[34] S. Modgil, Revisiting abstract argumentation frameworks, in: Proceedings of the 2nd International Workshop on Theory and Applications of Formal Argumentation, TAFA, Springer, 2014, pp. 1–15.
[35] M. Caminada, Semi-stable semantics, in: Proceedings of the 1st International Conference on Computational Model of Argument, COMMA, IOS Press, 2006, pp. 121–130.
[36] P.M. Dung, P. Mancarella, F. Toni, Computing ideal sceptical argumentation, Artif. Intell. 171 (2007) 642–674.
[37] P. Baroni, M. Caminada, M. Giacomin, An introduction to argumentation semantics, Knowl. Eng. Rev. 26 (2011) 365–410.
[38] S. Bistarelli, F. Santini, ConArg: a constraint-based computational framework for argumentation systems, in: Proceedings of the 23rd IEEE International Conference on Tools with Artificial Intelligence, ICTAI, IEEE, 2011, pp. 605–612.
[39] D. Grossi, G. Pigozzi, Judgment Aggregation: A Primer, Synthesis Lectures on Artificial Intelligence and Machine Learning, Morgan & Claypool Publishers, 2014.
[40] F. Dietrich, C. List, Judgment aggregation by quota rules: majority voting generalized, J. Theor. Polit. 19 (2007) 391–424.
[41] K.J. Arrow, Social Choice and Individual Values, 2nd ed., John Wiley and Sons, 1963, first edition published in 1951.
[42] C. List, P. Pettit, Aggregating sets of judgments: an impossibility result, Econ. Philos. 18 (2002) 89–110.
[43] I. McLean, A.B. Urken (Eds.), Classics of Social Choice, University of Michigan Press, 1995.
[44] S. Konieczny, J. Lang, P. Marquis, DA^2 merging operators, Artif. Intell. 157 (2004) 49–79.
[45] E. Hemaspaandra, H. Spakowski, J. Vogel, The complexity of Kemeny elections, Theor. Comput. Sci. 349 (2005) 382–391.
[46] U. Endriss, U. Grandi, D. Porello, Complexity of judgment aggregation, J. Artif. Intell. Res. 45 (2012) 481–514.
[47] C. Cayrol, M. Lagasquie-Schiex, Bipolar abstract argumentation systems, in: I. Rahwan, G.R. Simari (Eds.), Argumentation in Artificial Intelligence, Springer, 2009, pp. 65–84.
[48] T.J.M. Bench-Capon, Persuasion in practical argument using value-based argumentation frameworks, J. Log. Comput. 13 (2003) 429–448.
[49] C. Geist, D. Peters, Computer-aided methods for social choice theory, in: U. Endriss (Ed.), Trends in Computational Social Choice, AI Access, 2017, pp. 249–267.

[50] C. Shi, S. Smets, F.R. Velázquez-Quesada, Argument-based belief in topological structures, in: Proceedings of the 16th Conference on Theoretical Aspects of Rationality and Knowledge, TARK, 2017, pp. 489–503.
[51] I. Rahwan, K. Larson, Argumentation and game theory, in: I. Rahwan, G.R. Simari (Eds.), Argumentation in Artificial Intelligence, Springer, 2009, pp. 321–339.
[52] A.D. Taylor, Social Choice and the Mathematics of Manipulation, Cambridge University Press, 2005.