Interval Methods for Judgment Aggregation in Argumentation

Richard Booth, Edmond Awad and Iyad Rahwan

Abstract

Given a set of conflicting arguments, there can exist multiple plausible opinions about which arguments should be accepted, rejected, or deemed undecided. Recent work explored some operators for deciding how multiple such judgments should be aggregated. Here, we generalize this line of study by introducing a family of operators called *interval aggregation methods*, and show that they contain existing operators as instances. While these methods fail to output a complete labelling in general, we show that it is possible to *transform* a given aggregation method into one that *does* always yield collectively rational labellings. This employs the *down-admissible* and *up-complete* constructions of Caminada and Pigozzi. For interval methods, collective rationality is attained at the expense of a strong *Independence* postulate, but we show that an interesting weakening of the Independence postulate is retained by this transformation. Our results open up many new directions in the study of generalized judgment aggregation in argumentation.[a]

[a]A short version of this paper appears in the proceedings of KR 2014.

1 Introduction

A conflicting knowledge base can be viewed abstractly as a set of arguments (defeasible derivations), and a binary relation capturing conflicts among them, forming an *argumentation framework* (AF) [11]. Given a set of conflicting arguments, there can exist multiple plausible ways to identify (or *label*) which arguments should be accepted, rejected, or deemed undecided [1]. The question we explore here is how to aggregate the judgments of multiple agents who have different opinions about how to evaluate a given set of arguments. This is akin to a jury who all have access to the same set of arguments presented in a legal case, but must aggregate their different opinions to a single decision.

This problem of *Judgment Aggregation* (JA) has been explored extensively in classical logic [14]. But it was only recently that JA has been applied to collective argument evaluation [7, 17]. Early results showed that argument-wise plurality voting cannot guarantee that the outcome of aggregation is always rational (consistent)—thus simple voting violates *Collective Rationality* [17]. On the other hand, the aggregation operators of Caminada and Pigozzi are able to guarantee collective rationality, but do so at the expense of the *Independence* property [7].

In the present paper, we embark on a broader study of JA in argumentation. We define a general family of aggregation operators called *interval methods* and show that they contain existing operators as instances. Interval methods always satisfy a strong version of Independence, but will usually fail Collective Rationality. But despite this important barrier, we are able to fully axiomatize interval methods in terms of a set of fundamental postulates.

We also define a sub-family of *widening* interval methods, and present an axiomatization in which it satisfies a weaker form of collective rationality. In particular, the output of widening methods can always be *completed* to a labelling, over a larger argumentation framework, that satisfies collective rationality. This finding suggests that failure to achieve collective rationality in general is not as bad as initially thought.

More importantly, building on Caminada and Pigozzi's *down-admissible + up-complete* (DAUC) construction, we present an approach to transform any interval method into one satisfying *Collective Rationality* while preserving a weaker and more reasonable form of independence known as *Directionality.*

We believe our results represent valuable contributions to the state-of-the-art for the following reasons. First, our definition of aggregation methods is very general: in contrast to previous approaches to JA in argumentation and logic, where the AF (or the *agenda*) is primarily fixed, our method allows the AF to vary. This allows us to formulate more general postulates that constrain the output of aggregation across *different* AFs (such as the *Isomorphism* postulate below). We believe this generalization will be immensely valuable beyond the present paper, since it allows the derivation of results that apply more broadly.

The second valuable aspect of our contribution is the axiomatization of very broad classes of aggregation operators, in terms of fundamental postulates that we map out precisely. These results ensure that any future refinements of these operators ensure minimal guaratees on their properties, and thus focus research on circumventing only the violated properties.

Finally, our approaches to transforming any interval aggregation method to one satisfying collective rationality form a foundation for much further work on JA in argumentation. They constitute generic procedures that can be combined with any method that can be shown to belong to the 'interval' family.

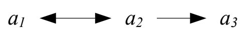
2 Preliminaries

We start by assuming a countably infinite set U of argument names, from which all possible argumentation frameworks are built. We restrict ourselves to *finite* argumentation frameworks.

Definition 1 *An* argumentation framework *(AF for short)* $\mathcal{A} = (Args, \rightharpoonup)$ *is a pair consisting of a finite set* $Args \subseteq U$ *of arguments and an* attack *relation* $\rightharpoonup \subseteq Args \times Args$*. Sometimes we use* $Args_{\mathcal{A}}$ *and* $\rightharpoonup_{\mathcal{A}}$ *to denote the arguments and attack relation of* $\mathcal{A}$*.*

An AF can be visualised as a directed graph, with nodes and edges representing arguments and attacks respectively.

Example 1 *The AF* $(\{a_1, a_2, a_3\}, \{(a_1, a_2), (a_2, a_1), (a_2, a_3)\}$ *can be pictured as follows.*



The setting for argumentation we consider in this paper is the abstract setting of [11]. Here, the notions of argument and attack are taken as primitive, with no internal structure assumed. This allows to focus on central questions of what makes an argument *acceptable* given its attacking arguments. In the argumentation literature once can find several concrete systems that form instantiations of this abstract framework, typically defining arguments as derivations of some claim from a set of premises (see, e.g., [3, 11, 12, 16]).

An AF is evaluated by assigning one of the labels `in`, `out` or `undec` to each argument in $Args_{\mathcal{A}}$, standing for *accepted*, *rejected* and *undecided* respectively [5, 6]. For notational convenience we define a unary "negation" operator on the set of labels by setting $\neg\mathtt{in} = \mathtt{out}$, $\neg\mathtt{out} = \mathtt{in}$ and $\neg\mathtt{undec} = \mathtt{undec}$. Given an AF $\mathcal{A}$, an $\mathcal{A}$*-labelling* is a function $L : Args_{\mathcal{A}} \to \{\mathtt{in}, \mathtt{out}, \mathtt{undec}\}$. For each $\mathtt{x} \in \{\mathtt{in}, \mathtt{out}, \mathtt{undec}\}$ we denote by $L^{-1}(\mathtt{x})$ the inverse image of $\mathtt{x}$ under L, and given $A \subseteq Args_{\mathcal{A}}$ we denote by $L[A]$ the restriction of L to A.

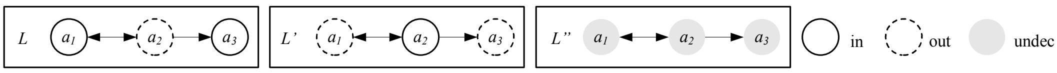
A major research question in abstract argumentation theory has been that of establishing when a given $\mathcal{A}$-labelling can be said to represent a *rational* evaluation of the arguments in $Args_{\mathcal{A}}$. Of course such an evaluation should somehow respect the attack relation. Several definitions, or so-called *argumentation semantics*, have been proposed. A fundamental concept is that of a *complete* labelling.

Definition 2 *Let $\mathcal{A} = (Args, \rightharpoonup)$ be an AF and L be an $\mathcal{A}$-labelling. We say L is a* complete *$\mathcal{A}$-labelling iff, for all $a \in Args$:*

- *If $L(a) = \texttt{in}$ then $L(b) = \texttt{out}$ for all $b \in Args$ s.t. $b \rightharpoonup a$.*
- *If $L(a) = \texttt{out}$ then $L(b) = \texttt{in}$ for some $b \in Args$ s.t. $b \rightharpoonup a$.*
- *If $L(a) = \texttt{undec}$ then $L(b) \neq \texttt{in}$ for all $b \in Args$ s.t. $b \rightharpoonup a$ and $L(c) = \texttt{undec}$ for some $c \in Args$ s.t. $c \rightharpoonup a$.*

We denote the set of complete $\mathcal{A}$-labellings by $Comp(\mathcal{A})$.

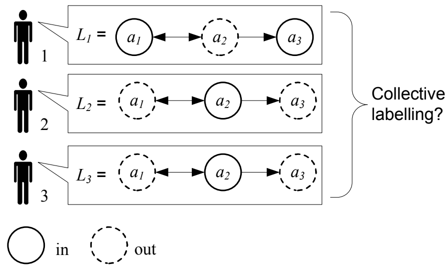
Example 2 *Consider the AF from Example 1. Then there are three possible complete labellings for this framework, which can be pictured as follows.*



Complete labellings form the basis of several other semantics such as *grounded*, *preferred*, *stable* [11]. Due to their fundamental nature and intuitiveness, we focus mainly on complete labellings in this paper. This is also in keeping with previous works on JA in abstract argumentation. We will, however, also use later the concept of *admissible* $\mathcal{A}$-labelling, which is an $\mathcal{A}$-labelling that satisfies the first two restrictions of Definition 2, but not necessarily the third. An example of an admissible $\mathcal{A}$-labelling in Example 2 would be L such that $L(a_1) = \texttt{in}$, $L(a_2) = \texttt{out}$ and $L(a_3) = \texttt{undec}$.

We assume a set of agents $Ag = \{1, \ldots, n\}$ (with $n \geqslant 2$) is fixed. An *$\mathcal{A}$-profile* is a sequence $\mathbf{L} = (L_1, \ldots, L_n)$ assigning a **complete** $\mathcal{A}$-labelling to each $i \in Ag$. Given $A \subseteq Args_{\mathcal{A}}$ we denote by $\mathbf{L}[A]$ the profile $(L_1[A], \ldots, L_n[A])$, i.e., the restriction of $\mathbf{L}$ to the arguments in A (writing $\mathbf{L}[a]$ rather than $\mathbf{L}[\{a\}]$ for singletons). For each label $\texttt{x} \in \{\texttt{in}, \texttt{out}, \texttt{undec}\}$ and $a \in Args_{\mathcal{A}}$ we denote the set of agents who *voted* for label $\texttt{x}$ for a by $V^{\mathbf{L}}_{a:\texttt{x}}$, i.e., $V^{\mathbf{L}}_{a:\texttt{x}} = \{i \in Ag \mid L_i(a) = \texttt{x}\}$.

Example 3 *Suppose $n = 3$ and let $\mathcal{A}$ be as in the previous example. Then one possible $\mathcal{A}$-profile $\mathbf{L}$ would be as follows:*



Here we have, for example, $V^{\mathbf{L}}_{a_1:\texttt{in}} = \{1\}$, $V^{\mathbf{L}}_{a_3:\texttt{out}} = \{2, 3\}$ and $V^{\mathbf{L}}_{a_2:\texttt{undec}} = \varnothing$.

What we seek in this paper is some method that, given *any* AF $\mathcal{A}$ and any $\mathcal{A}$-profile, will return another $\mathcal{A}$-labelling that represents the *collective labelling* of the group. Here is the central concept in this paper.

Definition 3 *An* aggregation method *is a function F that assigns to every AF $\mathcal{A}$ and $\mathcal{A}$-profile* $\mathbf{L}$ *an $\mathcal{A}$-labelling $F_{\mathcal{A}}(\mathbf{L})$.*

Note that an aggregation method takes as input **both** an AF $\mathcal{A}$ as well as an accompanying profile of complete $\mathcal{A}$-labellings. We emphasise, though, that in each particular aggregation situation, the agents aggregate labellings over a *single* AF that is shared by all agents. This contrasts with the argumentation aggregation operators of Coste-Marquis et al [9], whose input is a profile of *different* AFs (without labellings), one for each agent. Note also the above definition comes with an assumption of *universal domain* built in, i.e, an aggregation method must yield a result for *any* given AF $\mathcal{A}$ and $\mathcal{A}$-profile $\mathbf{L}$. Finally note that we require the output $F_{\mathcal{A}}(\mathbf{L})$ to be a *single* $\mathcal{A}$-labelling, as opposed to a *set* of $\mathcal{A}$-labellings.

3 Postulates for aggregation methods

We start by writing down some postulates for a good aggregation method. Some are inspired by postulates in [7, 17] (which in turn were inspired by postulates familiar from the JA literature), but we modify them to account for allowing the AF to vary. Note that free occurrences of $\mathcal{A}$ and $\mathbf{L}$ within the postulates are implicitly universally quantified.

In the definition of aggregation method the labellings of the agents are required to be complete, but the output labelling $F_{\mathcal{A}}(\mathbf{L})$ may be an arbitrary $\mathcal{A}$-labelling. Ideally, of course, we would like the output too to be complete.

> **Collective Rationality** $F_{\mathcal{A}}(\mathbf{L}) \in Comp(\mathcal{A})$.

Full *Collective Rationality* will turn out to be beyond the reach of the the simplest aggregation methods, including the family of interval methods in the next section. However, if we restrict to a particularly simple class of AFs it turns out to be relatively easy to satisfy. We call $\mathcal{A}$ a *2-loop AF* if it consists only of two arguments that mutually attack each other, i.e., $Args_{\mathcal{A}} = \{a, b\}$ and $\rightharpoonup_{\mathcal{A}} = \{(a, b), (b, a)\}$ for some distinct $a, b \in U$.

> **Minimal Collective Rationality** For any 2-loop AF $\mathcal{A}$ we have $F_{\mathcal{A}}(\mathbf{L}) \in Comp(\mathcal{A})$.

Some more weakenings of *Collective Rationality* will be considered in Section 5.

The next three postulates will be taken as fundamental in this paper. Given an $\mathcal{A}$-profile $\mathbf{L} = (L_1, \ldots, L_n)$, we say the $\mathcal{A}$-profile $\mathbf{L}'$ is a *permutation* of $\mathbf{L}$ if $\mathbf{L}' = (L_{\sigma(1)}, \ldots, L_{\sigma(n)})$ for some permutation σ on Ag.

> **Anonymity** If $\mathbf{L}'$ is a permutation of $\mathbf{L}$ then $F_{\mathcal{A}}(\mathbf{L}) = F_{\mathcal{A}}(\mathbf{L}')$.

Anonymity says the identity of which agent submitted which labelling does not matter when aggregating. The next rule says that if all agents submit exactly the same labelling then that should also be the collective labelling.

> **Unanimity** If there is some $\mathcal{A}$-labelling L such that $L_i = L$ for all $i \in Ag$ then $F_{\mathcal{A}}(\mathbf{L}) = L$.

Next we come to the first of our postulates that refers to varying the AF. The idea is that AFs that are *isomorphic* should be treated the same when aggregating (cf the *language independence* principle of [2]). Given $\mathcal{A}_1 = (Args_1, \rightharpoonup_1)$ and $\mathcal{A}_2 = (Args_2, \rightharpoonup_2)$, an isomorphism from $\mathcal{A}_1$ to $\mathcal{A}_2$ is a bijection $g : Args_1 \to Args_2$ such that, for all $a, b \in Args_1$ we have $a \rightharpoonup_1 b$ iff $g(a) \rightharpoonup_2 g(b)$. An isomorphism between $\mathcal{A}_1$ and $\mathcal{A}_2$ extends to a mapping between the $\mathcal{A}_1$-labellings and the $\mathcal{A}_2$-labellings. For any $\mathcal{A}_1$-labelling L we define the $\mathcal{A}_2$-labelling

$g(L)$ by setting, for all $a \in \mathcal{A}_2$, $[g(L)](a) = L(g^{-1}(a))$. The function g further extends naturally to a mapping between $\mathcal{A}_1$-profiles and $\mathcal{A}_2$-profiles by setting, for any $\mathcal{A}_1$-profile $\mathbf{L} = (L_1, \ldots, L_n)$, $g(\mathbf{L}) = (g(L_1), \ldots, g(L_n))$.

> **Isomorphism** Suppose $\mathcal{A}_1$ and $\mathcal{A}_2$ are connected by isomorphism g. Then, for any $\mathcal{A}_1$-profile $\mathbf{L}$ we have $g(F_{\mathcal{A}_1}(\mathbf{L})) = F_{\mathcal{A}_2}(g(\mathbf{L}))$.

Isomorphism enforces a certain kind of *neutrality* over the arguments. If we add it to *Anonymity* and *Minimal Collective Rationality* then it means that in any 2-loop AF$\{a, b\}$ in which the number of `in`-votes and `out` votes for a is equal, the collective label must be `undec` for both a and b.

Proposition 1 *Let F be an aggregation method that satisfies* Isomorphism*,* Anonymity *and* Minimal Collective Rationality *and let $\mathcal{A} = \{a, b\}$ be a 2-loop AF. If $\mathbf{L}$ is such that $|V^{\mathbf{L}}_{a:\texttt{in}}| = |V^{\mathbf{L}}_{a:\texttt{out}}|$ then $[F_{\mathcal{A}}(\mathbf{L})](a) = \texttt{undec} = [F_{\mathcal{A}}(\mathbf{L})](b)$.*

A standard idea in aggregation is that the group evaluation concerning some item should depend only on the individuals' evaluations over that item and no others. In our setting this means that the group label $[F_{\mathcal{A}}(\mathbf{L})](a)$ attached to an argument a in an AF $\mathcal{A}$ should depend only on the tuple $\mathbf{L}[a] = (L_i(a))_{i \in Ag}$.

> **Independence** If $\mathbf{L}_1$ and $\mathbf{L}_2$ are $\mathcal{A}$-profiles and $a \in Args_{\mathcal{A}}$ then $\mathbf{L}_1[a] = \mathbf{L}_2[a]$ implies $[F_{\mathcal{A}}(\mathbf{L}_1)](a) = [F_{\mathcal{A}}(\mathbf{L}_2)](a)$.

Given we allow the AF to vary, we strengthen this property somewhat.

> **AF-Independence** If $\mathbf{L}_1$ and $\mathbf{L}_2$ are profiles over $\mathcal{A}_1$ and $\mathcal{A}_2$ respectively and $a \in Args_{\mathcal{A}_1} \cap Args_{\mathcal{A}_2}$ then $\mathbf{L}_1[a] = \mathbf{L}_2[a]$ implies $[F_{\mathcal{A}_1}(\mathbf{L}_1)](a) = [F_{\mathcal{A}_2}(\mathbf{L}_2)](a)$.

This postulate implies the preceding *Independence* (just put $\mathcal{A}_1 = \mathcal{A}_2$). It roughly says that the collective label of a depends only on $\mathbf{L}[a]$ *no matter what other arguments might be present or absent in $\mathcal{A}$*. It is a very strong postulate. For example, It can be shown that any F satisfying it is completely determined by how it behaves on 2-loop AFs. Meanwhile it is easily seen that *Unanimity* becomes equivalent to its following "argument-wise" variant:

> **AW-Unanimity** For each $a \in Args_{\mathcal{A}}$, if there is some $\texttt{x} \in \{\texttt{in}, \texttt{out}, \texttt{undec}\}$ such that $L_i(a) = \texttt{x}$ for all $i \in Ag$ then $[F_{\mathcal{A}}(\mathbf{L})](a) = \texttt{x}$.

The next two postulates are *monotonicity* rules. `in`/`out`-*Monotonicity* says that if some agents change their individual labels of some arguments in profile $\mathbf{L}$ so that they agree with the collective labelling $F_{\mathcal{A}}(\mathbf{L})$, and assuming that those collective labels are in $\{\texttt{in}, \texttt{out}\}$, then the collective labelling does not change.

> `in`/`out`**-Monotonicity** Let $\mathbf{L}$, $\mathbf{L}'$ be $\mathcal{A}$-profiles such that for all $a \in Args_{\mathcal{A}}$ and all $i \in Ag$, $(L'_i(a) \neq L_i(a)$ implies $L'_i(a) = [F_{\mathcal{A}}(\mathbf{L})](a) \in \{\texttt{in}, \texttt{out}\})$. Then $F_{\mathcal{A}}(\mathbf{L}') = F_{\mathcal{A}}(\mathbf{L})$.

Note we require $[F_{\mathcal{A}}(\mathbf{L})](a) \in \{\texttt{in}, \texttt{out}\}$, i.e., we leave open the possibility that if the collective label is `undec` and some agents change their labels to `undec` then this might cause the collective label to change. Indeed several of our examples in the next section will exhibit this behaviour.

A stronger version of this postulate might also be expected to hold. The intuition behind *Strong* `in`/`out`-*Monotonicity* is that if some agents in $\mathbf{L}$ move their individual labels of some arguments *closer* towards the collective label (and those collective labels belong to $\{\texttt{in}, \texttt{out}\}$), then the resulting collective labelling remains unchanged. To formulate it we use the notion of one label being *between* another two labels. Given $\texttt{x}, \texttt{y}, \texttt{z} \in \{\texttt{in}, \texttt{out}, \texttt{undec}\}$ we say that $\texttt{y}$ is between $\texttt{x}$ and $\texttt{z}$ iff either $\texttt{y} = \texttt{x}$ or $\texttt{y} = \texttt{z}$ or $[\texttt{y} = \texttt{undec}$ and $\texttt{x} \neq \texttt{z}]$.

Strong in/out-Monotonicity Let $\mathbf{L}$, $\mathbf{L}'$ be $\mathcal{A}$-profiles such that for all $a \in Args_{\mathcal{A}}$ such that $[F_{\mathcal{A}}(\mathbf{L})](a) \in \{\mathtt{in}, \mathtt{out}\}$ and all $i \in Ag$, $L'_i(a)$ is between $L_i(a)$ and $[F_{\mathcal{A}}(\mathbf{L})](a)$. Then $F_{\mathcal{A}}(\mathbf{L}') = F_{\mathcal{A}}(\mathbf{L})$.

So, for example, if the collective label of a is in and one agent who voted out for a changes their vote to undec then we should expect the collective label to remain in.

The next postulate was a major motivation behind the operators in [7]. It says the collective label on any argument never goes against the individual label of any agent.

Compatibility For all $i \in Ag$ and $a \in Args_{\mathcal{A}}$ we have $[F_{\mathcal{A}}(\mathbf{L})](a) = \neg L_i(a)$ implies $[F_{\mathcal{A}}(\mathbf{L})](a) = \mathtt{undec}$.

This postulate implies that as soon as there is disagreement among the agents about the label of some argument then that argument *must* be labelled collectively undec. As such, if n is large, we are quite likely to end up with a lot of undec arguments in $F_{\mathcal{A}}(\mathbf{L})$.

Given any n-tuple (l_i) of labels the in/out-*winner in* (l_i) is the label among $\{\mathtt{in}, \mathtt{out}\}$ which appears more frequently in (l_i) (if such a label exists). For example the in/out-winner in $(\mathtt{in}, \mathtt{undec}, \mathtt{out}, \mathtt{undec}, \mathtt{in})$ is in. If x is the in/out-winner in (l_i) then we call $\neg\mathtt{x}$ the in/out-*loser*. A weaker version of *Compatibility* can then be formulated as follows:

in/out-Plurality If x is the in/out-loser in $(L_i(a))_{i \in Ag}$ then $[F_{\mathcal{A}}(\mathbf{L})](a) \neq \mathtt{x}$

This postulate says that the collective label of an argument is never set to in or out if more agents vote for the opposite label.

Proposition 2 *Let F be an aggregation method satisfying* Compatibility*. Then*
(i). *F satisfies* in/out-Plurality*.*
(ii). *If F satisfies* in/out-Monotonicity *then it satisfies* Strong in/out-Monotonicity*.*

4 Interval aggregation methods

Now we describe a family of aggregation methods that we call *interval methods*, which will include a number of interesting special cases. The idea is that we determine $[F_{\mathcal{A}}(\mathbf{L})](a)$ as follows. First, we establish the in/out-winner x in $(L_i(a))_{i \in Ag}$. If no winner exists, i.e., if in and out receive an equal number of votes, then $[F_{\mathcal{A}}(\mathbf{L})](a)$ is set to undec. Otherwise, at this point (and in keeping with in/out-*Plurality*) we discard the in/out-loser $\neg\mathtt{x}$ as a candidate to be the collective label and proceed to check whether the configuration of votes $(|V^{\mathbf{L}}_{a:\neg\mathtt{x}}|, |V^{\mathbf{L}}_{a:\mathtt{x}}|)$ represents a sufficiently acceptable "victory" of x over $\neg\mathtt{x}$. If it does then we set $[F_{\mathcal{A}}(\mathbf{L})](a)$ to x, otherwise to undec.

Formally, let Int_n be the set of *intervals* of non-zero length in $\{0, 1, \ldots, n\}$ (where recall n is the number of agents), i.e., $Int_n = \{(k, l) \mid k < l,\ k, l \in \{0, 1, \ldots, n\}\}$. Let $Y \subseteq Int_n$ be some subset of distinguished intervals in Int_n. Then we define aggregation method F^Y by setting, for each $\mathcal{A}$, $\mathcal{A}$-labelling profile $\mathbf{L}$ and $a \in Args_{\mathcal{A}}$:

$$[F^Y_{\mathcal{A}}(\mathbf{L})](a) = \begin{cases} \mathtt{x} & \text{if } \mathtt{x} \in \{\mathtt{in}, \mathtt{out}\} \text{ and } (|V^{\mathbf{L}}_{a:\neg\mathtt{x}}|, |V^{\mathbf{L}}_{a:\mathtt{x}}|) \in Y \\ \mathtt{undec} & \text{otherwise} \end{cases}$$

We will soon make several requirements on the set Y, but in general the only requirement we place (which is basically required to ensure *Unanimity* is satisfied) is:

(I1) $(0, n) \in Y$

Definition 4 *An* interval aggregation method *is an aggregation method F such that $F = F^Y$ for some $Y \subseteq Int_n$ satisfying* **(I1)***.*

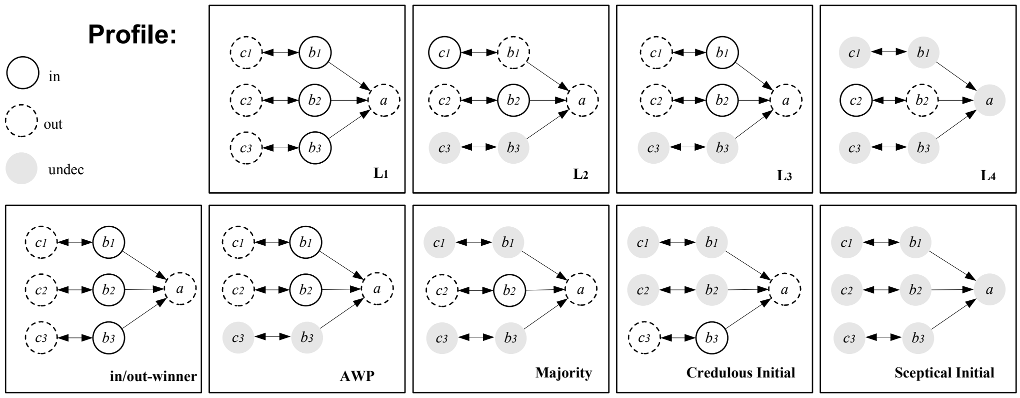


Figure 1: Examples of different interval methods.

We remark that interval methods are closely-related to the *quota rules* considered in judgment aggregation by [10]. In fact if there was no `undec` label then they would be more-or-less the same.

By making different choices of Y we find interesting special instances of interval methods.

Argument-wise plurality: Take the collective label of a to be the label among $\{\texttt{in}, \texttt{out}, \texttt{undec}\}$ that gets the most votes. If there is a tie then take `undec`. This corresponds to $Y_{\text{AWP}} = \{(k,l) \in Int_n \mid n - (k+l) < l\}$. We use F^{AWP} to denote $F^{Y_{\text{AWP}}}$.

Majority: Take the collective label of a to be $\texttt{x}$ if more than half of the agents voted for it, otherwise take `undec`. $Y_{\text{Maj}} = \{(k,l) \in Int_n \mid l > n/2\}$. We use F^{Maj} to denote $F^{Y_{\text{Maj}}}$.

Sceptical initial: [7] Take the `in`/`out` winner if it is the unanimous choice among the agents, otherwise `undec`. $Y_{\text{Scept}} = \{(0,n)\}$. We use F^{Scept} to denote $F^{Y_{\text{Scept}}}$.

Credulous initial: [7] Take the `in`/`out`-winner $\texttt{x}$ whenever no agent voted for $\neg\texttt{x}$, otherwise `undec`. $Y_{\text{Cred}} = \{(0,l) \in Int_n \mid l \geqslant 1\}$. We use F^{Cred} to denote $F^{Y_{\text{Cred}}}$.

`in`/`out`*-winner:* Take the `in`/`out`-winner whenever it exists.[1] $Y_{\text{iow}} = Int_n$. We use F^{iow} to denote $F^{Y_{\text{iow}}}$.

Fig. 1 shows an example comparing the labellings returned by these instances for a particular AF $\mathcal{A}$ and $\mathcal{A}$-profile, assuming $n = 4$.

We can compare different specific aggregation methods according to how *committed* their output collective labellings tend to be [7]. Given any two $\mathcal{A}$-labellings L_1, L_2 we write $L_1 \sqsubseteq L_2$ iff both $L_1^{-1}(\texttt{in}) \subseteq L_2^{-1}(\texttt{in})$ and $L_1^{-1}(\texttt{out}) \subseteq L_2^{-1}(\texttt{out})$. In other words every argument labelled `in` (resp. `out`) by L_1 is also labelled `in` (resp. `out`) by L_2. It is easy to see $\sqsubseteq$ forms a partial order over the set of all $\mathcal{A}$-labellings.

Proposition 3 *If $Y_1 \subseteq Y_2 \subseteq Int_n$ then, for any AF $\mathcal{A}$ and any $\mathcal{A}$-profile* $\mathbf{L}$*, we have* $F_{\mathcal{A}}^{Y_1}(\mathbf{L}) \sqsubseteq F_{\mathcal{A}}^{Y_2}(\mathbf{L})$.

The above result means we get $F_{\mathcal{A}}^{\text{Scept}}(\mathbf{L}) \sqsubseteq F_{\mathcal{A}}^{Y}(\mathbf{L}) \sqsubseteq F_{\mathcal{A}}^{\text{iow}}(\mathbf{L})$ for all $\mathcal{A}$ and $\mathcal{A}$-profiles $\mathbf{L}$ and all interval methods F^Y. So the Sceptical initial and the `in`/`out`-winner are interval methods standing at opposite ends of the commitment spectrum. Among all interval methods F^{Scept} will always return the *least* committed outcome, with arguments tending more often to be collectively labelled `undec`, while F^{iow} always returns the *most* committed outcome.

[1] This rule is sometimes known as *simple majority* in voting theory.

Which postulates from our previous section hold for the general family of interval methods? We obtain the following axiomatic characterisation (which can be compared to the characterisation of quota rules in JA given in [10])

Theorem 1 *Let F be an aggregation method. Then F is an interval aggregation method iff it satisfies:* Minimal Collective Rationality*,* Anonymity*,* Unanimity*,* Isomorphism*,* AF-Independence*,* `in`/`out`-Plurality*.*

Thus we see that most of the postulates from the previous section are sound for the interval methods. The postulates missing from Thm. 1 are the two *Monotonicity* postulates, *Compatibility* and, most significantly, *Collective Rationality*. None of these will hold in general for interval methods, at least not without placing some extra restrictions on Y beyond only **(I1)**. Looking first at `in`/`out`-*Monotonicity* we can construct counterexamples such as the following.

Example 4 *Assume $n = 4$ and let Y be such that $(1,2) \in Y$ but $(0,3) \notin Y$. Let $\mathcal{A}$ be a 2-loop AF $\{a,b\}$ with the three possible complete $\mathcal{A}$-labellings denoted by $L_{\mathtt{in}}$, $L_{\mathtt{out}}$ and $L_{\mathtt{undec}}$, where the subscript represents the label of a. If $\mathbf{L} = (L_{\mathtt{in}}, L_{\mathtt{in}}, L_{\mathtt{out}}, L_{\mathtt{undec}})$ then $[F^Y_{\mathcal{A}}(\mathbf{L})](a) = \mathtt{in}$, since $(1,2) \in Y$. Suppose $\mathbf{L}' = (L_{\mathtt{in}}, L_{\mathtt{in}}, L_{\mathtt{in}}, L_{\mathtt{undec}})$, i.e., agent 3 changes their labelling to $F^Y_{\mathcal{A}}(\mathbf{L})$. If F^Y satisfied* `in`/`out`-Monotonicity *then we would get again $[F^Y_{\mathcal{A}}(\mathbf{L}')] = \mathtt{in}$. But because we assumed $(0,3) \notin Y$ we get $[F^Y_{\mathcal{A}}(\mathbf{L}')](a) = \mathtt{undec}$.*

Although not sound for general interval methods, one can show that all our concrete instances of such methods satisfy `in`/`out`-*Monotonicity*.

Proposition 4 *F^{AWP}, F^{Maj}, F^{Scept}, F^{Cred} and F^{iow} all satisfy* `in`/`out`-Monotonicity*.*

However, not all of these methods satisfy *Strong* `in`/`out`-*Monotonicity*. Specifically, F^{AWP} fails it, as the following example shows:

Example 5 *As in the previous example assume $n = 4$ and $\mathcal{A} = \{a,b\}$ is a 2-loop AF and that $\mathbf{L} = (L_{\mathtt{in}}, L_{\mathtt{in}}, L_{\mathtt{out}}, L_{\mathtt{undec}})$. Then $F^{\mathrm{AWP}}_{\mathcal{A}}(\mathbf{L}) = L_{\mathtt{in}}$. Now suppose $\mathbf{L}' = (L_{\mathtt{in}}, L_{\mathtt{in}}, L_{\mathtt{undec}}, L_{\mathtt{undec}})$, i.e., agent 3 changes their labelling from $L_{\mathtt{out}}$ to $L_{\mathtt{undec}}$. We have that $[F^{\mathrm{AWP}}_{\mathcal{A}}(\mathbf{L})](a) \in \{\mathtt{in}, \mathtt{out}\}$ and that $L_{\mathtt{undec}}(a)$ is between $L_{\mathtt{out}}(a)$ and $[F^{\mathrm{AWP}}_{\mathcal{A}}(\mathbf{L})](a)$ (similarly for b), hence if F^{AWP} satisfied* Strong `in`/`out`-Monotonicity *we would again get $F^{\mathrm{AWP}}_{\mathcal{A}}(\mathbf{L}') = L_{\mathtt{in}}$. But $F^{\mathrm{AWP}}_{\mathcal{A}}(\mathbf{L}') = L_{\mathtt{undec}}$.*

We can obtain *Strong* `in`/`out`-*Monotonicity* for an interval method F^Y if we assume Y satisfies an extra condition saying that Y is closed under *widening* intervals:

(I2) If $(k,l) \in Y$ and $s \leqslant k,\ l \leqslant t$ then $(s,t) \in Y$.

Y_{AWP} does not satisfy **(I2)**. For instance if $n = 4$ then $(1,2) \in Y_{\mathrm{AWP}}$ but $(0,2) \notin Y_{\mathrm{AWP}}$, which enabled the above counterexample.

Proposition 5 *Let F^Y be an interval method. Then F^Y satisfies* Strong `in`/`out`-Monotonicity *iff Y satisfies* **(I2)***.*

Definition 5 *If $Y \subseteq Int_n$ satisfies both* **(I1)** *and* **(I2)** *then we say Y is* widening*. A* widening interval method *is an aggregation method F such that $F = F^Y$ for some widening Y.*

We remark that widening interval methods have essentially been previously proposed in voting theory with an alternative formulation under the name *quota systems* [18]. The main difference with our work is that they only consider aggregation over a single issue (i.e., a single argument in our setting).

Putting Thm. 1 and Prop. 5 together we can see that the class of widening interval methods is characterised by the six postulates of Thm. 1 plus *Strong* in/out*-Monotonicity*.

It can be checked that each of our previous examples of interval methods, apart from Y_{AWP}, are widening and so yield interval methods that satisfy *Strong* in/out*-Monotonicity*. However not all of them yield a method that satisfies *Compatibility*. This can be seen already in the example in Fig. 1, where we have, e.g., $[F^{\text{iow}}_{\mathcal{A}}(\mathbf{L})](b_1) = \texttt{out}$ but $L_2(b_1) = \texttt{in}$. If we want *Compatibility* to hold then we need to place a further restriction on Y:

(I3) If $(k, l) \in Y$ then $k = 0$.

Proposition 6 *Let F^Y be an interval method. Then F^Y satisfies* Compatibility *iff Y satisfies* **(I3)**.

Clearly, among our examples, Y_{Scept} and Y_{Cred} are the only Y that satisfy **(I3)**, which means that F^{Scept} and F^{Cred} are the only interval methods among our examples that satisfy *Compatibility*. Looking more generally, combining the previous proposition with Thm. 1 and Prop. 5 (and recalling the facts about *Compatibility* in Prop. 2) gives us the following result.

Theorem 2 *Let F be an aggregation method. Then the following are equivalent:*
(i). *$F = F^Y$ for some Y of the form $\{(0, t) \mid t \geqslant l\}$ for some $1 \leqslant l \leqslant n$.*
(ii). *F satisfies the following postulates:* Minimal Collective Rationality*,* Anonymity*,* Unanimity*,* Isomorphism*,* AF-Independence*,* Compatibility*,* in/out-Monotonicity*.*

Regarding *Collective Rationality*, we know already from [7, 17] that our examples of interval methods above fail to satisfy it. Is there any requirement we can place on Y to ensure it? The answer is no, as the following impossibility result (whose proof has a flavour of similar impossibility results commonly seen in JA, e.g., Thm. 1 of [13]) shows.

Theorem 3 *There is no aggregation method (for any $n > 1$) satisfying all of* Isomorphism*,* Anonymity*,* Unanimity*,* AF-Independence *and* Collective Rationality*.*

The above result says that, given the basic requirements *Isomorphism*, *Anonymity* and *Unanimity*, there is no hope to obtain *both* of *AF-Independence* and *Collective Rationality*. We have to weaken one of them. Next we'll look at some weakenings of *Collective Rationality*, and after that we'll look at relaxing *AF-Independence*.

5 Weakening Collective Rationality

We have already met one weakening of *Collective Rationality*, namely *Minimal Collective Rationality*. In this section we look at a couple of others which, rather than look for restricted classes of AF for which fully rational outcomes are guaranteed, as *Minimal Collective Rationality* does, look at relaxing the requirement of *completeness* of the collective labelling (see Definition 2).

Our first weakening says merely that the group never ends up collectively accepting two arguments that are directly *conflicting*:

Conflict-freeness If $[F_{\mathcal{A}}(\mathbf{L})](a) = \texttt{in}$ and $a \rightharpoonup_{\mathcal{A}} b$ then $[F_{\mathcal{A}}(\mathbf{L})](b) \neq \texttt{in}$

Thankfully every interval method already satisfies this basic requirement. This follows from the next proposition.

Proposition 7 *Every aggregation method F that satisfies* in/out-Plurality *also satisfies* Conflict-freeness*.*

A way to strengthen *Conflict-freeness* is to require that if the group collectively accepts an argument a, then the group also explicitly *rejects* every argument attacked by a:

`out`-semi-legal If $[F_{\mathcal{A}}(\mathbf{L})](a) = \texttt{in}$ and $a \rightharpoonup_{\mathcal{A}} b$ then $[F_{\mathcal{A}}(\mathbf{L})](b) = \texttt{out}$

Similarly a should be collectively accepted only if every attacker of a is collectively rejected.

`in`-semi-legal If $[F_{\mathcal{A}}(\mathbf{L})](a) = \texttt{in}$ and $b \rightharpoonup_{\mathcal{A}} a$ then $[F_{\mathcal{A}}(\mathbf{L})](b) = \texttt{out}$

(Note this corresponds to the first item in the definition of complete labelling (Definition 2.)) It turns out that, for interval methods F^Y, the satisfaction of both these postulates is intimately bound with the requirement that Y is widening.

Proposition 8 *Let F^Y be an interval method. Then the following are equivalent:* (i). *Y is widening,* (ii). *F^Y satisfies* `out`-semi-legal*,* (iii). *F^Y satisfies* `in`-semi-legal.

From this proposition we see that we could replace *Strong* `in`/`out`-*Monotonicity* in the characterisation of widening interval methods by either `out`-*semi-legal* or `in`-*semi-legal*.

An arbitrary $\mathcal{A}$-labelling L that simultaneously satisfies both of the conditions expressed in `out`-*semi-legal* and `in`-*semi-legal* is called *subcomplete* in [4].

Definition 6 ([4]) *Let $\mathcal{A}$ be an AF and L an $\mathcal{A}$-labelling. If, for all $a, b \in Args_{\mathcal{A}}$, it holds that if $L(a) = \texttt{in}$ and [$a \rightharpoonup_{\mathcal{A}} b$ or $b \rightharpoonup_{\mathcal{A}} a$] then $L(b) = \texttt{out}$, then we call L a* subcomplete *$\mathcal{A}$-labelling.*

This class of subcomplete labellings is of special interest because, as is shown in [4], they are precisely the labellings that may be *embedded* as part of a complete labelling over a *larger* framework $\mathcal{A}'$. We introduce some notation.

Definition 7 *Let $\mathcal{A}_1 = (A_1, \rightharpoonup_1)$ and $\mathcal{A}_2 = (A_2, \rightharpoonup_2)$ be two AFs. We say $\mathcal{A}_1$ is a* sub-framework *of $\mathcal{A}_2$, written $\mathcal{A}_1 \subseteq_{\mathrm{f}} \mathcal{A}_2$, iff $A_1 \subseteq A_2$ and $\rightharpoonup_1 = \rightharpoonup_2 \cap (A_1 \times A_1)$.*

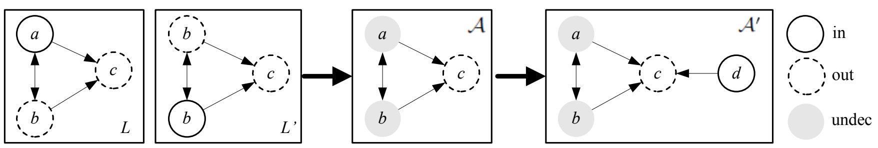
Proposition 9 ([4]) *Let $\mathcal{A}$ be an AF and L be an $\mathcal{A}$-labelling. Then L is a subcomplete $\mathcal{A}$-labelling iff there exists $\mathcal{A}'$ and $L' \in Comp(\mathcal{A}')$ s.t. $\mathcal{A} \subseteq_{\mathrm{f}} \mathcal{A}'$ and $L = L'[Args_{\mathcal{A}}]$.*

This result means the conjunction of `out`-*semi-legal* and `in`-*semi-legal* is equivalent to:

Collective Embeddibility There exists $\mathcal{A}'$ and $L' \in Comp(\mathcal{A}')$ s.t. $\mathcal{A} \subseteq_{\mathrm{f}} \mathcal{A}'$ and $F_{\mathcal{A}}(\mathbf{L}) = L'[Args_{\mathcal{A}}]$.

Collective Embeddibility relaxes *Collective Rationality* by saying that, although the collective labelling might not be a complete $\mathcal{A}$-labelling, it is always possible to *recover* completeness by adding some more arguments to $\mathcal{A}$.

Example 6 *Consider the AF $\mathcal{A}$ with three arguments $\{a, b, c\}$ and the two complete $\mathcal{A}$-labellings L, L' depicted on the left below.*



Assume we have an even number of agents and an $\mathcal{A}$-profile $\mathbf{L}$ in which half the agents submit L and half submit L'. Every interval method will yield the collective labelling $F^Y(\mathbf{L})$ in which a and b are `undec` *and c is* `out`*. This $F^Y(\mathbf{L})$ is not complete for $\mathcal{A}$, but if we now expand $\mathcal{A}$ to $\mathcal{A}'$ by adding a new argument d such that d attacks c as pictured on the right then we do get that $F^Y(\mathbf{L})$ can be extended to a complete $\mathcal{A}'$-labelling, namely the one in which additionally d is* `in`.

Given Props. 7 and 9, we can use *Collective Embeddibility* to obtain yet another alternative characterisation of the widening interval methods.

Theorem 4 *Let F be an aggregation method. Then F is a widening interval aggregation method iff it satisfies:* Minimal Collective Rationality*,* Anonymity*,* Unanimity*,* Isomorphism*,* `in`/`out`-Plurality*,* AF-Independence*,* Collective Embeddibility*.*

6 Weakening AF-Independence

We have seen that it is possible to define aggregation methods that satisfy *AF-Independence* while holding on to some interesting weakenings of *Collective Rationality*. Now we explore the possibility of satisfying *Collective Rationality* at the expense of *AF-Independence*. One might argue anyway that *AF-Independence* is too strong. Indeed how can it be expected to hold when part of the input to the aggregation explicitly contains information (in the form of the attack relation $\rightharpoonup_{\mathcal{A}}$) regarding dependencies between arguments? Instead we might expect the following weaker version, inspired by the work of Baroni and Giacomin [2], who proposed a similar postulate for argumentation semantics. The idea is that if we have a set of arguments in $\mathcal{A}$ that is *unattacked* (that is, no argument in the set is attacked by any argument outside the set) then we can aggregate just that part without looking at the arguments outside the set. In this way the collective label of an argument a might be influenced by what happens "upstream" from a (i.e., what labels are assigned to the attackers of a and, in turn, their attackers) but not "downstream".

> **Directionality** Suppose $\mathcal{A} \subseteq_{\mathrm{f}} \mathcal{A}'$ and suppose $Args_{\mathcal{A}}$ is unattacked in $\mathcal{A}'$. Then for any $\mathcal{A}'$-profile $\mathbf{L}$ and $a \in Args_{\mathcal{A}}$ we have $[F_{\mathcal{A}'}(\mathbf{L})](a) = [F_{\mathcal{A}}(\mathbf{L}[Args_{\mathcal{A}}])](a)$.

Proposition 10 *Every aggregation method F that satisfies* AF-Independence *also satisfies* Directionality*.*

The question is: by weakening *AF-Independence* to *Directionality* do we obtain a possibility result? That is, can we construct an aggregation method that satisfies *Directionality*, *Collective Rationality* and some of our other nice postulates? We show the answer is yes, using the *down-admissible* and *up-complete* constructions of [7].

6.1 Down-admissible and up-complete

We begin with the down-admissible construction, which uses the definition of the 'committedness' relation $\sqsubseteq$ that was introduced for Prop. 3.

Definition 8 ([7]) *Given an $\mathcal{A}$-labelling L, the* down-admissible *labelling of L, denoted by $\downharpoonright L$, is the (unique) greatest element (under $\sqsubseteq$) of the set of all admissible $\mathcal{A}$-labellings M such that $M \sqsubseteq L$.*

A constructive definition of $\downharpoonright L$ is given in [7]. It can be arrived at by just iteratively relabelling every argument that is illegally `in` or illegally `out` with `undec` until no illegal `in` or `out` labels remain. The end result is a labelling that is admissible.

Example 7 *Consider the sequence of the 3 leftmost labellings of the AF $\mathcal{A}$ shown in Fig. 2. The initial labelling L is on the far left. The only argument illegally labelled* `in` *or* `out` *in L is c (because it is* `out` *but none of its attackers is* `in`*), so its label is changed to* `undec` *(2nd labelling). This change in turn causes d to become illegally* `in`*, so then d's label is also changed to* `undec`*. At this point there are no illegally* `in` *or* `out` *arguments left and so the procedure stops with $\downharpoonright L$ as the 3rd labelling.*

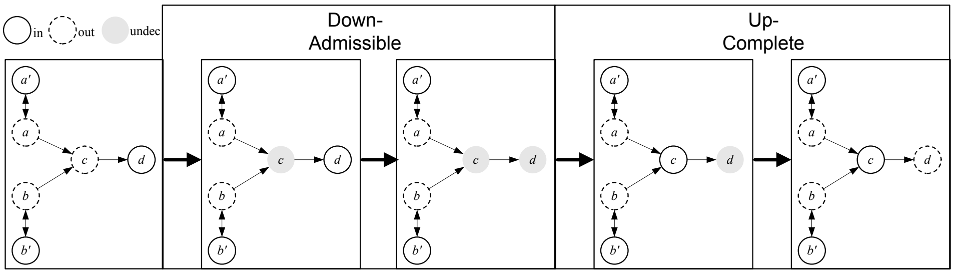


Figure 2: Example showing down-admissible and up-complete procedures.

As the example illustrates, $\downharpoonleft L$ might not be a complete labelling. To ensure a complete labelling we need an additional step which applies the up-complete operator.

Definition 9 ([7]) *Given an admissible $\mathcal{A}$-labelling L, the* up-complete *labelling of L, denoted by $\upharpoonleft L$, is the (unique) smallest element (under $\sqsubseteq$) of the set of all complete $\mathcal{A}$-labellings M such that $L \sqsubseteq M$.*

There is also a constructive definition of $\upharpoonleft L$ in [7]. Just iteratively change every illegally `undec` argument to `in` or `out` as appropriate, until no illegally labelled arguments remain.

Example 8 *Consider the sequence of the 3 rightmost labellings in Fig. 2, which starts with $L' = \downharpoonleft L$ from the previous example. There is one illegally* `undec` *argument, namely c. Since both attackers a, b are labelled* `out`*, we change c's label to* `in`*. At this point d becomes illegally* `undec` *because it now has an attacker c which is* `in`*. Hence we change d's label to* `out`*. Now there are no illegally* `undec` *arguments and so the process stops and returns the rightmost labelling as $\upharpoonleft L'$.*

We denote by $\upharpoonleft\!\downharpoonleft L$ the composite operation of performing the down-admissible followed by the up-complete procedures on L. Since $\upharpoonleft\!\downharpoonleft L$ is guaranteed to return a complete $\mathcal{A}$-labelling by construction, this opens up a way to transform *any* aggregation method F into one that satisfies *Collective Rationality.*

Definition 10 *Given any aggregation method F, the* DAUC version *of F is the aggregation method $\widehat{F}$ defined by setting, for any AF $\mathcal{A}$ and $\mathcal{A}$-labelling profile $\mathbf{L}$, $\widehat{F}_{\mathcal{A}}(\mathbf{L}) = \upharpoonleft\!\downharpoonleft(F_{\mathcal{A}}(\mathbf{L}))$.*

In other words, to aggregate using $\widehat{F}$, first aggregate using F and then perform the down-admissible followed by the up-complete on the result.

Example 9 *Consider again the AF $\mathcal{A}$ and $\mathcal{A}$-profile $\mathbf{L}$ from Example 6. We saw there that every interval method yields the non-complete $\mathcal{A}$-labelling $F_{\mathcal{A}}(\mathbf{L})$ in which a, b are both labelled* `undec` *and c is labelled* `out`*. Then applying the DAUC procedure yields the labelling $\widehat{F}^Y_{\mathcal{A}}(\mathbf{L})$, which labels all of a, b, c with* `undec`*.*

For the special cases of interval methods F^{Scept} and F^{Cred} this procedure was studied in detail in [7]. Their DAUC versions were called the *sceptical* and *super-credulous* aggregation methods respectively there.[2] Which postulates, apart from *Collective Rationality*, are satisfied by the DAUC versions of the general family of interval methods? We lose *AF-Independence* (and indeed plain *Independence*), as expected. This can be seen in the above Example 9 where we get $[\widehat{F}^Y_{\mathcal{A}}(\mathbf{L})](c) =$ `undec` but $[\widehat{F}^Y_{\mathcal{A}}(\mathbf{L}')](c) =$ `out` if every agent in $\mathbf{L}'$ submits the *same* labelling L such that $L(c) =$ `out`, despite the fact that $\mathbf{L}[c] = \mathbf{L}'[c]$. But some postulates satisfied by the initial method F can be *inherited* by $\widehat{F}$.

[2]Note there is also a third operator in [7] called the *credulous* method, which applies just the DA procedure (without UC) to F^{Cred}. However, as noted there, this doesn't always yield a complete labelling.

Proposition 11 *Let F be any aggregation method. For each of the following postulates, if F satisfies that postulate then so does $\widehat{F}$:* Anonymity*,* Unanimity*,* Isomorphism*,* Directionality*,* Compatibility*.*

For the case when the initial F is an interval method the above result gives the following:

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

Hence we have established that, for every interval method F, $\widehat{F}$ satisfies four of the six postulates that characterised the interval methods in Thm. 1, plus a weaker version (*Directionality*) of a fifth (*AF-Independence*). What about the remaining postulate from there, i.e., `in`/`out`-*Plurality*? From Prop. 11 above we know that if Y satisfies **(I3)** then $\widehat{F}^Y$ will satisfy *Compatibility* and hence `in`/`out`-*Plurality.* Thus **(I3)** is sufficient to obtain `in`/`out`-*Plurality.* Surprisingly, it turns out this condition is also necessary.

Proposition 12 *Let F^Y be an interval method. The $\widehat{F}^Y$ satisfies* `in`/`out`-Plurality *iff Y satisfies* **(I3)***.*

One last question concerns the circumstances under which $\widehat{F}^Y$ will satisfy *(Strong)* `in`/`out`-*Monotonicity.* Since for interval methods we have that *Strong* `in`/`out`-*Monotonicity* holds iff Y is widening, it might be expected that an analogous equivalence is preserved for the class of DAUC versions of the interval methods. However we have so far been unable to prove or disprove this, so it remains open for now.

7 Conclusion

We introduced a framework for exploring labelling aggregation operators in abstract argumentation, with the central concept being that of *aggregation method* that takes as input both an AF together with a profile of complete $\mathcal{A}$-labellings and returns a collective $\mathcal{A}$-labelling. We formulated some postulates for aggregation methods and axiomatically characterised the family of interval aggregation methods, which don't satisfy the desirable postulate of *Collective Rationality*, but do satisfy the strong postulate of *AF-Independence.* We showed that nevertheless some interesting weakenings of *Collective Rationality* do hold. We also showed that using the down-admissible plus up-complete construction of Caminada and Pigozzi allows a way to turn any aggregation method into one satisfying *Collective Rationality* while preserving a weaker and more reasonable form of independence known as *Directionality*, and we identified some postulates that are sound for (certain subclasses of) the DAUC versions of the interval methods. Table 1 summarises which of our postulates are satisfied by the various classes of aggregation methods we have looked at in this paper.

There are several avenues for future work. As well as answering the open question concerning the satisfaction or otherwise of *(Strong)* `in`/`out`-*Monotonicity* for the DAUC versions of the interval methods, we would like to find a complete axiomatisation of such methods (and their widening subclass). Another plan concerns investigation of *dialogue games* in the context of aggregating argument labellings. Such games are commonly used in abstract argumentation as a way of formulating argumentation semantics [15]. It would be interesting in particular to devise a dialogue game for the DAUC procedure. This would likely give insights into how a group might form a collective labelling via *discussion.* Finally we would like to look at questions involving *manipulation* of the aggregation process. The problems of *agenda manipulation* has been already studied in JA [8]. In our particular setting this would involve questions such as "which argument(s) could be added to the AF in order to achieve a particular collective outcome"?

Postulates	(I1)		(I1) + (I2)		(I1)+(I2) +(I3)	
	F^Y	$\widehat{F}^Y$	F^Y	$\widehat{F}^Y$	F^Y	$\widehat{F}^Y$
Independence	✓	✗	✓	✗	✓	✗
Anonymity	✓	✓	✓	✓	✓	✓
Unanimity	✓	✓	✓	✓	✓	✓
Isomorphism	✓	✓	✓	✓	✓	✓
AF-Independence	✓	✗	✓	✗	✓	✗
Directionality	✓	✓	✓	✓	✓	✓
Collective Rationality	✗	✓	✗	✓	✗	✓
Minimal Collective Rationality	✓	✓	✓	✓	✓	✓
Collective Embeddibility	✗	✓	✓	✓	✓	✓
in**-semi-legal**	✗	✓	✓	✓	✓	✓
out**-semi-legal**	✗	✓	✓	✓	✓	✓
Conflict-freeness	✓	✓	✓	✓	✓	✓
Strong in/out**-Monotonicity**	✗	✗	✓	?	✓	?
in/out**-Plurality**	✓	✗	✓	✗	✓	✓
Compatibility	✗	✗	✗	✗	✓	✓

Table 1: The postulates that are satisfied/violated by interval methods satisfying **(I1)**, **(I2)**, and **(I3)**.

Acknowledgments

Richard Booth is supported by the Fonds National de Recherche (DYNGBaT project). Edmond Awad benefitted from a SINTELNET grant. Thanks are due to the reviewers of COMSOC and KR, and to Martin Caminada and Souhila Kaci for helpful comments.

References

[1] Pietro Baroni, Martin Caminada, and Massimiliano Giacomin. An introduction to argumentation semantics. *Knowledge Engineering Review*, 26(4):365–410, 2011.

[2] Pietro Baroni and Massimiliano Giacomin. On principle-based evaluation of extension-based argumentation semantics. *Artificial Intelligence*, 171(10):675–700, 2007.

[3] Andrei Bondarenko, Phan Minh Dung, Robert Kowalski, and Francesca Toni. An abstract, argumentation-theoretic approach to default reasoning. *Artificial Intelligence*, 93:63–101, 1997.

[4] Richard Booth, Souhila Kaci, Tjitze Rienstra, and Leon van der Torre. Conditional acceptance functions. In *Proceedings of the 4th International Conference on Computational Models of Argument (COMMA 2012)*, volume 245, pages 470–477. IOS Press, 2012.

[5] Martin Caminada. On the issue of reinstatement in argumentation. In *Logics in artificial intelligence*, pages 111–123. Springer, 2006.

[6] Martin Caminada and Dov Gabbay. A logical account of formal argumentation. *Studia Logica*, 93(2-3):109–145, 2009.

[7] Martin Caminada and Gabriella Pigozzi. On judgment aggregation in abstract argumentation. *Autonomous Agents and Multi-Agent Systems*, 22(1):64–102, 2011.

[8] Fabrizio Cariani, Marc Pauly, and Josh Snyder. Decision framing in judgment aggregation. *Synthese*, 163(1):1–24, 2008.

[9] Sylvie Coste-Marquis, Caroline Devred, Sébastien Konieczny, Marie-Christine Lagasquie-Schiex, and Pierre Marquis. On the merging of dung's argumentation systems. *Artificial Intelligence*, 171(10):730–753, 2007.

[10] Franz Dietrich and Christian List. Judgment aggregation by quota rules: Majority voting generalized. *Journal of Theoretical Politics*, 19(4):391–424, 2007.

[11] Phan Minh Dung. On the acceptability of arguments and its fundamental role in nonmonotonic reasoning, logic programming and n-person games. *Artificial intelligence*, 77(2):321–357, 1995.

[12] Guido Governatori, Michael Maher, Grigoris Antoniou, and David Billington. Argumentation semantics for defeasible logic. *Journal of Logic and Computation*, 14(5):675–702, 2004.

[13] Christian List and Philip Pettit. Aggregating sets of judgments: An impossibility result. *Economics and Philosophy*, 18(01):89–110, 2002.

[14] Christian List and Clemens Puppe. Judgment aggregation: A survey. In *Handbook of Rational and Social Choice*. Oxford University Press, 2009.

[15] Sanjay Modgil and Martin Caminada. Proof theories and algorithms for abstract argumentation frameworks. In *Argumentation in artificial intelligence*, pages 105–129. Springer, 2009.

[16] Sanjay Modgil and Henry Prakken. A general account of argumentation with preferences. *Artificial Intelligence*, 195:361–397, 2013.

[17] Iyad Rahwan and Fernando Tohmé. Collective argument evaluation as judgement aggregation. In *Proceedings of the 9th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2010)*, pages 417–424, 2010.

[18] Stephen C Young, Alan D Taylor, and William S Zwicker. Counting quota systems: A combinatorial question from social choice theory. *Mathematics Magazine*, pages 331–342, 1995.

Appendix: Selected Proof Outlines

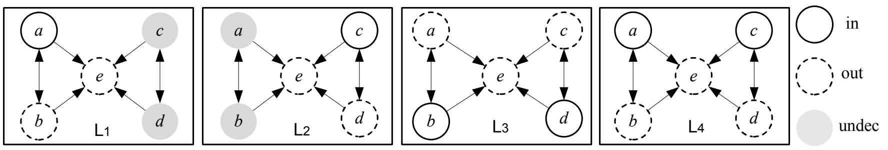
Theorem 1 Let F be an aggregation method. Then F is an interval aggregation method iff it satisfies: *Minimal Collective Rationality*, *Anonymity*, *Unanimity*, *Isomorphism*, *AF-Independence*, `in`/`out`-*Plurality*.

Proof: Soundness is relatively straightforward. For completeness we first show how to construct, from any given aggregation method F, a subset $Y(F) \subseteq Int_n$: Let $\mathcal{A}_0 = (Args_0, \rightharpoonup_0)$ be a 2-loop AF such that $Args_0 = \{a_0, b_0\}$. There are three complete labellings for $\mathcal{A}_0$, which we denote by $L_{\texttt{in}}$, $L_{\texttt{out}}$ and $L_{\texttt{undec}}$, where the subscript represents the label of a_0

(with the label of b_0 of course being always $\neg L(a_0)$). Then we define $Y(F)$ by setting $Y(F) = \{(k,l) \in Int_n \mid [F_{\mathcal{A}_0}(\mathbf{L}_{k,l})](a_0) = \mathtt{in}\}$, where $\mathbf{L}_{k,l}$ is any $\mathcal{A}_0$-profile such that precisely k agents provide labelling $L_{\mathtt{out}}$ and l agents provide $L_{\mathtt{in}}$. Note by *Anonymity* that the precise distribution of labellings among $\mathbf{L}_{k,l}$ doesn't matter. $Y(F)$ is well-defined, i.e., it doesn't matter which 2-loop AF we take to define it (by *Isomorphism*) and $Y(F)$ satisfies **(I1)** (by *Unanimity*). One can then show that F and $F^{Y(F)}$ agree on the 2-loop AF $\mathcal{A}_0$, i.e., that for every $\mathcal{A}_0$-profile $\mathbf{L}$ we have $F_{\mathcal{A}_0}(\mathbf{L}) = F^{Y(F)}_{\mathcal{A}_0}(\mathbf{L})$. This part depends on *Anonymity*, *Isomorphism*, *Minimal Collective Rationality* and `in`/`out`-*Plurality*. Then finally we extend this to hold for *any* argumentation framework $\mathcal{A}$ using *AF-Independence* and *Isomorphism*. ■

Theorem 3 There is no aggregation method (for any $n > 1$) satisfying all of *Isomorphism*, *Anonymity*, *Unanimity*, *AF-Independence* and *Collective Rationality.*

Proof: Assume for contradiction that F is an aggregation method satisfying the named postulates. Consider the following four complete labellings of an AF consisting of five arguments a, b, c, d, e.



Case n is even. Choose any profile $\mathbf{L}$ for this AF such that half the agents submit L_3 and half submit L_4. Then there is no `in`/`out`-winner for any of a, b, c, d and so they will be collectively labelled `undec` by $F_{\mathcal{A}}(\mathbf{L})$ by Prop. 1 and *AF-Independence.* By *Collective Rationality* this implies e is also collectively labelled `undec`. But *AF-Independence* and *Unanimity* together say that e must be collectively labelled `out` - contradiction.
Case n is odd. Define profile $\mathbf{L}'$ by setting $L'_1 = L_1$, $L'_2 = L_2$ and $L'_i = L_3$ if $i \geqslant 3$ and i is odd, $L'_i = L_4$ if $i \geqslant 3$ and i is even. Then again there is no `in`/`out`-winner for any of a, b, c, d and we can repeat the same reasoning as in the previous case to obtain a contradiction. ■

Proposition 8 Let F^Y be an interval method. Then the following are equivalent: *(i).* Y is widening, *(ii).* F^Y satisfies `out`-*semi-legal*, *(iii).* F^Y satisfies `in`-*semi-legal.*

Proof: *(i) ⇒ (ii).* Suppose Y is widening, $[F^Y_{\mathcal{A}}(\mathbf{L})](a) = \mathtt{in}$ and $a \rightharpoonup_{\mathcal{A}} b$. We must show $[F^Y_{\mathcal{A}}(\mathbf{L})](b) = \mathtt{out}$. Since $[F^Y_{\mathcal{A}}(\mathbf{L})](a) = \mathtt{in}$ we know $(|V^{\mathbf{L}}_{a:\mathtt{out}}|, |V^{\mathbf{L}}_{a:\mathtt{in}}|) \in Y$. By the definition of complete $\mathcal{A}$-labelling we know $|V^{\mathbf{L}}_{b:\mathtt{in}}| \leqslant |V^{\mathbf{L}}_{a:\mathtt{out}}|$ and $|V^{\mathbf{L}}_{a:\mathtt{in}}| \leqslant |V^{\mathbf{L}}_{b:\mathtt{out}}|$. Hence, since Y is widening, $(|V^{\mathbf{L}}_{b:\mathtt{in}}|, |V^{\mathbf{L}}_{b:\mathtt{out}}|) \in Y$ and so $[F^Y_{\mathcal{A}}(\mathbf{L})](b) = \mathtt{out}$ as required.
(ii) ⇒ (i). Suppose Y is not widening, i.e., there exist $(k,l) \in Y$, $(s,t) \notin Y$ such that $s \leqslant k$, $l \leqslant t$. We construct $\mathcal{A}$ and $\mathcal{A}$-profile $\mathbf{L}$ such that $a \rightharpoonup_{\mathcal{A}} b$, $[F^Y_{\mathcal{A}}(\mathbf{L})](a) = \mathtt{in}$ but $[F^Y_{\mathcal{A}}(\mathbf{L})](b) \neq \mathtt{out}$. Define $\mathcal{A}$ as follows:

We construct $\mathbf{L}$ via a preliminary profile $\mathbf{L}'$ defined as follows: (1) assume l agents provide labelling $\{(a', \mathtt{out}), (a, \mathtt{in}), (b, \mathtt{out}), (b', \mathtt{in})\}$, (2) k agents provide labelling $\{(a', \mathtt{in}), (a, \mathtt{out}), (b, \mathtt{in}), (b', \mathtt{out})\}$, (3) $n - (k + l)$ agents provide the all `undec` labelling. Note that in $\mathbf{L}'$ we have $(|V_{b:\mathtt{in}}|, |V_{b:\mathtt{out}}|) = (k, l)$. Now transform $\mathbf{L}'$ into $\mathbf{L}$ as follows: (4) change the labellings of $k - s$ agents from group (2) above to $\{(a', \mathtt{in}), (a, \mathtt{out}), (b, \mathtt{undec}), (b', \mathtt{undec})\}$ (so that now $(|V_{b:\mathtt{in}}|, |V_{b:\mathtt{out}}|) = (s, l)$), (5) change the labellings of $t - l$ agents from groups (3) and (4) to $\{(a', \mathtt{in}), (a, \mathtt{out}), (b, \mathtt{out}), (b', \mathtt{in})\}$ (if agent is from group (4)) or $\{(a', \mathtt{undec}), (a, \mathtt{undec}), (b, \mathtt{out}), (b', \mathtt{in})\}$ (if agent is from

group (3)). We now have $(|V^{\mathbf{L}}_{a:\texttt{out}}|, |V^{\mathbf{L}}_{a:\texttt{in}}|) = (k, l) \in Y$ and $(|V^{\mathbf{L}}_{b:\texttt{in}}|, |V^{\mathbf{L}}_{b:\texttt{out}}|) = (s, t) \notin Y$. Thus $[F^Y_{\mathcal{A}}(\mathbf{L})](a) = \texttt{in}$ but $[F^Y_{\mathcal{A}}(\mathbf{L})](b) = \texttt{undec} \neq \texttt{out}$ as required.
(i) ⇒ (iii) and *(iii) ⇒ (i)*. Proved along similar lines as the above. ■

Proposition 11 Let F be any aggregation method. For each of the following postulates, if F satisfies that postulate then so does $\widehat{F}$: *Anonymity*, *Unanimity*, *Isomorphism*, *Directionality*, *Compatibility*.

Proof: The proofs that *Anonymity* and *Unanimity* are preserved are straightforward. The preservation of the other postulates can be seen by considering the constructive definitions of $\downharpoonright$ and $\upharpoonleft$ from [7] (i.e., iteratively relabelling illegally in or out arguments to undec in the case of $\downharpoonright$, and then iteratively relabelling illegally undec arguments to in or out in the case of $\upharpoonleft$).

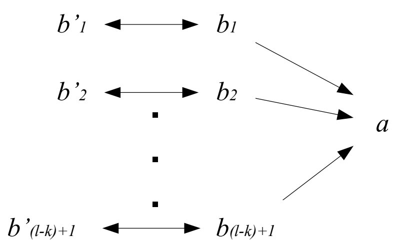
The preservation of *Isomorphism* comes mainly from the fact that, given $\mathcal{A}_1$ and $\mathcal{A}_2$ connected by isomorphism g and $\mathcal{A}_1$-labelling L, then $a \in \mathit{Args}_{\mathcal{A}_1}$ is illegally labelled in L iff $g(a)$ is illegally labelled in $g(L)$.

The preservation of *Directionality* follows from the properties that $\downharpoonright(L[A]) = (\downharpoonright L)[A]$ and $\upharpoonleft(L[A]) = (\upharpoonleft L)[A]$ for any $\mathcal{A}$-labelling L and any unattacked set $A \subseteq \mathit{Args}_{\mathcal{A}}$. These two properties hold since the relabelling of an argument a in the down-admissible and up-complete procedures depends only on the labels of the attackers of a.

Finally *Compatibility* is preserved since if for no a do we have $[F_{\mathcal{A}}(\mathbf{L})](a) = \neg L_i(a)$ and $[F_{\mathcal{A}}(\mathbf{L})](a) \neq \texttt{undec}$ then clearly the same will hold true for $\downharpoonright F_{\mathcal{A}}(\mathbf{L})$ (since $\downharpoonright$ only relabels arguments to undec). The fact that the same also holds true in turn for $\upharpoonleft\downharpoonright F_{\mathcal{A}}(\mathbf{L})$ essentially follows from Lemma 7 of [7]. ■

Proposition 12 Let F^Y be an interval method. Then $\widehat{F}^Y$ satisfies in/out-*Plurality* iff Y satisfies **(I3)**.

Proof: Suppose Y does not satisfy **(I3)** and so $(k, l) \in Y$ for some $k > 0$. We construct an AF $\mathcal{A}$, an $\mathcal{A}$-profile $\mathbf{L}$ and $a \in \mathit{Args}_{\mathcal{A}}$ such that $[\widehat{F}^Y_{\mathcal{A}}(\mathbf{L})](a)$ is the in/out-loser for a in $\mathbf{L}$. We use the following AF:



We define $\mathbf{L}$ by setting, for each $i \in Ag$ and $j = 1, \ldots, l - k + 1$:

$$L_i(b_j) = \begin{cases} \texttt{out} & \text{if } i < j \\ \texttt{in} & \text{if } j \leqslant i \leqslant k + j - 1 \\ \texttt{out} & \text{if } k + j - 1 < i \leqslant k + l \\ \texttt{undec} & \text{otherwise.} \end{cases} \qquad L_i(a) = \begin{cases} \texttt{out} & \text{if } 1 \leqslant i \leqslant l \\ \texttt{in} & \text{if } l < i \leqslant k + l \\ \texttt{undec} & \text{otherwise} \end{cases}$$

$$L_i(b'_j) = \neg L_i(b_j)$$

We have $L_i \in \mathit{Comp}(\mathcal{A})$ for each $i \in Ag$. We also have $|V^{\mathbf{L}}_{b_j:\texttt{out}}| = |V^{\mathbf{L}}_{b'_j:\texttt{in}}| = |V^{\mathbf{L}}_{a:\texttt{out}}| = l$ and $|V^{\mathbf{L}}_{b_j:\texttt{in}}| = |V^{\mathbf{L}}_{b'_j:\texttt{out}}| = |V^{\mathbf{L}}_{a:\texttt{in}}| = k$. Thus, since $(k, l) \in Y$ we have $[F^Y_{\mathcal{A}}(\mathbf{L})](b_j) = [F^Y_{\mathcal{A}}(\mathbf{L})](a) = \texttt{out}$ and $[F^Y_{\mathcal{A}}(\mathbf{L})](b'_j) = \texttt{in}$. Hence$[\widehat{F}^Y_{\mathcal{A}}(\mathbf{L})](a) = \texttt{in}$, i.e., $\widehat{F}^Y_{\mathcal{A}}(\mathbf{L})$ yields the `in`/`out`-loser for a. ■

Richard Booth
University of Luxembourg
Luxembourg

Edmond Awad
Masdar Institute of Science and Technology
Abu Dhabi, UAE

Iyad Rahwan
Masdar Institute of Science and Technology
Abu Dhabi, UAE