

Judgment Aggregation in Abstract Dialectical Frameworks

Richard Booth

Université du Luxembourg
6 rue Richard Coudenhove-Kalergi, Luxembourg
ribooth@gmail.com

Abstract. Abstract dialectical frameworks (ADFs) are a knowledge representation formalism introduced as a generalisation of Dung’s abstract argumentation frameworks (AFs) by Gerhard Brewka and co-authors. We look at a judgment aggregation problem in ADFs, namely the problem of aggregating a profile of complete interpretations. We generalise the family of *interval aggregation methods*, studied in the AF case in our previous work, to the ADF case. Along the way we define the notions of *down-admissible* and *up-complete* interpretations, that were already previously defined for the AF case by Caminada and Pigozzi. These aggregation methods may open the way to define interesting new semantics for ADFs, such as a generalisation to the ADF case of the *ideal* semantics for AFs.

Keywords: Abstract dialectical frameworks, argumentation frameworks, judgment aggregation, interval methods

1 Introduction

Abstract Dialectical Frameworks (ADFs) [5, 4] have recently been introduced by Gerhard Brewka and colleagues as a knowledge representation formalism that generalises the popular Abstract Argumentation Frameworks (AFs) introduced by Dung [9]. Several different *semantics* for ADFs have been defined which each provide a way to map any ADF to its set of *models* or *interpretations*. Usually these different semantics are defined so as to generalise an existing semantics of AFs, such as *admissible* and *complete* semantics. In this paper we likewise seek to explore within the wider setting of ADFs a problem that has recently received attention in the more confined setting of AFs. Namely we are interested in the problem of *aggregation opinions* in ADFs.

Our aggregation problem can be described as follows. Suppose we have a group of agents who share a given ADF D . Each agent has a particular opinion about the truth-value of the statements in D , and we assume each agent is *rational* in the sense that their interpretation is a model of D according to some commonly agreed semantics (which, in this paper, will be the complete semantics). The question is, can we aggregate these opinions into a single group interpretation that represents the opinion of the group as a whole? In the AF

setting, this question (which can be thought of as a special case of the problem of *judgment aggregation* [11, 12]) has been investigated in [3]. There, a general family of methods for aggregating complete AF labellings called *interval methods* was defined and studied. In this paper we will show how this family can be extended to cover the case of aggregation in ADFs.

Although simple to define and understand, the interval methods of [3] suffered from the rather severe drawback that they were not guaranteed to output a rational (i.e., complete) labelling for every possible AF. To remedy this, it was suggested to add a post-processing *repair* step to the output. This step consisted of two sub-procedures, the *down-admissible* followed by the *up-complete* procedures, which were first introduced in [7]. We will see that these procedures, that were so far only defined in the AF setting, can be generalised to the ADF case, thus yielding a family of ADF aggregation operators - the *DAUC interval methods* - that guarantee a rational outcome.

The plan of this paper is as follows. We start in the next section by giving preliminary background on ADFs. Then in Section 3 we formally introduce the ADF aggregation problem and give some postulates for aggregation methods. In Section 4 we define interval aggregation methods and axiomatically characterise them. Then in Section 5 we present the down-admissible and up-complete procedures for ADFs and define the DAUC interval methods. Lastly we conclude, including a discussion about possible applications for defining new semantics for ADFs, in particular a counterpart of ideal semantics from AFs.

2 Abstract Dialectical Frameworks

Abstract dialectical frameworks were first introduced by Brewka and Woltran [5] and then further developed by Brewka et al [4] as a useful generalisation of Dung's abstract argumentation frameworks [9]. The idea is that we have a collection S of atomic *statements* (essentially just propositional atoms), and each $s \in S$ has a propositional formula φ_s associated to it that intuitively represents the *justification* for accepting s . Roughly, if we have enough grounds for holding φ_s to be true, then that gives us license to hold s to be true. The following setup of ADFs is based on [4] (which was, in turn, inspired by the algebraic approach to non-monotonic reasoning from [8] - see also [13] for a comprehensive study of algebraic semantics for ADFs).

To begin, we assume a countable universal set U of possible *statements* from which all ADFs are formed. Then an ADF may be defined as follows:

Definition 1. An abstract dialectical framework is a pair $D = (S, C)$ where

- $S \subseteq U$ is a finite set of statements.
- C is a collection $\{\varphi_s\}_{s \in S}$ of propositional formulas built from S . Formula φ_s is called the condition of s .

We will sometimes (if D is not clear from the context) denote the sets of statements and conditions of an ADF D by S_D and C_D respectively.

For simplicity we assume for each $s \in S$ that φ_s doesn't contain redundant statements, that is, for each statement t appearing in φ_s , the set of (2-valued)

interpretations satisfying $\varphi_s(t/\mathbf{t})$ differs from the set of (2-valued) interpretations satisfying $\varphi_s(t/\mathbf{f})$, where $\varphi_s(t/\mathbf{t})$ denotes the formula resulting from substituting t everywhere by \mathbf{t} , and similarly for $\varphi_s(t/\mathbf{f})$, where \mathbf{t} and \mathbf{f} denote propositional truth and falsity respectively. That is, the truth or falsity of t can make a difference to the truth value of φ_s .

An ADF comes with an implicit graph structure, with S as the nodes, reflecting the dependencies between statements. To each ADF D we associate the set of links L_D by setting $(t, s) \in L_D$ iff $s, t \in S_D$ and t appears in φ_s . We denote the set of parents of s by $\text{Par}_D(s)$, i.e., $\text{Par}_D(s) = \{t \in S_D \mid (t, s) \in L_D\}$. Dung argumentation frameworks (AFs) form a subclass of the class of ADFs. Indeed an ADF D is an AF iff each φ_s is equivalent to $\bigwedge_{t \in \text{Par}_D(s)} \neg t$.

Example 1. Following [4], we may represent an ADF by listing its statements, with the condition of each statement written in square brackets immediately after it. For example one possible ADF D_0 with $S_{D_0} = \{a, b, c\}$ may be written as follows.

$$a \ [\mathbf{t}], \quad b \ [b], \quad c \ [-a \vee b].$$

A D -interpretation is just a function $v : S \rightarrow \{\mathbf{t}, \mathbf{f}, \mathbf{u}\}$ assigning one of the truth-values \mathbf{t} (true), \mathbf{f} (false) or \mathbf{u} (unknown) to each statement in S . For notational convenience we define a negation operator on the set of truth-values by setting $\neg \mathbf{t} = \mathbf{f}$, $\neg \mathbf{f} = \mathbf{t}$ and $\neg \mathbf{u} = \mathbf{u}$. The ordering \sqsubseteq between D -interpretations is defined by $v_1 \sqsubseteq v_2$ iff for all $s \in S$, $v_1(s) \leq v_2(s)$, where \leq is the (reflexive) information ordering between truth-values given by $\mathbf{u} \leq \mathbf{t}$ and $\mathbf{u} \leq \mathbf{f}$ (and no other pair (\mathbf{x}, \mathbf{y}) for $\mathbf{x} \neq \mathbf{y}$ is in \leq). The set of truth-values forms a complete meet-semi-lattice¹ under \leq , with meet operation \sqcap behaving as a ‘‘consensus’’ operator, i.e., $\mathbf{t} \sqcap \mathbf{t} = \mathbf{t}$, $\mathbf{f} \sqcap \mathbf{f} = \mathbf{f}$, and the meet of all other pairs returning \mathbf{u} . The set of all D -interpretations, equipped with ordering \sqsubseteq , inherits this semi-lattice structure with meet operation \sqcap defined by $(v_1 \sqcap v_2)(s) = v_1(s) \sqcap v_2(s)$.

We say a D -interpretation is *2-valued* if it assigns only values in $\{\mathbf{t}, \mathbf{f}\}$. Given a D -interpretation v , the set of 2-valued interpretations \sqsubseteq -extending v is denoted by $[v]_2$. Then we define a function Γ_D taking D -interpretations to D -interpretations as follows by setting, for all $s \in S_D$:

$$[\Gamma_D(v)](s) = \bigsqcap \{w(\varphi_s) \mid w \in [v]_2\}.$$

That is, to determine $[\Gamma_D(v)](s)$ we look at all possible ways the 3-valued interpretation v may be completed to a 2-valued one. If there is consensus on the value of the condition φ_s among all of these then $[\Gamma_D(v)](s)$ is set to that value. Otherwise $[\Gamma_D(v)](s) = \mathbf{u}$. An alternative formulation of Γ_D can be given as follows. For each D -interpretation v let v^\wedge denote the conjunction

¹ A complete meet-semi-lattice is such that (i) every non-empty finite subset has a greatest lower bound, and (ii) every non-empty directed subset has a least upper bound, where a subset X is directed if any two elements of X have an upper bound in X .

$\bigwedge\{s \mid v(s) = \mathbf{t}\} \wedge \bigwedge\{\neg s \mid v(s) = \mathbf{f}\}$. Then, for each $s \in S_D$

$$[\Gamma_D(v)](s) = \begin{cases} \mathbf{t} & \text{if } v^\wedge \models \varphi_s \\ \mathbf{f} & \text{if } v^\wedge \models \neg\varphi_s \\ \mathbf{u} & \text{otherwise} \end{cases}$$

where \models denotes entailment in classical propositional logic. It is shown in [5] that Γ_D is monotonic in \sqsubseteq , i.e., if $v_1 \sqsubseteq v_2$ then $\Gamma_D(v_1) \sqsubseteq \Gamma_D(v_2)$.

The notions of *admissible* and *complete* D -interpretations can then be defined in terms of the Γ_D -function as follows.

Definition 2. *A D -interpretation is admissible iff $v \sqsubseteq \Gamma_D(v)$. It is complete iff $v = \Gamma_D(v)$.*

Intuitively, a D -interpretation is admissible if it doesn't assign \mathbf{t} or \mathbf{f} to any statement s without *justification* for doing so. An admissible D -interpretation is complete if it assigns \mathbf{t} or \mathbf{f} to every statement for which justification is at hand. As a fixed point of Γ_D , a complete D -interpretation can be thought of a *rational, internally coherent* belief state regarding the truth or falsity of the statements in S_D . For the special case in which D is an AF, these notions coincide with the notions of admissible and complete *argument labellings* of [6] (see also [1]).

Example 2. Consider ADF D_0 from Example 1. We can write D_0 -interpretations as triples (p, q, r) of truth-values expressing the values of a, b, c in that order. There are three possible complete D_0 -interpretations: $v_1 = (\mathbf{t}, \mathbf{t}, \mathbf{t})$, $v_2 = (\mathbf{t}, \mathbf{u}, \mathbf{u})$ and $v_3 = (\mathbf{t}, \mathbf{f}, \mathbf{f})$. An example of an interpretation which is admissible but not complete is $v_4 = (\mathbf{u}, \mathbf{t}, \mathbf{t})$. The all-unknown interpretation - in this case $v_5 = (\mathbf{u}, \mathbf{u}, \mathbf{u})$ - is always admissible.

The notions of admissible or complete D -interpretations provide just two possible *semantics* for ADFs. Others are possible (see [4, 13]) but for this paper we will focus on only these.

3 Aggregating complete interpretations: postulates

The aggregation setting we have in mind is as follows. We assume a fixed set $Ag = \{1, \dots, n\}$ of *agents* (for some fixed $n \geq 2$). The idea is that, given some arbitrary ADF D , each agent forms some opinion over the truth or falsity of each statement, subject to the constraints encoded in D . Each agent i 's opinion is expressed as a complete D -interpretation v_i , and they are collected in a *D -profile* $\mathbf{v} = (v_1, \dots, v_n)$. For any $T \subseteq S_D$ and D -interpretation v we denote by $v[T]$ the projection of v to just the statements in T , and we denote by $\mathbf{v}[T]$ the n -tuple $(v_1[T], \dots, v_n[T])$. We would like to determine a *single* D -interpretation that reflects the opinion of the group as a whole.

Definition 3. *An ADF aggregation method (hereafter aggregation method for short) is a function \mathcal{F} that assigns, to each ADF D and each profile \mathbf{v} of complete D -interpretations, a D -interpretation $\mathcal{F}_D(\mathbf{v})$.*

How should we define a good aggregation method? Before describing some concrete families of such methods in the next sections, we take a look at a few desirable postulates for aggregation methods. These are inspired by and appropriately modified from postulates studied in the AF case in [3] (which, in turn, have been mostly inspired by postulates from the judgment aggregation literature [11, 12]). Free variables in these postulates, e.g., D , \mathbf{v} in the first three postulates below, are implicitly universally quantified.

The first, basic, group of postulates is as follows:

Collective Completeness $\mathcal{F}_D(\mathbf{v})$ is a complete D -interpretation.

Anonymity If \mathbf{v}' is a permutation of \mathbf{v} then $\mathcal{F}_D(\mathbf{v}') = \mathcal{F}_D(\mathbf{v})$.

Unanimity If $v_i = v$ for all $i \in Ag$ then $\mathcal{F}_D(\mathbf{v}) = v$.

Collective Completeness requires that the output of the aggregation should be a rational interpretation. For *Anonymity* we say “ \mathbf{v}' is a permutation of \mathbf{v} ” to mean that $\mathbf{v} = (v_1, \dots, v_n)$ and $\mathbf{v}' = (v_{\sigma(1)}, \dots, v_{\sigma(n)})$ for some permutation σ on Ag . Thus this postulate says the identity of the agents does not matter in the aggregation process. *Unanimity* says that if all agents agree on the same D -interpretation then this should be the output.

Although a basic requirement, *Collective Completeness* will turn out not to be satisfied by the family of aggregation methods we present in the next section. However, restricting it to a particularly simple kind of ADF - in fact a kind of AF - brings it to within much easier reach. We say an AF D is a *2-loop AF* iff $S_D = \{s, t\}$ and $L_D = \{(s, t), (t, s)\}$ for some distinct $s, t \in U$.

Minimal Collective Completeness If D is a 2-loop AF, then $\mathcal{F}_D(\mathbf{v})$ is a complete D -interpretation.

The next two postulates try to ensure minimum levels of *satisfaction* for the agents with the collective outcome. Given a tuple $(\mathbf{x}_i)_{i \in Ag}$ of truth-values the *t/f-winner* (resp. *t/f-loser*) in $(\mathbf{x}_i)_{i \in Ag}$ is that value among $\{\mathbf{t}, \mathbf{f}\}$ which appears more (resp. less) often in $(\mathbf{x}_i)_{i \in Ag}$. For example the *t/f-loser* in $(\mathbf{t}, \mathbf{u}, \mathbf{u}, \mathbf{f}, \mathbf{t})$ is \mathbf{f} .

t/f-Plurality If \mathbf{x} is the *t/f-loser* in $(v_i(s))_{i \in Ag}$ then $[\mathcal{F}_D(\mathbf{v})](s) \neq \mathbf{x}$.

Compatibility $v_i(s) = \neg[\mathcal{F}_D(\mathbf{v})](s)$ implies $v_i(s) = \mathbf{u}$.

t/f-Plurality thus says the collective value assigned to s cannot be $\mathbf{x} \in \{\mathbf{t}, \mathbf{f}\}$ if strictly more agents voted for it to be $\neg\mathbf{x}$. *Compatibility* is a stronger property that says the collective value cannot be $\mathbf{x} \in \{\mathbf{t}, \mathbf{f}\}$ if *at least one agent* voted for it to be $\neg\mathbf{x}$. The postulate is so-called because it says that the collective interpretation $\mathcal{F}_D(\mathbf{v})$ must be *compatible* with the interpretation of *every* agent i .

Definition 4. Two D -interpretations u, v are compatible iff there is no statement s such that $u(s) = \neg v(s) \neq \mathbf{u}$.

This notion of compatibility plays a leading role in the AF aggregation setting of Caminada and Pigozzi [7]. The following lemma regarding the interplay between the notions of compatibility and completeness will be used in the proof of Proposition 5 in Section 5.

Lemma 1. *Let u be a complete D -interpretation. Then, for any D -interpretation v , if v is compatible with u then so is $\Gamma_D(v)$.*

Proof. Suppose u is complete but $\Gamma_D(v)$ is not compatible with u . Then there must be some $s \in S_D$ such that $[\Gamma_D(v)](s) = \neg u(s) \neq \mathbf{u}$. By completeness of u this gives $[I_D(v)](s) = \neg[I_D(u)](s) \neq \mathbf{u}$, which implies $v^\wedge \wedge u^\wedge \models \perp$. But this can only happen if $v(t) = \neg u(t) \neq \mathbf{u}$ for some $t \in S_D$, i.e., if v is not compatible with u . \square

For the next aggregation postulate, we say that a given truth-value \mathbf{y} is *between* truth-values \mathbf{x} and \mathbf{z} iff $\mathbf{y} = \mathbf{x}$ or $\mathbf{y} = \mathbf{z}$ or $[\mathbf{y} = \mathbf{u}$ and $\mathbf{x} = \neg\mathbf{z}]$. The next postulate implies that if a particular collective outcome $\mathbf{x} \in \{\mathbf{t}, \mathbf{f}\}$ is obtained for statement s , and if some of the agents then change the truth-value they assign to s so that they move *closer* to this collective outcome, then the collective outcome does not change.

Monotonicity Let \mathbf{v}, \mathbf{v}' be D -profiles such that for all $s \in S_D$ and all $i \in Ag$, if $v_i(s) \neq v'_i(s)$ then $([\mathcal{F}_D(\mathbf{v})](s) \in \{\mathbf{t}, \mathbf{f}\}$ and $v'_i(s)$ is between $v_i(s)$ and $[\mathcal{F}_D(\mathbf{v})](s)$). Then $\mathcal{F}_D(\mathbf{v}') = \mathcal{F}_D(\mathbf{v})$.

All the postulates until now referred to only a single ADF D . The remaining postulates deal with restricting the behaviour of \mathcal{F} across different, but related, ADFs. The first enforces a certain *neutrality* over the *names* of the statements used in an ADF. Given two ADFs $D = (S, C)$ and $D' = (S', C')$ a *renaming* from D to D' is a bijection $\tau : S \rightarrow S'$ such that $\varphi_{\tau(s)} = \tau(\varphi_s)$ for each $s \in S$, where $\tau(\varphi_s)$ is obtained from φ_s by replacing every occurrence of each statement t with $\tau(t)$. A renaming lifts to a function taking D -interpretations v to D' -interpretations $\tau(v)$ by taking $[\tau(v)](s) = v(\tau^{-1}(s))$ for each $s \in S'$. As the following result shows, both the admissible and complete semantics for ADFs are invariant under renaming.²

Proposition 1. *Let τ be a renaming from D to D' . If v is an admissible, resp. complete, D -interpretation then $\tau(v)$ is an admissible, resp. complete, D' -interpretation.*

Proof. (Outline) Follows mainly from that fact that, for each $s \in S_D$ and any D -interpretation v we have $v^\wedge \models \varphi_s$ iff $\tau(v)^\wedge \models \varphi_{\tau(s)}$, and $v^\wedge \models \neg\varphi_s$ iff $\tau(v)^\wedge \models \neg\varphi_{\tau(s)}$. \square

A renaming τ further extends naturally to a mapping from D -profiles \mathbf{v} to D' -profiles $\tau(\mathbf{v}) = (\tau(v_1), \dots, \tau(v_n))$. The *Renaming* postulate for aggregation methods can then be formalised as follows:

² In the AF setting, this is known as the *language independence* property of AF semantics [2].

Renaming If τ is a renaming from D to D' then $\tau(\mathcal{F}_D(\mathbf{v})) = \mathcal{F}_{D'}(\tau(\mathbf{v}))$.

We remark that in the restricted case of AFs this postulate can be simplified so that it talks about graph isomorphisms rather than renamings (see the *Isomorphism* postulate in [3]).

The next postulate is a strong version of the *Independence* postulate which forms the basis of several important results (especially impossibility results) in judgment aggregation. It says that the collective truth-value of s depends at most on the tuple of individual truth-values assigned to s by the agents, regardless of which other statements may or may not be present in D .

ADF-Independence If \mathbf{v} is a D -profile and \mathbf{v}' is a D' -profile and $s \in S_D \cap S_{D'}$, then $\mathbf{v}[s] = \mathbf{v}'[s]$ implies $[\mathcal{F}_D(\mathbf{v})](s) = [\mathcal{F}_{D'}(\mathbf{v}')](s)$.

As expected, this postulate turns out to be too strong, and anyway could be argued against on the basis that it asks us to disregard dependency information between statements (in the form of the set L_D) that is explicitly submitted as part of the input to the problem. We thus formulate a weaker version, inspired by the *Directionality* property for AF semantics [2]. Given an ADF $D = (S, \{\varphi_s\}_{s \in S})$ and $T \subseteq S$, we say T is *primary in D* if for no $t \in T, s \in S \setminus T$ do we have $(s, t) \in L_D$, i.e., each statement in T depends only on statements within T . We denote by $D \downarrow T$ the ADF $(T, \{\varphi_s\}_{s \in T})$, with the φ_s “inherited” from D . Then the *Directionality* postulate for ADF aggregation can be formulated as follows:

Directionality If $T \subseteq S_D$ is primary in D then $\mathcal{F}_{D \downarrow T}(\mathbf{v}[T]) = \mathcal{F}_D(\mathbf{v})[T]$.

(Note one can straightforwardly show that if v is admissible, resp. complete, in D and T is primary in D , then $v[T]$ is admissible, resp. complete, in $D \downarrow T$.) This property says that the outcome of aggregation for a primary set T of statements is independent of statements outside the set.

4 Interval aggregation methods

We now describe a family of ADF aggregation methods, generalised from the family of AF aggregation methods from [3] known as *interval methods*. Let Int_n denote the set of non-zero *intervals* over $\{0, 1, \dots, n\}$, i.e., $Int_n = \{(k, l) \mid k, l \in \{0, 1, \dots, n\} \text{ and } k < l\}$. Then to define a member of this family, we just choose some distinguished set $Y \subseteq Int_n$. We say Y is *widening*³ if $(a, b) \in Y$ whenever $(k, l) \in Y$ and $a \leq k, l \leq b$, and is *zero-based* if $k = 0$ whenever $(k, l) \in Y$. Each possible choice of Y yields an aggregation method \mathcal{F}^Y by setting, for each ADF D , D -profile \mathbf{v} and $s \in S_D$:

$$[\mathcal{F}_D^Y(\mathbf{v})](s) = \begin{cases} \mathbf{x} & \text{if } \mathbf{x} \in \{\mathbf{t}, \mathbf{f}\} \text{ and } (|N_{s: \neg \mathbf{x}}^{\mathbf{v}}|, |N_{s: \mathbf{x}}^{\mathbf{v}}|) \in Y. \\ \mathbf{u} & \text{otherwise} \end{cases}$$

³ The widening interval methods are very closely related to *quota systems* studied in voting theory by Young et al [14].

Here, $N_{s:\mathbf{x}}^{\mathbf{v}}$ denotes the set of agents who assign value \mathbf{x} to s in \mathbf{v} , i.e., $N_{s:\mathbf{x}}^{\mathbf{v}} = \{i \in Ag \mid v_i(s) = \mathbf{x}\}$. Thus $\mathcal{F}_D^Y(\mathbf{v})$ sets the collective truth-value of s to be the \mathbf{t}/\mathbf{f} -winner \mathbf{x} in $(v_i(s))_{i \in Ag}$ provided such a winner exists and $(|N_{s:\neg\mathbf{x}}^{\mathbf{v}}|, |N_{s:\mathbf{x}}^{\mathbf{v}}|) \in Y$. Otherwise the collective value is set to \mathbf{u} .

Definition 5. *An aggregation method \mathcal{F} will be called an interval method iff $\mathcal{F} = \mathcal{F}^Y$ for some $Y \subseteq Int_n$ such that $(0, n) \in Y$. If, furthermore, Y is widening, resp. zero-based, then we say \mathcal{F} is a widening, resp. zero-based, interval method.*

The restriction $(0, n) \in Y$ is essentially made to ensure \mathcal{F}^Y satisfies the *Unanimity* postulate.

The family of interval methods contains a number of interesting special cases. We mention three here, the first two of which were first studied in the AF case in [7]:

- Sceptical: $Y^{\text{Scep}} = \{(0, n)\}$. Take the collective value of a statement s to be \mathbf{x} if *all* agents voted for \mathbf{x} , otherwise take \mathbf{u} . We use $\mathcal{F}^{\text{Scep}}$ to denote $\mathcal{F}^{Y^{\text{Scep}}}$. Note that $\mathcal{F}^{\text{Scep}}(\mathbf{v}) = \prod_{i \in Ag} v_i$.
- Credulous: $Y^{\text{Cred}} = \{(0, l) \mid l \geq 1\}$. Take the collective value to be $\mathbf{x} \in \{\mathbf{t}, \mathbf{f}\}$ if at least one agent voted for \mathbf{x} and none voted for the opposite $\neg\mathbf{x}$. Otherwise take \mathbf{u} . We use $\mathcal{F}^{\text{Cred}}$ to denote $\mathcal{F}^{Y^{\text{Cred}}}$.
- Simple majority: $Y^{\text{SMaj}} = Int_n$. Here we just take the \mathbf{t}/\mathbf{f} -winner whenever it exists, and take \mathbf{u} otherwise. We use $\mathcal{F}^{\text{SMaj}}$ to denote $\mathcal{F}^{Y^{\text{SMaj}}}$.

Notice that all three of Y^{Scep} , Y^{Cred} and Y^{SMaj} are widening, and all except Y^{SMaj} are zero-based.

Example 3. Consider the following ADF D_1 with $S_{D_1} = \{a, b, c, d\}$.

$$a [a], \quad b [b], \quad c [c], \quad d [\neg a \wedge \neg b \wedge \neg c].$$

Assume $n = 4$ and that $\mathbf{v} = (v_1, v_2, v_3, v_4)$ with $v_1 = (\mathbf{t}, \mathbf{t}, \mathbf{t}, \mathbf{f})$, $v_2 = (\mathbf{f}, \mathbf{t}, \mathbf{u}, \mathbf{f})$, $v_3 = (\mathbf{t}, \mathbf{t}, \mathbf{u}, \mathbf{f})$, $v_4 = (\mathbf{u}, \mathbf{f}, \mathbf{u}, \mathbf{u})$. Then $\mathcal{F}^{\text{Scep}}(\mathbf{v}) = (\mathbf{u}, \mathbf{u}, \mathbf{u}, \mathbf{u})$, $\mathcal{F}^{\text{Cred}}(\mathbf{v}) = (\mathbf{u}, \mathbf{u}, \mathbf{t}, \mathbf{f})$ and $\mathcal{F}^{\text{SMaj}}(\mathbf{v}) = (\mathbf{t}, \mathbf{t}, \mathbf{t}, \mathbf{f})$.

We can characterise the family of interval methods in terms of postulates as follows.

Theorem 1. *Let \mathcal{F} be an aggregation method. Then \mathcal{F} is an interval method iff it satisfies Anonymity, Unanimity, Minimal Collective Completeness, \mathbf{t}/\mathbf{f} -Plurality, Renaming and ADF-Independence.*

Proof. (Outline). Soundness is relatively straightforward. The completeness part largely follows the same pattern as the proof of the corresponding theorem for the restricted AF case in [3]. We first show how to construct, from any given aggregation method \mathcal{F} , a subset $Y(\mathcal{F}) \subseteq Int_n$: Let D_0 be a 2-loop AF such that $S_{D_0} = \{a_0, b_0\}$. There are three complete D_0 -interpretations, which we denote by $v_{\mathbf{t}}$, $v_{\mathbf{f}}$ and $v_{\mathbf{u}}$, where the subscript represents the value of a_0 (with the label of b_0 of course being always $\neg v(a_0)$). Then we define $Y(\mathcal{F})$ by setting

$Y(\mathcal{F}) = \{(k, l) \in \text{Int}_n \mid [\mathcal{F}_{D_0}(\mathbf{v}_{k,l})](a_0) = \mathbf{t}\}$, where $\mathbf{v}_{k,l}$ is any D_0 -profile such that precisely k agents provide labelling $v_{\mathbf{f}}$ and l agents provide $v_{\mathbf{t}}$. Note by *Anonymity* that the precise distribution of labellings among $\mathbf{v}_{k,l}$ doesn't matter. $Y(\mathcal{F})$ is well-defined, i.e., it doesn't matter which 2-loop AF we take to define it (by *Renaming*) and $(0, n) \in Y(\mathcal{F})$ (by *Unanimity*). One can then show that \mathcal{F} and $\mathcal{F}^{Y(\mathcal{F})}$ agree on the 2-loop AF D_0 , i.e., that for every D_0 -profile \mathbf{v} we have $\mathcal{F}_{D_0}(\mathbf{v}) = \mathcal{F}_{D_0}^{Y(\mathcal{F})}(\mathbf{v})$. This part depends on *Anonymity*, *Renaming*, *Minimal Collective Completeness* and *t/f-Plurality*. Then finally we extend this to hold for any ADF D using *ADF-Independence* and *Renaming*. \square

Regarding the other postulates mentioned in the previous section, *Directionality* is satisfied by every interval method, since it is a direct weakening of *ADF-Independence*, but it can be shown that none of the remaining three, i.e., *Collective Completeness*, *Compatibility* and *Monotonicity* is satisfied in general. The last two, however, can be obtained by adding restrictions on Y .

Proposition 2. *Let \mathcal{F}^Y be an interval method.*

- (i) \mathcal{F}^Y satisfies *Monotonicity* iff Y is widening.
- (ii) \mathcal{F}^Y satisfies *Compatibility* iff Y is zero-based.

(The proof of this is straightforward. We remark that the “only if” directions of these two results are essentially already covered by the analogous results proved for the AF case in [3].) As a corollary we see that all three of our example interval methods $\mathcal{F}^{\text{Scep}}$, $\mathcal{F}^{\text{Cred}}$ and $\mathcal{F}^{\text{SMaj}}$ satisfy *Monotonicity*, while all except $\mathcal{F}^{\text{SMaj}}$ satisfy *Compatibility*.

What about *Collective Completeness*? Readers familiar with the judgment aggregation literature will not be surprised to learn that there is *no* interval method satisfying *Collective Completeness*, even if we restrict to AFs, as shown in [3].

Theorem 2 ([3]). *There is no aggregation method (for any $n > 1$) satisfying all of Anonymity, Unanimity, Renaming, ADF-Independence and Collective Completeness.*

One response to this result in the AF case which was followed in [3] (thereby generalising the approach of [7] who focussed only on the special cases $\mathcal{F}^{\text{Scep}}$ and $\mathcal{F}^{\text{Cred}}$) was to give up *ADF-Independence* by applying the *down-admissible* and *up-complete* procedures to the outcome of aggregation. We will follow this route here.

5 Down-admissible and up-complete procedures

The purpose of the down-admissible procedure in [7] was to take any AF-labelling and to revise it *downwards* (along the ordering \sqsubseteq), just enough so that it becomes an admissible labelling. It turns out that this procedure quite easily generalises to the ADF setting. Suppose we start from any given D -interpretation v . We

then iteratively construct a sequence v_0, v_1, \dots of D -interpretations by setting $v_0 = v$ and $v_{i+1} = v_i \sqcap \Gamma_D(v_i)$ for $i \geq 0$. Clearly $v_{i+1} \sqsubseteq v_i$ for each i . Let $a = \min\{i \mid v_{i+1} = v_i\}$. By the finiteness of S_D a is guaranteed to exist.

Definition 6. *Let D be an ADF and v a D -interpretation. The down-admissible interpretation of v , denoted by $\downarrow v$ is defined by $\downarrow v = v_a$, where the sequence v_0, v_1, \dots, v_a is defined as above.*

Since $v_a = v_a \sqcap \Gamma_D(v_a)$ we have $v_a \sqsubseteq \Gamma_D(v_a)$ and so $\downarrow v$ is admissible. In fact it is the largest admissible D -interpretation that is \sqsubseteq -smaller than v , as the next result confirms.

Proposition 3. *Let v' be an admissible D -interpretation such that $v' \sqsubseteq v$. Then $v' \sqsubseteq \downarrow v$.*

Proof. We show by induction on i that $v' \sqsubseteq v_i$ for all $i = 0, 1, \dots, a$ in the above procedure. Since $v_0 = v$ the base case $i = 0$ holds by assumption. So now assume $v' \sqsubseteq v_i$. We will show also $v' \sqsubseteq v_{i+1}$, i.e., $v' \sqsubseteq (v_i \sqcap \Gamma_D(v_i))$. To show this it is enough to show both $v' \sqsubseteq v_i$ and $v' \sqsubseteq \Gamma_D(v_i)$. The first of these holds by inductive hypothesis. For the second, we know $v' \sqsubseteq \Gamma_D(v')$ from the assumption that v' is admissible. We also know $\Gamma_D(v') \sqsubseteq \Gamma_D(v_i)$ from $v' \sqsubseteq v_i$ and the monotonicity of Γ_D . From these two we conclude $v' \sqsubseteq \Gamma_D(v_i)$ as required. \square

Example 4. Let D_2 be the following ADF, with $S_{D_2} = \{a, b, c, d\}$.

$$a [a], \quad b [b], \quad c [\neg a \vee b], \quad d [c],$$

and consider the D_2 -interpretation $v = v_0 = (\mathbf{t}, \mathbf{f}, \mathbf{t}, \mathbf{t})$. We have $\Gamma_{D_2}(v_0) = (\mathbf{t}, \mathbf{f}, \mathbf{f}, \mathbf{t})$, so $v_1 = v_0 \sqcap \Gamma_{D_2}(v_0) = (\mathbf{t}, \mathbf{f}, \mathbf{u}, \mathbf{t})$. Now $\Gamma_{D_2}(v_1) = (\mathbf{t}, \mathbf{f}, \mathbf{f}, \mathbf{u})$, so $v_2 = (\mathbf{t}, \mathbf{f}, \mathbf{u}, \mathbf{u})$. Since v_2 is admissible, the procedure stops here with $\downarrow v = v_2 = (\mathbf{t}, \mathbf{f}, \mathbf{u}, \mathbf{u})$.

As the previous example shows, the down-admissible interpretation of v need not be a complete D -interpretation. In [7] the purpose of the up-complete procedure was to take any admissible AF-labelling and to revise it *upwards* (along \sqsubseteq), just enough so that it becomes a complete labelling. As with the down-admissible procedure, it is relatively straightforward to generalise this procedure to the ADF case. Starting with an admissible D -interpretation v we can construct a sequence $v = v_0, v_1, v_2, \dots$ of D -interpretations by setting $v_{i+1} = \Gamma_D(v_i)$ for $i \geq 0$. From the assumption that v is admissible (so $v_0 \sqsubseteq v_1$) and the monotonicity of Γ_D we have $v_i \sqsubseteq v_{i+1}$ for $i \geq 0$. Let $c = \min\{i \mid v_{i+1} = v_i\}$. Again, by finiteness c is guaranteed to exist.

Definition 7. *Let D be an ADF and v an admissible D -interpretation. The up-complete interpretation of v , denoted by $\uparrow v$ is defined by $\uparrow v = v_c$, where the sequence $v = v_0, v_1, \dots, v_c$ is defined as above.*

Clearly v_c is complete, and the next result confirms it to be the smallest complete D -interpretation that is \sqsubseteq -larger than v .

Proposition 4. *Let v' be a complete D -interpretation such that $v \sqsubseteq v'$. Then $\uparrow v \sqsubseteq v'$.*

Proof. We show by induction on i that $v_i \sqsubseteq v'$. Since $v_0 = v$ the base case $i = 0$ holds by assumption. So now assume $v_i \sqsubseteq v'$. Then, by monotonicity of Γ_D we know $\Gamma_D(v_i) \sqsubseteq \Gamma_D(v')$, i.e., $v_{i+1} \sqsubseteq \Gamma_D(v')$. But since $v' = \Gamma_D(v')$ by the assumption that v' is complete, this gives us $v_{i+1} \sqsubseteq v_i$ as required. \square

Example 5. Let us continue Example 4. Let $v_0 = \downarrow v = (\mathbf{t}, \mathbf{f}, \mathbf{u}, \mathbf{u})$. Then $v_1 = \Gamma_{D_2}(v_0) = (\mathbf{t}, \mathbf{f}, \mathbf{f}, \mathbf{u})$, $v_2 = \Gamma_{D_2}(v_1) = (\mathbf{t}, \mathbf{f}, \mathbf{f}, \mathbf{f}) = \Gamma_{D_2}(v_2)$. Thus $\uparrow(\downarrow v) = (\mathbf{t}, \mathbf{f}, \mathbf{f}, \mathbf{f})$.

We denote the composite operation $\uparrow(\downarrow v)$ of taking the down-admissible followed by the up-complete interpretations of v by $\uparrow\downarrow v$. Taken in combination, these procedures provide a way of transforming any aggregation method into one that is guaranteed to satisfy *Collective Completeness*.

Definition 8. *Let \mathcal{F} be an aggregation method. The DAUC version of \mathcal{F} , denoted by $\widehat{\mathcal{F}}$, is defined by setting, for any ADF D and D -profile \mathbf{v} , $\widehat{\mathcal{F}}_D(\mathbf{v}) = \uparrow\downarrow(\mathcal{F}_D(\mathbf{v}))$.*

What can we say about the properties of $\widehat{\mathcal{F}}$, other than *Collective Completeness*? The next proposition gives us some other properties of $\widehat{\mathcal{F}}$, provided that \mathcal{F} satisfies them.

Proposition 5. *Let \mathcal{F} be an aggregation method. For each of the following postulates, if \mathcal{F} satisfies that postulate then so does $\widehat{\mathcal{F}}$: Anonymity, Unanimity, Renaming, Compatibility and Directionality.*

Proof. (Outline) The proofs for *Anonymity* and *Unanimity* are straightforward. *Renaming* is preserved mainly due to the fact that, for any D -interpretation v and renaming τ (to some D') we have $v^\wedge \models \varphi_s$ iff $\tau(v)^\wedge \models \varphi_{\tau(s)}$, and $v^\wedge \models \neg\varphi_s$ iff $\tau(v)^\wedge \models \neg\varphi_{\tau(s)}$. For *Compatibility* suppose \mathcal{F} satisfies that postulate. Then for every D, \mathbf{v} , $\mathcal{F}_D(\mathbf{v})$ is compatible with every agent's interpretation v_i . Since $\downarrow\mathcal{F}_D(\mathbf{v}) \sqsubseteq \mathcal{F}_D(\mathbf{v})$ we know $\downarrow\mathcal{F}_D(\mathbf{v})$ must also be compatible with every v_i . By Lemma 1 this compatibility is then preserved at each step of the up-complete procedure for $\uparrow(\downarrow\mathcal{F}_D(\mathbf{v}))$ and so finally $\widehat{\mathcal{F}}_D(\mathbf{v})$ is also compatible. Finally, *Directionality* is preserved due to that fact that if t is a statement not appearing in φ_s , then $v^\wedge \models \varphi_s$ iff $v^\wedge \models \varphi_s$ (and similarly for $\neg\varphi_s$), where v^\wedge is the same as v^\wedge but with any literal t or $\neg t$ removed. \square

Combined with the results of the previous section, this gives us a list of sound postulates for the DAUC versions of the interval methods.

Corollary 1. *Let \mathcal{F} be an interval method. Then $\widehat{\mathcal{F}}$ satisfies Collective Completeness, Anonymity, Unanimity, Renaming and Directionality.*

Of course the DAUC interval methods do not satisfy *ADF-Independence* by Theorem 2. Regarding *t/f-plurality*, we know from Propositions 2 and 5 that if Y is zero-based then $\hat{\mathcal{F}}^Y$ satisfies *Compatibility* and hence also *t/f-plurality*. In fact it turns out that being zero-based is a *necessary* condition for $\hat{\mathcal{F}}^Y$ to satisfy *t/f-plurality*. This follows already from the analogous result proved for the AF case in [3].

Proposition 6 ([3]). *Let \mathcal{F}^Y be an interval method. Then $\hat{\mathcal{F}}^Y$ satisfies t/f-plurality iff Y is zero-based.*

Regarding *Monotonicity*, in view of Proposition 2 one might expect that a necessary and sufficient condition for $\hat{\mathcal{F}}^Y$ to satisfy that postulate is that Y is widening. However we have thus far been unable to prove or disprove this, and so it remains an open question for now.

6 Conclusion

We looked at the problem of defining methods for aggregation that take any profile of complete D -interpretations, over any given ADF D , and return a group D -interpretation. We showed that much of the same machinery used in the more specialised case of aggregating complete labellings of AFs can be applied to this problem. In particular we were able to define and axiomatically characterise a generalised version of the interval methods of [3], and to apply the down-admissible and up-complete procedures to transform the output of any interval method into a complete D -interpretation.

As noted in [7] for the AF case, one imaginative use for aggregation methods is as a route to define a (single-status) *semantics* for AFs. The role of an AF semantics is to prescribe, for every possible AF and for each argument a in the AF, which label represents the “common sense” label that a should be assigned in the context of that AF. One way to obtain this common sense labelling is to *aggregate* all possible rational labellings. In [7] this manoeuvre was carried out using both $\hat{\mathcal{F}}^{\text{Scep}}$ and $\hat{\mathcal{F}}^{\text{Cred}}$, and both were shown to correspond to some already existing semantics. Specifically, the result of aggregating all possible complete labellings⁴ using $\hat{\mathcal{F}}^{\text{Scep}}$ coincides with the *grounded labelling* [6, 9], while aggregating all possible complete labellings using $\hat{\mathcal{F}}^{\text{Cred}}$ results in the *ideal labelling* [7, 10]. This latter result is interesting, since until now there has been no generalised version of ideal semantics proposed for ADFs. The above discussion suggests that we can obtain such an ADF semantics by taking the output of the result of aggregating *all* complete D -interpretations using $\hat{\mathcal{F}}^{\text{Cred}}$. A related question is: what happens when we aggregate all complete D -interpretations using *other* members of the family of DAUC interval methods, such as $\hat{\mathcal{F}}^{\text{SMaj}}$. Does this give rise to other meaningful ADF semantics? These questions will be left for further study.

⁴ [7] also considered the results of aggregating other sets of labellings, such as all possible *preferred* labellings.

Another open question regards the axiomatic characterisation of the DAUC versions of the interval methods. In this paper we have managed to give a list of sound postulates for this family (Corollary 1). It remains to be proved that this list is complete. Finally, in this paper we have restricted ourselves to the problem of aggregating *complete* interpretations. It would be interesting to look at aggregation using other ADF semantics, such as those described in [4].

Acknowledgments

This work is supported by Fonds National de Recherche (DYNGBaT project). Thanks are due to Edmond Awad and Iyad Rahwan for earlier helping to develop some of the ideas for the restricted case of AFs, and to Johannes Wallner for some helpful comments.

References

1. Baroni, P., Caminada, M., Giacomin, M.: An introduction to argumentation semantics. *Knowledge Engineering Review* 26(4), 365–410 (2011)
2. Baroni, P., Giacomin, M.: On principle-based evaluation of extension-based argumentation semantics. *Artificial Intelligence* 171(10), 675–700 (2007)
3. Booth, R., Awad, E., Rahwan, I.: Interval methods for judgment aggregation in argumentation. In: *Proceedings of the 14th International Conference on Principles of Knowledge Representation and Reasoning (KR 2014)*. pp. 594–597 (2014)
4. Brewka, G., Ellmauthaler, S., Strass, H., Wallner, J.P., Woltran, S.: Abstract dialectical frameworks revisited. In: *Proceedings of the 23rd International Joint Conference on Artificial Intelligence (IJCAI 2013)*. pp. 803–809 (2013)
5. Brewka, G., Woltran, S.: Abstract dialectical frameworks. In: *Proceedings of the 12th International Conference on Principles of Knowledge Representation and Reasoning (KR 2010)*. pp. 102–111 (2010)
6. Caminada, M., Gabbay, D.: A logical account of formal argumentation. *Studia Logica* 93(2-3), 109–145 (2009)
7. Caminada, M., Pigozzi, G.: On judgment aggregation in abstract argumentation. *Autonomous Agents and Multi-Agent Systems* 22(1), 64–102 (2011)
8. Denecker, M., Marek, V.W., Truszczyński, M.: Ultimate approximation and its application in nonmonotonic knowledge representation systems. *Information and Computation* 192(1), 84–121 (2004)
9. Dung, P.M.: On the acceptability of arguments and its fundamental role in non-monotonic reasoning, logic programming and n -person games. *Artificial intelligence* 77(2), 321–357 (1995)
10. Dung, P.M., Mancarella, P., Toni, F.: Computing ideal sceptical argumentation. *Artificial Intelligence* 171(10), 642–674 (2007)
11. Grossi, D., Pigozzi, G.: *Judgment Aggregation: A Primer*. Morgan and Claypool (2014)
12. List, C., Puppe, C.: *Judgment aggregation: A survey*. In: *Handbook of Rational and Social Choice*. Oxford University Press (2009)
13. Strass, H.: Approximating operators and semantics for abstract dialectical frameworks. *Artificial Intelligence* 205, 39–70 (2013)
14. Young, S.C., Taylor, A.D., Zwicker, W.S.: Counting quota systems: A combinatorial question from social choice theory. *Mathematics Magazine* pp. 331–342 (1995)