Interval Methods for Judgment Aggregation in Argumentation

Richard Booth¹, Edmond Awad² and Iyad Rahwan^{2,3}

¹University of Luxembourg, Luxembourg ²Masdar Institute of Science & Technology, UAE ³University of Edinburgh, UK

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, which 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 downadmissible 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.

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) (Dung 1995). 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 (Baroni, Caminada, and Giacomin 2011). 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 problem of *Judgment Aggregation* (JA) has been explored extensively in classical logic (List and Puppe 2009). But it was only recently that JA has been applied to collective argument evaluation (Caminada and Pigozzi 2011; Rahwan and Tohmé 2010). Early results showed that argumentwise plurality voting cannot guarantee that the outcome of aggregation is always rational (consistent)—thus simple voting violates *Collective Rationality* (Rahwan and Tohmé 2010). 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 (Caminada and Pigozzi 2011).

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. Then, 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*.

Preliminaries

We assume a countably infinite set U of argument names, from which all possible argumentation frameworks are built.

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

An AF is evaluated by assigning one of the labels in, out or undec to each argument in $Args_A$, standing for *accepted*, *rejected* and *undecided* respectively (Caminada 2006). We define a unary "negation" operator on the set of labels by setting $\neg in = out$, $\neg out = in$ and $\neg undec = undec$. Given an AF A, an A-labelling is a function $L : Args \rightarrow$ {in, out, undec}. For each $x \in$ {in, out, undec} we denote by $L^{-1}(x)$ the inverse image of x under L, and given $A \subseteq$ $Args_A$ we denote by L[A] the restriction of L to A.

A *rational* evaluation should somehow respect the attack relation, as captured by the notion of *complete* labelling.

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

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

Copyright © 2014, Association for the Advancement of Artificial Intelligence (www.aaai.org). All rights reserved.

An admissible A-labelling is one that satisfies the first two conditions above. We denote the set of complete Alabellings by Comp(A).

We assume a set of agents $Ag = \{1, \ldots, n\}$ (with $n \ge 2$) 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 $(L_1[A], \ldots, L_n[A])$ (writing $\mathbf{L}[a]$ rather than $\mathbf{L}[\{a\}]$ for the singleton case). For each label $\mathbf{x} \in \{in, out, undec\}$ and $a \in Args_{\mathcal{A}}$ we denote the set of agents who voted for label \mathbf{x} for a by $V_{a:\mathbf{x}}^{\mathbf{L}}$, i.e., $V_{a:\mathbf{x}}^{\mathbf{L}} = \{i \in Ag \mid L_i(a) = \mathbf{x}\}.$

The central concept of this paper is the following.

Definition 3 An aggregation method is a function F that assigns to every AF A and A-profile L an A-labelling $F_A(L)$.

Postulates for aggregation methods

We start with some postulates for a good aggregation method. Some are inspired by postulates in (Caminada and Pigozzi 2011; Rahwan and Tohmé 2010), which in turn were inspired by those familiar from the JA literature. We modify them to account for allowing the AF to vary. Note that free occurrences of A and L within the postulates are implicitly universally quantified. 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. However, a very weak version 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 $\rightarrow_{\mathcal{A}} = \{(a, b), (b, a)\}$ for some distinct $a, b \in U$.

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

Given an \mathcal{A} -profile $\mathbf{L} = (L_1, \ldots, L_n)$, we say \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 L' is a permutation of L then $F_{\mathcal{A}}(\mathbf{L}) = F_{\mathcal{A}}(\mathbf{L}')$.

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

The idea behind the next postulate is that AFs that are *isomorphic* should be treated the same when aggregating. Given $\mathcal{A}_1 = (Args_1, \rightarrow_1)$ and $\mathcal{A}_2 = (Args_2, \rightarrow_2)$, an isomorphism from \mathcal{A}_1 to \mathcal{A}_2 is a bijection $g : Args_1 \rightarrow Args_2$ such that, for all $a, b \in Args_1$ we have $a \rightarrow_1 b$ iff $g(a) \rightarrow_2 g(b)$. Such a g 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 A_1 and A_2 are connected by isomorphism g. Then, for any A_1 -profile \mathbf{L} we have $g(F_{A_1}(\mathbf{L})) = F_{A_2}(g(\mathbf{L})).$

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. 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 more commonly used version of Independence (just put $A_1 = A_2$). 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 A.

Our first monotonicity postulate, in/out-Monotonicity, says that if some agents change their individual labels of some arguments in profile L so that they agree with the collective labelling $F_{\mathcal{A}}(\mathbf{L})$, assuming those collective labels are in {in, 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 $L'_i(a) = [F_{\mathcal{A}}(\mathbf{L})](a) \in \{\text{in, out}\}$). Then $F_{\mathcal{A}}(\mathbf{L}') = F_{\mathcal{A}}(\mathbf{L})$.

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

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

The next postulate says the collective label on any argument never goes against the individual label of any agent (Caminada and Pigozzi 2011).

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 $[F_{\mathcal{A}}(\mathbf{L})](a) =$ undec.

Given any *n*-tuple (l_i) of labels the in/out-winner in (l_i) is the label among {in, out} which appears more frequently in (l_i) (if such a label exists). E.g. the in/out-winner in (in, undec, out, undec, in) is in. If x is the in/out-winner in (l_i) then we call $\neg x$ 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 $[F_{\mathcal{A}}(\mathbf{L})](a) \neq \mathbf{x}$

Proposition 1 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.

Interval aggregation methods

Now we describe the family of *interval aggregation methods*, which will include a number of interesting special cases

(and which are closely-related to the quota rules considered in JA by (Dietrich and List 2007)). Formally, let Int_n be the set of *intervals* of non-zero length in $\{0, 1, ..., n\}$ (recall n is the number of agents), i.e., $Int_n = \{(k,l) \mid k < l\}$ $l, k, l \in \{0, 1, ..., n\}$. 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 L and $a \in Args_A$:

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

Definition 4 An interval aggregation method is an aggregation method F such that $F = F^Y$ for some $Y \subseteq Int_n$ satisfying (I1): $(0, n) \in Y$.

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

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

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

Sceptical initial: (Caminada and Pigozzi 2011) 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: (Caminada and Pigozzi 2011) Take the in/out-winner x whenever no agent voted for $\neg x$, otherwise undec. $Y_{\text{Cred}} = \{(0, l) \in Int_n \mid l \ge 1\}$. We use F^{Cred} to denote $F^{\tilde{Y}_{Cred}}$.

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

We obtain the following axiomatic characterisation.

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 and 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 say the following:

Proposition 2 (i). There exists an interval method that does not satisfy in/out-Monotonicity. (ii). F^{AWP} , F^{Maj} , F^{Scept} , F^{Cred} and F^{iow} all satisfy

in/out-Monotonicity.

We 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 $s \leq k, l \leq t$ then $(s, t) \in Y$.

Proposition 3 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.

Putting Thm. 1 and Prop. 3 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 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 4 Let F^Y be an interval method. Then F^Y satisfies Compatibility iff Y satisfies (I3).

Clearly, among our examples, $Y_{\rm Scept}$ and $Y_{\rm Cred}$ 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. 3 (and recalling the facts about Compatibility in Prop. 1) 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 $\{(0, t) \mid t \ge l\}$ for some $1 \leq l \leq n.$

(ii). F satisfies Minimal Collective Rationality, Anonymity, Unanimity, Isomorphism, AF-Independence, Compatibility and in/out-Monotonicity

Regarding Collective Rationality, we know already from (Caminada and Pigozzi 2011; Rahwan and Tohmé 2010) that our examples of interval methods above fail to satisfy it. Is there any requirement we can place on Y to ensure it? Unfortunately 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 (List and Pettit 2002)) shows.

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

Thus, given the basic requirements Isomorphism, Anonymity and Unanimity, there is no hope to obtain both of AF-Independence and Collective Rationality. We now look at relaxing AF-Independence.

Weakening AF-Independence

One might argue that AF-Independence cannot be expected to hold when part of the input to the aggregation explicitly contains information (in the form of the attack relation $\rightarrow_{\mathcal{A}}$) regarding dependencies between arguments. Instead we might expect the following weaker version, inspired by a similar postulate for argumentation semantics from (Baroni and Giacomin 2007). The idea is that if we have a set of arguments in \mathcal{A} that is *unattacked* then we can aggregate just that part without looking at the arguments outside the set. Note $\mathcal{A} \subseteq_{\mathrm{f}} \mathcal{A}'$ indicates that $Args_{\mathcal{A}} \subseteq Args_{\mathcal{A}'}$ and $\rightarrow_{\mathcal{A}} = \rightarrow_{\mathcal{A}'} \cap (Args_{\mathcal{A}} \times Args_{\mathcal{A}})$.

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 5 Every aggregation method F that satisfies AF-Independence also satisfies Directionality.

Can we construct an aggregation method that satisfies *Directionality, Collective Rationality* and some other desirable postulates? We show the answer is yes, using the *down-admissible* and *up-complete* constructions of (Caminada and Pigozzi 2011). We begin with the down-admissible construction, which uses the definition of the 'committedness' relation \sqsubseteq according to which $L_1 \sqsubseteq L_2$ iff both $L_1^{-1}(in) \subseteq L_2^{-1}(in)$ and $L_1^{-1}(out) \subseteq L_2^{-1}(out)$.

Definition 6 ((Caminada and Pigozzi 2011)) Given an \mathcal{A} labelling L, the down-admissible labelling of L, denoted by $\downarrow L$, is the (unique) greatest element (under \sqsubseteq) of the set of all admissible \mathcal{A} -labellings M such that $M \sqsubseteq L$.

As described in (Caminada and Pigozzi 2011), it can be arrived at by iteratively relabelling every argument that is illegally in or illegally out with undec until no illegal in or out labels remain. The result is a labelling that is admissible, but not necessarily complete. To ensure a complete labelling we need to additionally apply the up-complete operator.

Definition 7 ((Caminada and Pigozzi 2011)) Given an admissible A-labelling L, the up-complete labelling of L, denoted by $\uparrow L$, is the (unique) smallest element (under \sqsubseteq) of the set of all complete A-labellings M such that $L \sqsubseteq M$.

To obtain |L we iteratively change every illegally undec argument to in or out as appropriate (Caminada and Pigozzi 2011). We denote by ||L| the composite operation of performing the down-admissible followed by the up-complete procedures on an A-labelling L.

Definition 8 Given any aggregation method F, the DAUC version of F is the aggregation method \hat{F} defined by setting, for any AF A and A-labelling profile \mathbf{L} , $\hat{F}_{A}(\mathbf{L}) =$ $\parallel (F_{A}(\mathbf{L})).$

For the special cases of interval methods F^{Scept} and F^{Cred} this procedure was studied in detail in (Caminada and Pigozzi 2011). Their DAUC versions were called the *sceptical* and *super-credulous* aggregation methods respectively there. We lose *AF-Independence* as expected. But we can show that some postulates satisfied by the initial method F can be *inherited* by \hat{F} .

Proposition 6 Let F be any aggregation method. For each of the following postulates, if F satisfies that postulate then so does \hat{F} : Anonymity, Unanimity, Isomorphism, Directionality, Compatibility.

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

Hence we have established that, for every interval method F, \hat{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 Props. 4 and 6 we know that if Y satisfies (**I3**) then \hat{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 7 Let F^Y be an interval method. The \hat{F}^Y satisfies in/out-Plurality iff Y satisfies (I3).

One last question concerns the circumstances under which \hat{F}^Y will satisfy (*Strong*) in/out-*Monotonicity*. Since for interval methods we have that *Strong* in/out-*Monotonicity* holds iff Y is widening, one might expect that an analogous equivalence is preserved for the class of DAUC versions of the interval methods. However this remains open for now.

Acknowledgments

Booth is supported by FNR (DYNGBaT project). Awad benefited from a SINTELNET grant. Thanks to Martin Caminada and anonymous referees for comments.

References

Baroni, P., and Giacomin, M. 2007. On principle-based evaluation of extension-based argumentation semantics. *Artificial Intelligence* 171(10):675–700.

Baroni, P.; Caminada, M.; and Giacomin, M. 2011. An introduction to argumentation semantics. *Knowledge Engineering Review* 26(4):365–410.

Caminada, M., and Pigozzi, G. 2011. On judgment aggregation in abstract argumentation. *Autonomous Agents and Multi-Agent Systems* 22(1):64–102.

Caminada, M. 2006. On the issue of reinstatement in argumentation. In *Proc. JELIA*, 111–123.

Dietrich, F., and List, C. 2007. Judgment aggregation by quota rules: Majority voting generalized. *Journal of Theoretical Politics* 19(4):391–424.

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

List, C., and Pettit, P. 2002. Aggregating sets of judgments: An impossibility result. *Economics and Philosophy* 18(01):89–110.

List, C., and Puppe, C. 2009. Judgment aggregation: A survey. In *Handbook of Rational and Social Choice*. Oxford University Press.

Rahwan, I., and Tohmé, F. 2010. Collective argument evaluation as judgement aggregation. In *Proc. AAMAS*, 417–424.