AUTHOR/EDITOR QUERIES

Article ID: LOGCOM-exv055			
Please respond to all queries and send any additional proof corrections. Failure to do so could result in delayed publication			
Query No	**Section**	**Paragraph**	**Query**
Q1	Author names		Please check that all names have been spelled correctly and appear in the correct order. Please also check that all initials are present. Please check that the author surnames (family name) have been correctly identified by a pink background. If this is incorrect, please identify the full surname of the relevant authors. Occasionally, the distinction between surnames and forenames can be ambiguous, and this is to ensure that the authors' full surnames and forenames are tagged correctly, for accurate indexing online. Please also check all author affiliations.
Q2			Please check the suggested short article title.
Q3			Please provide keywords for this article.
Q4	Figures		Figures have been placed as close as possible to their first citation. Please check that they have no missing sections and that the correct figure legend is present.

MAKING CORRECTIONS TO YOUR PROOF

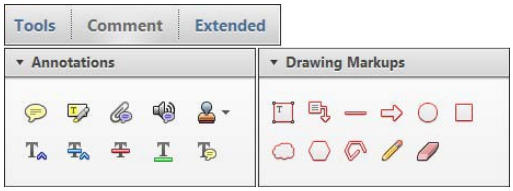
NOT FOR PUBLIC RELEASE

These instructions show you how to mark changes or add notes to the document using the Adobe Acrobat Professional version 7 (or onwards) or Adobe Reader X (or onwards). To check what version you are using go to **Help** then **About**.
The latest version of Adobe Reader is available for free from get.adobe.com/reader.

Displaying the toolbars

Adobe Professional X, XI and Reader X, XI

Select **Comment, Annotations and Drawing Markups.**
If this option is not available, please let me know so that I can enable it for you.

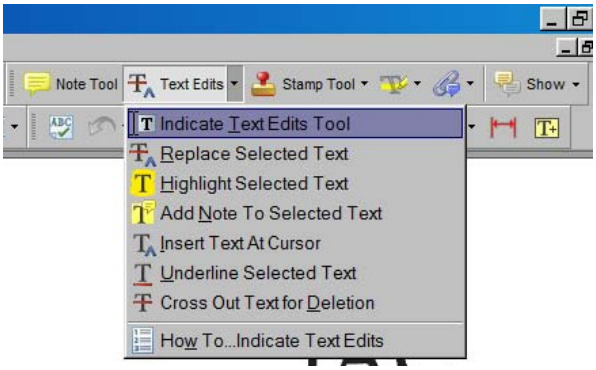


Acrobat Professional 7, 8 and 9

Select **Tools, Commenting, Show Commenting Toolbar**.

Using Text Edits

This is the quickest, simplest and easiest method both to make corrections, and for your corrections to be transferred and checked.

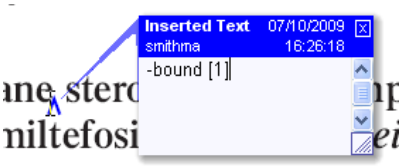


1. Click **Text Edits**
2. Select the text to be annotated or place your cursor at the insertion point.
3. Click the **Text Edits** drop down arrow and select the required action.

   You can also right click on selected text for a range of commenting options.

Pop up Notes

With *Text Edits* and other markup, it is possible to add notes. In some cases (e.g. inserting or replacing text), a pop-up note is displayed automatically.



To **display** the pop-up note for other markup, right click on the annotation on the document and selecting **Open Pop-Up Note**.

To **move** a note, click and drag on the title area.

To **resize** of the note, click and drag on the bottom right corner.

To **close** the note, click on the cross in the top right hand corner.

To **delete** an edit, right click on it and select **Delete**. The edit and associated note will be removed.

SAVING COMMENTS

In order to save your comments and notes, you need to save the file (**File, Save**) when you close the document.
A full list of the comments and edits you have made can be viewed by clicking on the Comments tab in the bottom-left-hand corner of the PDF.

Copyedited by: ES MANUSCRIPT CATEGORY: Corner Article

Judgement aggregation in multi-agent argumentation

Q1 EDMOND AWAD, *Masdar Institute of Science & Technology, UAE.*
E-mail: xxx

RICHARD BOOTH, *Mahasarakham University, Thailand.*
E-mail: xxx

FERNANDO TOHMÉ, *Universidad Nacional del Sur, Argentina.*
E-mail: xxx

IYAD RAHWAN, *Masdar Institute of Science & Technology, UAE; University of Edinburgh, UK; MIT, USA*
E-mail: irahwan@acm.org

Abstract

Given a set of conflicting arguments, there can exist multiple plausible opinions about which arguments should be accepted, rejected or deemed undecided. We study the problem of how multiple such judgements can be aggregated. We define the problem by adapting various classical social-choice-theoretic properties for the argumentation domain. We show that while argument-wise plurality voting satisfies many properties, it fails to guarantee the *collective rationality* of the outcome. We then present more general results, proving multiple impossibility results on the existence of *any* good aggregation operator. After characterizing the sufficient and necessary conditions for satisfying collective rationality, we study whether restricting the domain of argument-wise plurality voting to classical semantics allows us to escape the impossibility result. We close by mentioning a couple of graph-theoretical restrictions under which the argument-wise plurality rule does produce collectively rational outcomes. In addition to identifying fundamental barriers to collective argument evaluation, our results contribute to research at the intersection of the argumentation and computational social choice fields.

Q3 *Keywords*: xxx

1 Introduction

Argumentation has recently become one of the key approaches to automated reasoning and rational interaction in Artificial Intelligence [5, 28]. A key milestone in the development of argumentation in AI has been Dung's landmark framework [15], known as abstract argumentation framework (AAF). Arguments are viewed as abstract entities (a set $\mathcal{A}$), with a binary *defeat* relation (denoted $\rightharpoonup$) over them. The defeat relation captures the fact that one argument somehow attacks or undermines another. This view of argumentation enables high-level analysis while abstracting away from the internal structure of individual arguments. In Dung's approach, given a set of arguments and a defeat relation, a rule specifies which arguments should be accepted.

Often, there are multiple reasonable ways in which an agent may evaluate a given argument structure (e.g. accepting only conflict-free, self-defending sets of arguments). Each possible

doi:10.1093/logcom/exv055

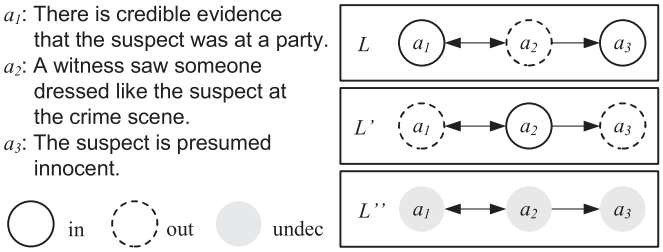


FIGURE 1. Argument graph with three possible labellings.

evaluation corresponds to a so-called *extension* [15] or *labelling* [8, 9]. Different argumentation semantics yield different restrictions on the possible extensions. Most previous research has focused on evaluating and comparing different semantics based on the (objective) logical properties of their extensions [3].

One of the essential properties, which is common, is the condition of *admissibility*: that accepted arguments must not attack one another, and must defend themselves against counter-arguments, by attacking them back. A stronger notion is called *completeness*, and is captured, in terms of labelling, in the following two conditions:

(1) An argument is labelled *accepted* (or `in`) if and only if all its defeaters are rejected (or `out`).
(2) An argument is labelled *rejected* (or `out`) if and only if at least one of its defeaters is accepted (or `in`).

Otherwise, an argument may be labelled `undec`. Thus, evaluating a set of arguments amounts to labelling each argument using a labelling function $L : \mathcal{A} \to \{\texttt{in}, \texttt{out}, \texttt{undec}\}$ to capture these three possible labels. Any labelling that satisfies the above conditions is also called a *legal labelling*. We will often use *legal labelling* and *complete labelling* interchangeably.

The above conditions attempt to evaluate arguments from a single point of view. Indeed, most research on formal models of argumentation discounts the fact that argumentation takes place among self-interested agents, who may have conflicting opinions and preferences over which arguments end up being accepted, rejected or undecided. Consider the following simple example.

EXAMPLE 1 (A Murder Case)
A murder case is under investigation. To start with, there is an argument that the suspect should be presumed innocent (a_3). However, there is evidence that he may have been at the crime scene at the time (a_2), which would counter the initial presumption of innocence. There is also, however, evidence that the suspect was attending a party that day (a_1). Clearly, a_1 and a_2 are mutually defeating arguments since the suspect can only be in one place at any given time. Hence, we have a set of arguments $\{a_1, a_2, a_3\}$ and a defeat relation $\rightharpoonup = \{(a_1, a_2), (a_2, a_1), (a_2, a_3)\}$. There are three possible labellings that satisfy the above conditions:

- $L(a_1) = \texttt{in}, L(a_2) = \texttt{out}, L(a_3) = \texttt{in}$.
- $L'(a_1) = \texttt{out}, L'(a_2) = \texttt{in}, L'(a_3) = \texttt{out}$.
- $L''(a_1) = \texttt{undec}, L''(a_2) = \texttt{undec}, L''(a_3) = \texttt{undec}$.

The graph and possible labellings are depicted in Figure 1.

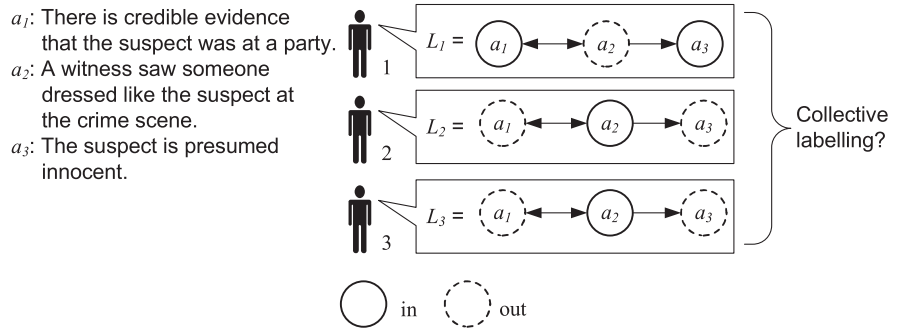


FIGURE 2. Three detectives with different judgements.

Example 1 highlights a situation in which multiple points of view can be taken, depending on whether one decides to accept the argument that the suspect was at the party or the crime scene. The question we explore in this article can be highlighted through the following example, extending Example 1.

EXAMPLE 2 (Three Detectives)
A team of three detectives, named 1, 2 and 3, have been assigned to the murder case described in Example 1. Each detective's judgement can only correspond to a legal labelling (otherwise, her judgement can be discarded). Suppose that each detective's judgement is such that $L_1 = L$, $L_2 = L'$ and $L_3 = L'$. That is, detectives 2 and 3 agree but differ with detective 1. These labellings are depicted in the labelled graph of Figure 2. The detectives must decide which (aggregated) argument labelling best reflects their collective judgement.

Example 2 highlights an aggregation problem, similar to the problem of preference aggregation [2, 16, 33] and the problem of judgement aggregation on propositional formulae [19, 20, 22, 23]. It is perhaps obvious in this particular example that a_3 must be rejected (and thus the defendant be considered guilty), since most detectives seem to think so. For the same reason, a_1 must be rejected and a_2 must be accepted. Thus, labelling L' (see Example 1) wins by majority. As we shall see in our analysis below, things are not that simple, and counter-intuitive situations may arise. We summarize the main question asked in the article as follows.

> Given a set of agents, each with a specific subjective evaluation (i.e. labelling) of a given set of conflicting arguments, how can agents reach a collective decision on how to evaluate those arguments?

While Arrow's Impossibility Theorem can be expected to ensue for this problem [1],[1] there exist many differences between labellings and preference relations (for which Arrow's result apply), stemming from their corresponding order-theoretic characterizations. In other words, aggregating preferences assumes that agents submit a full order of preferences over candidates, while in labelling aggregation, agents submit their top labelling for a set of logically connected arguments.

[1]Arrow's Theorem claims that four quite natural constraints, that capture abstractly the properties of a democratic aggregation process, cannot be simultaneously satisfied.

The problem of labelling aggregation is more comparable to the judgement aggregation problem [19, 20, 22, 23], by considering arguments as propositions which are logically connected by the conditions of legal labelling. However, one important difference is that in judgement aggregation, each proposition can have two values: True or False. In labelling aggregation, on the other hand, each argument can have three values: `in`, `out` or `undec`. This makes labelling aggregation be more comparable to non-binary evaluations [13, 14]. Considering the general framework in [14], our settings can be considered as focusing on special classes of feasible evaluations, which are the conditions imposed by the legal labelling (or other semantics). Additionally, the possible evaluations of each issue (argument, in our case) are to accept (labels as `in`), reject (labels as `out`) or be undecided (labels as `undec`).

In this article, we conduct an extensive social-choice-theoretic analysis of argument evaluation semantics by means of labellings. We assume that individuals are presented with a shared argumentation framework (AF) and need to make a decision about how to evaluate this AF. Individuals are assumed to have different, but reasonable, evaluations. There can be many scenarios in which such settings are present. For example, consider a jury members that are all provided with the same information, each of them has a different opinion about these information and yet they all need to come up with a collective decision. Another example is a company board committee who need to make an informed decision. They can be all presented with the same information about the current economic status and the possible strategies, each one of them has his/her own opinion about what should be done, yet they all need to reach a collective decision.

The article makes three distinct contributions to the state-of-the-art in the computational modelling of argumentation. First, the article introduces the study of aggregating different individual judgements on how a given set of arguments is to be evaluated.[2] This requires adapting classical social-choice properties to the argumentation domain, and sometimes demands special treatment (e.g. different versions of some properties).

The second contribution of this article is proving the impossibility of the existence of *any* aggregation operator that satisfies some minimal properties. In doing so, we show impossibility results that concern dealing with ties and producing a collectively rational evaluation of arguments. These results establish the limits of aggregation in the context of argumentation, and come in accordance with the impossibility results in the topics of aggregation such as preference aggregation [1, 17, 25, 30, 31] and judgement aggregation [21]. Hence, as is the case with other aggregation domains, the aggregation paradox in argument evaluation is an example of a more fundamental barrier. These results are important because they give conclusive answers and focus research in more constructive directions (e.g. weakening the desired properties in order to avoid the paradox). Aiming to investigate possible relaxations in order to circumvent the impossibility in the context of argumentation, we broke down the *Collective Rationality* postulate into sub-postulates. This helps in taking a deeper look at the distinct parts of the postulate. As a consequence, satisfying any of these parts can be used to weaken the collective rationality.

The third contribution of this article is an extensive analysis of an aggregation rule, namely *argument-wise plurality rule*. We analyse the properties of the argument-wise plurality rule in general,

[2] In fact, this idea was first introduced in [29] for which this article is a substantially extended and revised version. Section 6 which introduces the impossibility of good aggregation operator is significantly enhanced by adding three impossibility results. Sections 8 and 9 are completely new. Section 10 contains more elaborate discussion of related and future work. Finally, further explanation, motivation, discussion and background is added to the other sections to improve clarity and presentation of the article.

Copyedited by: ES MANUSCRIPT CATEGORY: Corner Article

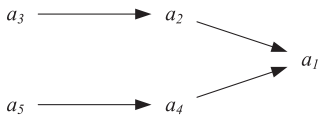


FIGURE 3. A simple argument graph.

and investigate whether the restriction of the domain of votes to a particular classical semantics would ensure the fulfillment of these conditions. This highlights a novel use of classical semantics, which are originally used to resolve issues in single-agent non-monotonic reasoning. Finally, we provide graph-theoretical restrictions on argumentation frameworks under which the argument-wise plurality rule would be guaranteed to produce collectively rational outcomes.

The article is organized as follows. In Section 2, we start by giving a brief background on abstract argumentation systems. Sections 3, 4, 6 and 7 focus on the problem of aggregating sets of judgements over argument evaluation. Sections 5, 8 and 9 focus on introducing and analysing the argument-wise plurality rule. We conclude the article and discuss some related work in Section 10.

2 Background

In this section, we briefly outline key elements of abstract argumentation frameworks. We begin with Dung's abstract characterization of an argumentation system [15]. We restrict ourselves to finite sets of arguments.

DEFINITION 1 (Argumentation framework)
An *argumentation framework* is a pair $AF = \langle \mathcal{A}, \rightharpoonup \rangle$ where $\mathcal{A}$ is a finite set of arguments and $\rightharpoonup \subseteq \mathcal{A} \times \mathcal{A}$ is a defeat relation. We say that an argument a *defeats* an argument b if $(a, b) \in \rightharpoonup$ (sometimes written $a \rightharpoonup b$).

For an argument $a \in \mathcal{A}$, we use a^- to denote the set of arguments that defeat a i.e. $a^- = \{b \in \mathcal{A} | b \rightharpoonup a\}$.

An argumentation framework can be represented as a directed graph in which vertices are arguments and directed arcs characterize defeat among arguments. An example of argument graph is shown in Figure 3. Argument a_1 has two defeaters (i.e. counter-arguments) a_2 and a_4, which are themselves defeated by arguments a_3 and a_5, respectively.

There are two approaches to define semantics that assess the acceptability of arguments. One of them is extension-based semantics by Dung [15], which produces a set of arguments that are accepted together. Another equivalent labelling-based semantics is proposed by Caminada [8, 9], which gives a labelling for each argument. With argument labellings, we can accept arguments (by labelling them as `in`), reject arguments (by labelling them as `out`) and abstain from deciding whether to accept or reject (by labelling them as `undec`). Caminada [8, 9] established a correspondence between properties of labellings and the different extensions. In this article, we employ the labelling approach.

DEFINITION 2 (Argument Labelling)
Let $AF = \langle \mathcal{A}, \rightharpoonup \rangle$ be an argumentation framework. An *argument labelling* is a total function $L : \mathcal{A} \to \{\texttt{in}, \texttt{out}, \texttt{undec}\}$.

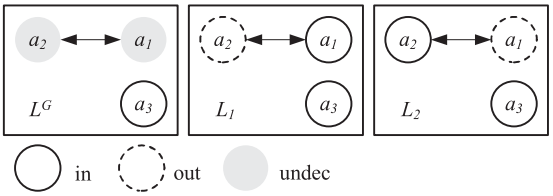


FIGURE 4. Graph with three complete labellings.

We write $\texttt{in}(L)$ (resp. $\texttt{out}(L)$, $\texttt{undec}(L)$) for the set of arguments that are labelled $\texttt{in}$ (resp. $\texttt{out}$, $\texttt{undec}$) by L. A labelling L can be represented as $L=(\texttt{in}(L),\texttt{out}(L),\texttt{undec}(L))$.

However, labellings should follow some given conditions. A minimal reasonable condition is the *conflict-freeness*.

DEFINITION 3 (Conflict-freeness)
A labelling L satisfies conflict-freeness iff $\forall a,b\in\texttt{in}(L)$, $\neg(a\rightharpoonup b)$.

One of the essential semantics, which satisfies conflict-freeness is the *complete semantics*. We already informally defined *complete* labellings via two conditions in the introduction. We find it convenient to equivalently formulate it as three conditions as follows.

DEFINITION 4 (Complete labelling)
Let $AF=\langle\mathcal{A},\rightharpoonup\rangle$ be an argumentation framework. A *complete labelling* is a total function $L:\mathcal{A}\rightarrow\{\texttt{in},\texttt{out},\texttt{undec}\}$ such that:

- $\forall a\in\mathcal{A}$: if $L(a)=\texttt{in}$ then $\forall b\in\mathcal{A}$: $(b\rightharpoonup a\Rightarrow L(b)=\texttt{out})$;
- $\forall a\in\mathcal{A}$: if $L(a)=\texttt{out}$ then $\exists b\in\mathcal{A}$ s.t. $(b\rightharpoonup a\wedge L(b)=\texttt{in})$; and
- $\forall a\in\mathcal{A}$: if $L(a)=\texttt{undec}$ then
  - $\exists b\in\mathcal{A}$: $(b\rightharpoonup a\wedge L(b)=\texttt{undec})$; and
  - $\nexists b\in\mathcal{A}$: $(b\rightharpoonup a\wedge L(b)=\texttt{in})$

We will use $Comp(AF)$ to denote the set of all *complete* labellings for AF.

As an example, consider the following.

EXAMPLE 3
Consider the graph in Figure 4. Here, we have three complete labellings: $L^G=(\{a_3\},\{\},\{a_1,a_2\})$, $L_1=(\{a_1,a_3\},\{a_2\},\{\})$ and $L_2=(\{a_2,a_3\},\{a_1\},\{\})$.

In addition to the *complete* labelling, there are other semantics which assume further conditions.

Copyedited by: ES MANUSCRIPT CATEGORY: Corner Article

DEFINITION 5 (Other Labelling Semantics)
Let $AF = \langle \mathcal{A}, \rightharpoonup \rangle$ be an argumentation framework. Let $L : \mathcal{A} \to \{\texttt{in}, \texttt{out}, \texttt{undec}\}$ be a complete labelling.

- L is a grounded labelling if and only if $\texttt{in}(L)$ is minimal, or equivalently $\texttt{out}(L)$ is minimal, or equivalently $\texttt{undec}(L)$ is maximal (w.r.t set inclusion) among all *complete* labellings.
- L is a preferred labelling if and only if $\texttt{in}(L)$ is maximal, or equivalently $\texttt{out}(L)$ is maximal (w.r.t set inclusion) among all *complete* labellings.
- L is a semi-stable labelling if and only if $\texttt{undec}(L)$ is minimal (w.r.t set inclusion) among all *complete* labellings.
- L is a stable labelling if and only if $\texttt{undec}(L) = \emptyset$.

Note that the grounded labelling is always unique, and stable labellings might not exist. Consider the following example.

EXAMPLE 4
Consider the graph in Figure 4. Here, we have the *grounded* labelling is $L^G = (\{a_3\}, \{\}, \{a_1, a_2\})$. We have only two *preferred* labellings: $L_1 = (\{a_1, a_3\}, \{a_2\}, \{\})$, and $L_2 = (\{a_2, a_3\}, \{a_1\}, \{\})$. These are also the only *stable* and *semi-stable* labellings for this framework.

Clearly, for any AF, $Stab(AF) \subseteq Semi(AF) \subseteq Pref(AF) \subseteq Comp(AF)$, and $Grnd(AF) \subseteq Comp(AF)$, where $Stab(AF)$, $Semi(AF)$, $Pref(AF)$ and $Grnd(AF)$ refer to the set of *stable*, *semi-stable*, *preferred* and *grounded* labellings for AF. We refer to the previous semantics as *classical semantics*. There exist other semantics which we do not consider in this work.

3 Aggregation of argument labellings

To date, most analyses inspired by Dung's framework have focused on analysing and comparing the properties of various types of extensions/labellings (i.e. semantics) [3]. The question is, therefore, whether a particular type of labelling is appropriate for a particular type of reasoning task in the presence of conflicting arguments.

In contrast with most existing work on Dung frameworks, our concern here is with multi-agent systems. Since each labelling captures a particular rational point of view, we ask the following question: *Given an argumentation framework and a set of agents, each with a legitimate subjective evaluation of the given arguments, how can the agents reach a collective compromise on how to evaluate those arguments?*

Thus, the problem we face is that of *judgement aggregation* [21] in the context of argumentation frameworks. This problem can be formulated as a set of individuals that collectively decide how an argumentation framework $AF = \langle \mathcal{A}, \rightharpoonup \rangle$ must be labelled.

DEFINITION 6 (Labelling aggregation problem)
Let $Ag = \{1, \ldots, n\}$ be a finite non-empty set of agents, and $AF = \langle \mathcal{A}, \rightharpoonup \rangle$ be an argumentation framework. A labelling aggregation problem is a pair $\mathcal{LAP} = \langle Ag, AF \rangle$.

Each individual $i \in Ag$ has a labelling L_i which expresses the evaluation of AF by this individual. A labelling profile is an $|Ag|$-tuple of labellings.

DEFINITION 7 (Labelling profile)
Let $\mathcal{LAP} = \langle Ag, AF \rangle$ be a labelling aggregation problem. We use $\mathcal{L} = (L_1, \ldots, L_n) \in \mathbf{L}(AF)^{|Ag|}$ to denote a labelling profile, where $\mathbf{L}(AF)$ is the class of labellings of AF. Additionally, we use $\mathcal{L}(a)$ to denote the labelling profile (i.e. an $|Ag|$-tuple) of an argument $a \in \mathcal{A}$ i.e. $\mathcal{L}(a) = (L_1(a), \ldots, L_n(a))$.

The aggregation of individuals' labellings can be defined as a partial function.[3]

DEFINITION 8 (Aggregation function)
Let $\mathcal{LAP} = \langle Ag, AF \rangle$ be a labelling aggregation problem. An aggregation function for $\mathcal{LAP}$ is a function $F : \mathbf{L}(AF)^n \to \mathbf{L}(AF)$.

For each $a \in \mathcal{A}$, $[F(\mathcal{L})](a)$ denotes the collective label assigned to a, if F is defined for $\mathcal{L} = (L_1, \ldots, L_n)$.

4 Desirable properties of aggregation operators

Aggregation involves comparing and assessing different points of view. There are, of course, many ways of doing this, as extensively discussed in the literature of Social Choice Theory [16]. In this literature, a consensus on some normative ideals has been reached, identifying what a 'fair' way of adding up votes should be. So for instance, if everybody agrees, the outcome must reflect that agreement; no single agent can impose her view on the aggregate; the aggregation should be performed in the same way in each possible case, *etc*. These informal requirements can be formally stated as properties that F should satisfy [12, 21]. In all of the following postulates, it is assumed that a fixed labelling aggregation problem $\mathcal{LAP} = \langle Ag, AF \rangle$ is given. The postulates can be grouped as follows:[4]

Group 1: Domain and co-domain postulates

In judgement aggregation, two postulates that are commonly assumed are those of *Universal Domain* and *Collective Rationality*. The former requires that any profile of labellings chosen from a pre-specified set of *feasible* labellings can be used as input to F and F will return an answer. The question is: what do we take to be the set of feasible labellings in our setting? This depends on which semantics we assume is being used. Theoretically, we can have a different version of *Universal Domain* for each semantics. However since *complete* semantics represent reasonable and self-defending points of views, it represents the best counterpart for the logical consistency in judgement aggregation:

Universal Domain F can take as input all profiles $\mathcal{L} = (L_1, \ldots, L_n)$ such that $\mathcal{L} \in Comp(AF)^n$

However, in Subsection 8.2 we will use other semantics as a domain for $\mathcal{L}$.

Similarly we could have a different version of *Collective Rationality*—one for each semantics—stating that the output of the aggregation should also be feasible. Again, since we focus on complete semantics, we focus on the following version:

Collective Rationality For all profiles $\mathcal{L}$ such that $F(\mathcal{L})$ is defined, $F(\mathcal{L}) \in Comp(AF)$.

[3] We state that the function is partial to allow for cases in which collective judgement may be undefined (e.g. when there is a tie in voting).

[4] This style of presentation of postulates was inspired by [18] which is on binary aggregation.

Copyedited by: ES MANUSCRIPT CATEGORY: Corner Article

Later, in Section 7, we will break this postulate down into further constituents.

Group 2: Fundamental postulates

Next we come to the standard property that forms the cornerstone of the usual impossibility results in judgement aggregation. It says the collective label of an argument depends only on the votes on that argument, independent of the other arguments.

Independence For any two profiles $\mathcal{L}=(L_1,\ldots,L_n)$, $\mathcal{L}'=(L'_1,\ldots,L'_n)$ such that $F(\mathcal{L})$ and $F(\mathcal{L}')$ are defined, and for all $a\in\mathcal{A}$, if $L_i(a)=L'_i(a)$ for all $i\in Ag$, then $[F(\mathcal{L})](a)=[F(\mathcal{L}')](a)$.

The effect of *Independence* is that aggregation is done 'argument-by-argument'. To be slightly more precise, each argument $a\in\mathcal{A}$ essentially has its own aggregation operator I_a associated with it, that takes an n-tuple of labels $\mathbf{x}=(l_1,\ldots,l_n)$ as input (representing the 'vote' of each agent on the label of a) and returns another label $I_a(\mathbf{x})$ as output (the 'collective label') of a. Then $[F(\mathcal{L})](a)=I_a((L_1(a),\ldots,L_n(a)))$. Note that the necessity of *Independence* is questionable in our settings because of the dependencies between arguments that come already encoded in the form of the attack relation. Nevertheless, it is usually investigated in the judgement aggregation and preference aggregation literature because of its role in analysing strategy-proofness. Though the relation between *Independence* and strategy-proofness is not established yet in our settings, our task in this article is to stick close to the methodology in judgement aggregation, and there it is often assumed.

Next, we have *Anonymity*, which says the identity of which agent submits which labelling is irrelevant.

Anonymity For any profile $\mathcal{L}=(L_1,\ldots,L_n)$, if $\mathcal{L}'=(L_{\rho(1)},\ldots,L_{\rho(n)})$ for some permutation ρ on Ag, and $F(\mathcal{L})$ and $F(\mathcal{L}')$ are both defined, then $F(\mathcal{L})=F(\mathcal{L}')$.

If we add *Anonymity* to *Independence*, then it means the outputs of the functions I_a described above depend only on the *number* of votes that each label gets in $\mathbf{x}$. Essentially it means I_a outputs a collective label just taking as input the triple $(\#\mathtt{in},\#\mathtt{out},\#\mathtt{undec})$ of numbers denoting, respectively, the number of votes for `in`, `out` and `undec` in $\mathbf{x}$.

PROPOSITION 1
Let F be an aggregation operator. Then F satisfies both *Independence* and *Anonymity* iff for each $a\in\mathcal{A}$ there exists a function $I_a:\mathbb{N}^3\to\{\mathtt{in},\mathtt{out},\mathtt{undec}\}$ such that, for all $\mathcal{L}$ we have $[F(\mathcal{L})](a)=I_a(\#\mathtt{in},\#\mathtt{out},\#\mathtt{undec})$.

Outline. The 'if' case is straightforward, since permuting the rows does not change the vote distribution and so *Anonymity* will hold. *Independence* is also clear.

For the 'only if' case, *Independence* gives us the existence of the function I_a such that $[F(\mathcal{L})](a)=I_a(L_1(a),\ldots,L_n(a))$ and then *Anonymity* implies that two vectors that have the same vote distribution will give the same results, so we can set $I_a(\#\mathtt{in},\#\mathtt{out},\#\mathtt{undec})=I_a(L_1(a),\ldots,L_n(a))$ where $(L_1(a),\ldots,L_n(a))$ is any vote which has $(\#\mathtt{in},\#\mathtt{out},\#\mathtt{undec})$ as its distribution. ■

A weakening of *Anonymity* is *Non-Dictatorship*:[5]

Non-Dictatorship There is no $i\in Ag$ such that, for every profile $\mathcal{L}=(L_1,\ldots,L_n)$ for which $F(\mathcal{L})$ is defined, we have $F(\mathcal{L})=L_i$.

[5] Since a violation of the latter would imply a violation of the former.

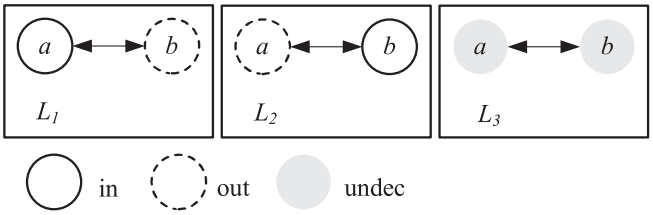


FIGURE 5. An example illustrating *Strong Systematicity*.

Group 3: Unanimity postulates

Next we move to *Unanimity*, and some other postulates related to it.

Unanimity If $\mathcal{L}$ is such that $F(\mathcal{L})$ is defined and there exists some L s.t. $L_i = L$ for all $i \in Ag$, then $F(\mathcal{L}) = L$.

This postulate is also familiar from judgement aggregation, but the move to three-valued labellings rather than the two usually seen in judgement aggregation opens up the possibility to define other variants of *Unanimity*, one of which is used by Dokow and Holzman [14], called *Supportiveness*:

Supportiveness For any profile $\mathcal{L}$ such that $F(\mathcal{L})$ is defined, and for all $a \in \mathcal{A}$, there exists $i \in Ag$ such that $[F(\mathcal{L})](a) = L_i(a)$.

Supportiveness says that, for each argument a and label l, the collective judgement cannot be set to l without at least one agent voting for that l. Clearly *Supportiveness* implies *Unanimity*.

It might seem natural to have the collective label of an argument as `undec` even when nobody votes for it, if we interpret `undec` as a halfway label between `in` and `out`. Then if half the agents say `in` and the other half says `out` then `undec` might be a reasonable compromise. Given this, a weaker version of *Supportiveness* that only applies to `in` and `out` can be defined. We call it `in/out`-*Supportiveness*.

`in/out`-Supportiveness For any profile $\mathcal{L}$ such that $F(\mathcal{L})$ is defined, and for all $a \in \mathcal{A}$, if $[F(\mathcal{L})](a) \neq$ `undec` then there exists some agent i such that $[F(\mathcal{L})](a) = L_i(a)$.

Group 4: Systematicity postulates

Now we come to the *Systematicity* postulates which deal with neutrality issues across arguments and labels. We can list two variants, both of which imply *Independence*. We start with the stronger version:

Strong Systematicity For any two profiles $\mathcal{L} = (L_1, \ldots, L_n)$ and $\mathcal{L}' = (L'_1, \ldots, L'_n)$ such that $F(\mathcal{L})$ and $F(\mathcal{L}')$ are defined, and for all $a, b \in \mathcal{A}$, and for every permutation ρ on the set of labels $\{\texttt{in}, \texttt{out}, \texttt{undec}\}$, if $\forall i \in Ag$: $L_i(a) = \rho(L'_i(b))$, then $[F(\mathcal{L})](a) = \rho([F(\mathcal{L}')](b))$.

To illustrate *Strong Systematicity*, consider the example in Figure 5. We have the following three labellings: $L_1 = (\{a\}, \{b\}, \{\})$, $L_2 = (\{b\}, \{a\}, \{\})$, $L_3 = (\{\}, \{\}, \{a, b\})$.

Consider the profiles $\mathcal{L} = (L_1, L_1, L_2, L_3)$ and $\mathcal{L}' = (L_3, L_3, L_2, L_1)$. Then, $\mathcal{L}(a) = (\texttt{in}, \texttt{in}, \texttt{out}, \texttt{undec})$ and $\mathcal{L}'(b) = (\texttt{undec}, \texttt{undec}, \texttt{in}, \texttt{out})$. Let ρ be the permutation on

Copyedited by: ES MANUSCRIPT CATEGORY: Corner Article

labels such that $\rho(\texttt{in})=\texttt{undec}$, $\rho(\texttt{out})=\texttt{in}$, and $\rho(\texttt{undec})=\texttt{out}$. Then, we can see that in this example $\forall i \in Ag: L'_i(b)=\rho(L_i(a))$. *Strong Systematicity* requires that $[F(\mathcal{L}')](b)=\rho([F(\mathcal{L})](a))$. The postulate forces us to give an even-handed treatment to the labels `in`, `out` and `undec` (in addition to treating each argument independently and similarly). This makes sense if we consider `in`, `out` and `undec` as three independent labels. However, one might be tempted to consider `undec` as a middle label between `in` and `out`. Hence, the equal treatment might not be desirable in this case. One might suggest a version of *Systematicity* that treats `in` and `out` equally. Following, we define this version (which we call `in/out`-*Systematicity*).

`in/out`-Systematicity For any two profiles $\mathcal{L}=(L_1,\ldots,L_n)$ and $\mathcal{L}'=(L'_1,\ldots,L'_n)$ such that $F(\mathcal{L})$ and $F(\mathcal{L}')$ are defined, and for all $a,b \in \mathcal{A}$, and for every `undec`-preserving permutation ρ on the set of labels $\{\texttt{in},\texttt{out},\texttt{undec}\}$ (*i.e.*$\rho(\texttt{undec})=\texttt{undec}$), if $\forall i \in Ag: L_i(a)=\rho(L'_i(b))$, then $[F(\mathcal{L})](a)=\rho([F(\mathcal{L}')](b))$.

`in/out`-*Systematicity* lies in the middle between *Strong Systematicity* and the following version of *Systematicity* which can be obtained by restricting the class of permutations, until we only consider the identity.

Weak Systematicity For any two profiles $\mathcal{L}=(L_1,\ldots,L_n)$ and $\mathcal{L}'=(L'_1,\ldots,L'_n)$ such that $F(\mathcal{L})$ and $F(\mathcal{L}')$ are defined, and for all $a,b \in \mathcal{A}$, if $\forall i \in Ag: L_i(a)=L'_i(b)$, then $[F(\mathcal{L})](a)=[F(\mathcal{L}')](b)$.

Clearly *Independence* follows from *Weak Systematicity* by just setting $a=b$. If we strengthen *Independence* to *Weak Systematicity* then the functions I_a, mentioned earlier, are identical for all arguments.

Group 5: Monotonicity postulates

Our final group relates to *Monotonicity*.

Monotonicity Let $l_a \in \{\texttt{in},\texttt{out},\texttt{undec}\}$ be such that given two profiles $\mathcal{L}=(L_1,\ldots,L_i,\ldots,L_{i+k},\ldots,L_n)$ and $\mathcal{L}'=(L_1,\ldots,L'_i,\ldots,L'_{i+k},\ldots,L_n)$ (differing only in the labellings of agents $i,i+1,\ldots,i+k$) such that $F(\mathcal{L})$ and $F(\mathcal{L}')$ are defined, where $i \in \{1,\ldots,n\}$ and $k \in \{0,\ldots,n-i\}$, if $L_j(a) \neq l_a$ while $L'_j(a)=l_a$ for all $j \in \{i,\ldots,i+k\}$, then $[F(\mathcal{L})](a)=l_a$ implies that $[F(\mathcal{L}')](a)=l_a$.

Monotonicity states that if a set of agents switch their label of argument a to the collective label of a then the collective label of a remains the same. Similar to *Supportiveness* and *Systematicity*, a weaker version of *Monotonicity* that only apply to `in` and `out` can be defined. We call it `in/out`-*Monotonicity*.

`in/out`-Monotonicity Let $l_a \in \{\texttt{in},\texttt{out}\}$ be such that given two profiles $\mathcal{L}=(L_1,\ldots,L_i,\ldots,L_{i+k},\ldots,L_n)$ and $\mathcal{L}'=(L_1,\ldots,L'_i,\ldots,L'_{i+k},\ldots,L_n)$ (differing only in the labellings of agents $i,i+1,\ldots,i+k$) such that $F(\mathcal{L})$ and $F(\mathcal{L}')$ are defined, where $i \in \{1,\ldots,n\}$ and $k \in \{0,\ldots,n-i\}$, if $L_j(a) \neq l_a$ while $L'_j(a)=l_a$ for all $j \in \{i,\ldots,i+k\}$, then $[F(\mathcal{L})](a)=l_a$ implies that $[F(\mathcal{L}')](a)=l_a$.

Copyedited by: ES MANUSCRIPT CATEGORY: Corner Article

5 The argument-wise plurality rule

An obvious candidate aggregation operator to check out is the *plurality* voting operator M. In this section, we analyse a number of key properties of this operator. Intuitively, for each argument, it selects the label that appears most frequently in the individual labellings.

DEFINITION 9 (Argument-Wise Plurality Rule (AWPR))
Let $AF=\langle\mathcal{A},\rightharpoonup\rangle$ be an argumentation framework. Given any argument $a\in\mathcal{A}$ and any profile $\mathcal{L}=(L_1,\ldots,L_n)$, it holds that $[M(\mathcal{L})](a)=l_a\in\{\texttt{in},\texttt{out},\texttt{undec}\}$ iff

$$|\{i:L_i(a)=l_a\}|>\max_{l'_a\neq l_a}|\{i:L_i(a)=l'_a\}|$$

Note that M is defined for all profiles that cause no ties, i.e. $M(\mathcal{L})$ is defined iff there does not exist any argument $a\in\mathcal{A}$ for which we have at least two labels l_a and l'_a with $l_a\neq l'_a$ and

$$|\{i:L_i(a)=l_a\}|=|\{i:L_i(a)=l'_a\}|=\max_{l}|\{i:L_i(a)=l\}|$$

One can directly notice that AWPR violates *Universal Domain*, because it is not defined for all profiles in $Comp(AF)$.

EXAMPLE 5 (Three Detectives (cont.))
Continuing on Example 2, applying the argument-wise plurality rule, we have $[M((L_1,L_2,L_3))](a_1)=\texttt{out}$, $[M((L_1,L_2,L_3))](a_2)=\texttt{in}$, and $[M((L_1,L_2,L_3))](a_3)=\texttt{out}$.

5.1 *Properties of argument-wise plurality rule*

We now analyse whether AWPR satisfies the properties listed above.

PROPOSITION 2
The argument-wise plurality rule operator M satisfies *Supportiveness*, *Anonymity*, *Strong Systematicity*, and *Monotonicity*.

PROOF. In this proof, the considered profiles are restricted to those for which $[M(\mathcal{L})]$ is defined.

- *Supportiveness*: consider any profile $\mathcal{L}=(L_1,\ldots,L_n)$. Suppose, towards a contradiction, that for some argument a, there exists no agent i such that $L_i(a)=l_a$ where $l_a=[M(\mathcal{L})](a)$. Then $|\{i:L_i(a)=l_a\}|=0$. But, $|\{i:L_i(a)=l_a\}|>\max_{l'_a\neq l_a}|\{i:L_i(a)=l'_a\}|>0$ (the last inequality holds since Ag is non-empty). Contradiction.
- *Anonymity*: consider any profile $\mathcal{L}=(L_1,\ldots,L_n)$. $[M(\mathcal{L})](a)=l_a$ if and only if $|\{i:L_i(a)=l_a\}|>\max_{l'_a\neq l_a}|\{i:L_i(a)=l'_a\}|$ if and only if $|\{\rho(i):L_{\rho(i)}(a)=l_a\}|>\max_{l'_a\neq l_a}|\{\rho(i):L_{\rho(i)}(a)=l'_a\}|$, which is equivalent to $[M((L_{\rho(1)},\ldots,L_{\rho(i)},\ldots,L_{\rho(n)}))](a)=l_a$.
- *Strong Systematicity*: consider, for any two profiles $\mathcal{L}=(L_1,\ldots,L_n)$ and $\mathcal{L}'=(L'_1,\ldots,L'_n)$, and for any $a,b\in\mathcal{A}$, the permutation $\rho:\{\texttt{in},\texttt{out},\texttt{undec}\}\to\{\texttt{in},\texttt{out},\texttt{undec}\}$. Suppose, towards a contradiction, that for any i, $L_i(a)=\rho(L'_i(b))$, and $[M(\mathcal{L})](a)=l_a$ but $\rho(M(\mathcal{L}')[b])\neq\rho(l_a)$. But then, $|\{i:L_i(a)=l_a\}|=|\{i:L'_i(b)=\rho(l_a)\}|$ while for any $l'_a\neq l_a$, $|\{i:L_i(a)=l'_a\}|=|\{i:L'_i(b)=\rho(l'_a)\}|$. So, if $|\{i:L_i(a)=l_a\}|>\max_{l'_a\neq l_a}|\{i:L_i(a)=l'_a\}|$ then, we have $|\{i:L'_i(b)=\rho(l_a)\}|>\max_{l'_a\neq l_a}|\{i:L'_i(b)=\rho(l'_a)\}|$ as well. Contradiction.

Copyedited by: ES MANUSCRIPT CATEGORY: Corner Article

- *Monotonicity*: consider the following two profiles $\mathcal{L}=(L_1,\ldots,L_i,\ldots,L_{i+k},\ \ldots,L_n)$ and $\mathcal{L}'=(L_1,\ldots,L'_i,\ldots,L'_{i+k},\ldots,L_n)$ (differing only in the labellings of agents $i,i+1,\ldots,i+k$) where $i\in\{1,\ldots,n\}$ and $k\in\{0,\ldots,n-i\}$. Suppose, towards a contradiction, that for $a\in\mathcal{A}$ and a label l_a we have that $L_h(a)\neq l_a$ while $L'_h(a)=l_a$ for all $h\in\{i,\ldots,i+k\}$, and we have that $[M(\mathcal{L})](a)=l_a$ while $[M(\mathcal{L}')](a)\neq l_a$. But then, $|\{j:L_j(a)=l_a\}|>\max_{l'_a\neq l_a}|\{j:L_j(a)=l'_a\}|$ in the profile $\mathcal{L}$ while in the profile $(\hat{L}_1,\ldots,\hat{L}_n)\equiv\mathcal{L}'$, we have $\{j:\hat{L}_j(a)=l_a\}=\{j:L_j(a)=l_a\}\cup\{i,\ldots,i+k\}$ and $\{j:\hat{L}_j(a)=l'_a\}\subseteq\{j:L_j(a)=l'_a\}$ for every other labelling l'_a. Then $|\{j:\hat{L}_j(a)=l_a\}|>\max_{l'_a\neq l_a}|\{j:\hat{L}_j(a)=l'_a\}|$. Contradiction.

■

COROLLARY 1
The argument-wise plurality rule operator M satisfies *Unanimity*, *Weak Systematicity*, *Independence* and *Non-Dictatorship*.

PROOF. *Weak Systematicity* and *Independence* follow from *Strong Systematicity*, *Unanimity* follows from *Supportiveness* and *Non-Dictatorship* follows from *Anonymity*. ■

Despite all these promising results, it turns out that plurality operator violates *Universal Domain* and *Collective Rationality* postulates. The violation of *Universal Domain* is because AWPR is not defined for profiles that cause ties, which means that it cannot take as input every possible profile $\mathcal{L}\in Comp(AF)^n$. However, a weaker version of *Universal Domain* can be defined.

No-Tie Universal Domain An aggregation operator F can take as input all profiles $\mathcal{L}=(L_1,\ldots,L_n)$ such that $\mathcal{L}$ does not cause a tie and $\mathcal{L}\in Comp(AF)^n$.

Since there are no restrictions (other than having no ties) on how labellings are defined, AWPR satisfies *No-Tie Universal Domain*. Note that one might be tempted to make AWPR satisfy *Universal Domain* by adding a deterministic[6] tie-breaking rule to deal with ties. However, as we show in the next section, the use of any tie-breaking rule would result in violating *Anonymity*, and/or *Strong Systematicity*. While the violation of *Universal Domain* represents a minor inconvenience that can be justified, the violation of *Collective Rationality* poses a serious issue as the collective decision is usually expected to be reasonable. The following example shows how AWPR violates *Collective Rationality*.

EXAMPLE 6
Suppose argument c has two defeaters, a and b, and argument a (resp. b) defeats and is defeated by argument a' (resp. b'). Suppose we have three agents, with votes as shown in Figure 6. We have $[M(\mathcal{L})](c)=\texttt{out}$, but it is not the case that $[M(\mathcal{L})](a)=\texttt{in}$ or $[M(\mathcal{L})](b)=\texttt{in}$.

Interestingly, the above counterexample demonstrates a variant of the discursive dilemma [21] in the context of argument evaluation, which itself is a variant of the well-known Condorcet paradox.

6 The impossibility of good aggregation operators

In the previous section, we analyzed a particular judgement aggregation operator (namely, argument-wise plurality rule). We showed that while it satisfies most key properties, it fails to satisfy *Universal*

[6]The use of a non-deterministic tie-breaking rule has its own issues too, such as producing different outcomes given the same profile.

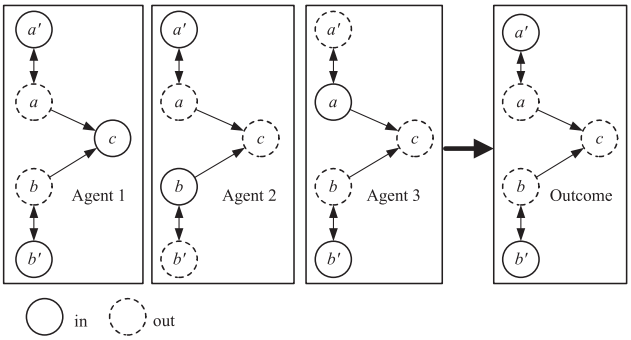


FIGURE 6. Counterexample to *Collective Rationality*.

Domain and *Collective Rationality*. In this section, we show a couple of impossibility results that involve these two postulates. The following result shows that introducing a tie-breaking rule to satisfy *Universal Domain* would result in violating *Anonymity* and/or *Strong Systematicity*.

THEOREM 1
There exists an argumentation framework AF such that, for any set of agents whose cardinality is divisible by three, there exists no labelling aggregation operator satisfying *Universal Domain*, *Anonymity* and *Strong Systematicity*.

PROOF. It is enough to assume an AF that contains at least one argument a which can feasibly take on any label, i.e. there exist complete labellings $L_{\mathtt{in}}, L_{\mathtt{out}}$ and $L_{\mathtt{undec}}$ over AF such that $L_{\mathtt{in}}(a)=\mathtt{in}$, $L_{\mathtt{out}}(a)=\mathtt{out}$ and $L_{\mathtt{undec}}(a)=\mathtt{undec}$. Divide n agents into three groups G_1, G_2, G_3 of equal size. By *Universal Domain*, all profiles consisting of legal labellings are valid input. Assume a profile in which everyone in G_1 provides labelling $L_{\mathtt{in}}$, everyone in G_2 provides $L_{\mathtt{out}}$ and everyone in G_3 provides $L_{\mathtt{undec}}$. For now let us denote this profile by $\mathcal{L}=([G_1:L_{\mathtt{in}}],[G_2:L_{\mathtt{out}}],[G_3:L_{\mathtt{undec}}])$. Now, assume for contradiction that F is an aggregation operator for AF satisfying *Universal Domain*, *Anonymity* and *Strong Systematicity*. Let $\rho:\{\mathtt{in},\mathtt{out},\mathtt{undec}\}\to\{\mathtt{in},\mathtt{out},\mathtt{undec}\}$ be any permutation on the set of labels such that $\rho(l)\neq l$ for all labels l (for instance, $\rho(\mathtt{in})=\mathtt{out}$, $\rho(\mathtt{out})=\mathtt{undec}$, $\rho(\mathtt{undec})=\mathtt{in}$), and let $\mathcal{L}'$ denote the profile $([G_1:L_{\rho(\mathtt{in})}],[G_2:L_{\rho(\mathtt{out})}],[G_3:L_{\rho(\mathtt{undec})}])$. Since $L'_i(a)=\rho(L_i(a))$ for all $i\in Ag$, *Strong Systematicity* implies $[F(\mathcal{L}')](a)=\rho([F(\mathcal{L})](a))$. However, we chose ρ s.t. $\rho(l)\neq l$. Hence, $[F(\mathcal{L}')](a)\neq[F(\mathcal{L})](a)$. But *Anonymity* implies $[F(\mathcal{L})](a)=[F(\mathcal{L}')](a)$. Contradiction. Hence no such F can exist. ■

The previous result can be read in two ways: first, the AWPR cannot be made to satisfy *Universal Domain* without violating *Strong Systematicity* or *Anonymity*. Second, there exists no aggregation operator at all that satisfies *Universal Domain*, *Strong Systematicity* and *Anonymity*.

Note that the previous theorem was stated for a set of agents divisible by three. Essentially, three-way ties would only happen if the cardinality of the agents is divisible by three (since there are only three possible labels for each argument, and each individual has to submit one label for each argument). Hence, one might wonder whether we could rule out the possibility of three-way ties, by assuming n cannot be a multiple of three.[7] However, with even number of agents, we can show

[7] It was shown in [24] that *Anonymity*, *Neutrality* (a weaker version of *Strong Systematicity*) and *Resolution* can be satisfied together if and only if the number of alternatives cannot be written as the sum of non-trivial dividers of the number of voters.

Copyedited by: ES MANUSCRIPT CATEGORY: Corner Article

that there is still a large class of *AF*s which do not have an operator satisfying those three postulates without violating *Collective Rationality*.

THEOREM 2
There exists an argumentation framework AF such that, for any set of agents of even cardinality, there exists no labelling aggregation operator satisfying *Universal Domain*, *Anonymity*, *Strong Systematicity* and *Collective Rationality*.

PROOF. It is enough to assume an AF that contains at least one argument a that can feasibly take on just two out of the three possible labels. For concreteness suppose a can only take on labels $\mathtt{out}$ and $\mathtt{undec}$ (An example of such a framework and an argument can be seen in the proof of Theorem 3 below, in which c can only be either $\mathtt{out}$ or $\mathtt{undec}$). Let $L_{\mathtt{undec}}$ and $L_{\mathtt{out}}$ be two complete labellings such that $L_{\mathtt{undec}}(a)=\mathtt{undec}$ and $L_{\mathtt{out}}(a)=\mathtt{out}$. Divide the agents into two groups G_1, G_2 of equal size. By *Universal Domain*, all profiles consisting of legal labellings are valid input, so assume a profile in which everyone in G_1 provides labelling $L_{\mathtt{undec}}$ and everyone in G_2 provides $L_{\mathtt{out}}$. Denote the resulting profile by $\mathcal{L}=([G_1:L_{\mathtt{undec}}],[G_2:L_{\mathtt{out}}])$ and assume for contradiction that F is an aggregation operator for this AF that satisfies *Universal Domain*, *Anonymity*, *Strong Systematicity* and *Collective Rationality*. Let ρ be the permutation that swaps $\mathtt{undec}$ and $\mathtt{out}$, i.e. $\rho(\mathtt{undec})=\mathtt{out}$ and $\rho(\mathtt{out})=\mathtt{undec}$, and let $\mathcal{L}'=([G_1:L_{\mathtt{out}}],[G_2:L_{\mathtt{undec}}])$.[8] By *Anonymity* we know $[F(\mathcal{L})](a)=[F(\mathcal{L}')](a)$. Then it cannot be that $[F(\mathcal{L})](a)=\mathtt{undec}$, for if so then *Strong Systematicity* would imply $[F(\mathcal{L}')](a)=\rho(\mathtt{undec})=\mathtt{out}\neq[F(\mathcal{L})](a)$, and similarly it cannot be that $[F(\mathcal{L})](a)=\mathtt{out}$. Thus we must have $[F(\mathcal{L})](a)=\mathtt{in}$. But by *Collective Rationality* $[F(\mathcal{L})](a)\in\{\mathtt{undec},\mathtt{out}\}$. Contradiction. ■

The careful reader can realize that *Collective Rationality* can be substituted with *Supportiveness* in the previous theorem. As for the proof, the last sentence becomes: 'Thus we must have $[F(\mathcal{L})](a)=\mathtt{in}$. But by *Supportiveness* $[F(\mathcal{L})](a)\in\{\mathtt{undec},\mathtt{out}\}$. Contradiction'.

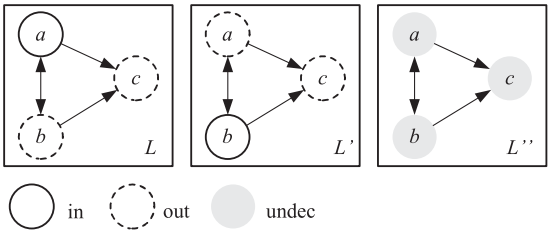
However, one might argue that *Strong Systematicity* is quite a strong condition. Treating $\mathtt{in}$, $\mathtt{out}$, and $\mathtt{undec}$ differently can be tolerated. Then, it is interesting to ask: 'Does there exist an operator that satisfies *Universal Domain*, *Weak Systematicity*, and *Anonymity*?'. The answer for this question is positive. Consider a modified version of the AWPR that deals with ties by labelling every argument that has a tie with $\mathtt{undec}$. One can show that this operator satisfies these three properties together. However, this operator still violates *Collective Rationality* (Example 6 holds as a counterexample). In fact, we show that any operator that satisfies *Universal Domain*, *Weak Systematicity*, and *Anonymity*, would violate either *Collective Rationality* or *Unanimity*.

THEOREM 3
There exists an argumentation framework *AF* such that, for any set of agents of even cardinality, there exists no labelling aggregation operator satisfying *Universal Domain*, *Weak Systematicity*, *Anonymity*, *Collective Rationality* and *Unanimity*.

PROOF. Consider the following argumentation framework. An argument c is defeated by two arguments a and b which defeat each others.

Resolute rules always produce a single outcome, so it resembles *No-Tie Universal Domain*. Also, in our settings, the number of candidates is three. So this result says that we can have these postulates together if the number of voters is not a multiple of three.

[8] Note here that all labellings in the profile $\mathcal{L}'$ are still *complete* labellings. This is because ρ does not uniformly exchange all labels in a given labelling, it is just a permutation on the set of labels.



Consider the two labellings $L=(\{a\},\{b,c\},\{\})$ and $L'=(\{b\},\{a,c\},\{\})$. Assume, towards a contradiction, that there exists an aggregation operator F that satisfies *Universal Domain*, *Collective Rationality*, *Weak Systematicity*, *Anonymity* and *Unanimity*.

By *Universal Domain*, we may consider any profile consisting of legal labellings. Consider the two profiles $\mathcal{L}=(L,\ldots,L,L',\ldots,L')$ and $\mathcal{L}'=(L',\ldots,L',L,\ldots,L)$. That is, in $\mathcal{L}$ half the agents give L and the other half give L', and then in $\mathcal{L}'$ the agents switch from L to L' and vice versa.

By *Unanimity* we know

$$[F(\mathcal{L})](c)=\texttt{out}. \tag{2a}$$

By *Weak Systematicity* we also know $[F(\mathcal{L})](a)=[F(\mathcal{L}')](b)$. But since $\mathcal{L}$ and $\mathcal{L}'$ are permutations of each other we know $F(\mathcal{L})=F(\mathcal{L}')$ by *Anonymity* and so we obtain

$$[F(\mathcal{L})](a)=[F(\mathcal{L})](b). \tag{2b}$$

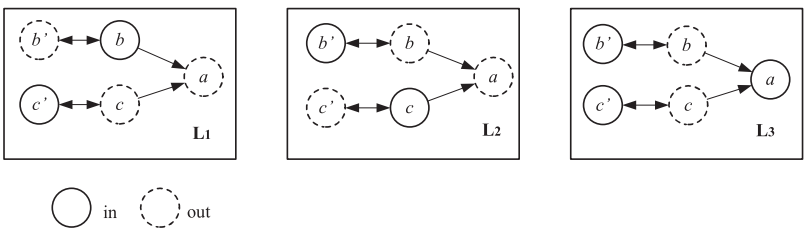
But there is no complete labelling simultaneously satisfying (2a) and (2b). Contradiction. Hence no F can exist.

■

One might note that all of the above theorems exploit the use of profiles that include ties. Then, one would ask: what if we relax *Universal Domain* to *No-Tie Universal Domain*? Do we still have impossibility results then? Following, we show that an aggregation operator which satisfies *No-Tie Universal Domain* (but not necessarily *Universal Domain*) cannot also satisfy *Weak Systematicity*, *Anonymity*, *Collective Rationality* and *Supportiveness* together.

THEOREM 4
There exists an argumentation framework AF such that, for any set of agents whose cardinality is divisible by three, there exists no labelling aggregation operator satisfying *No-Tie Universal Domain*, *Weak Systematicity*, *Anonymity*, *Collective Rationality* and *Supportiveness*.

PROOF. Consider the following argumentation framework. An argument a is defeated by two arguments b and c. Argument b (resp. c) defeats and is defeated by argument b' (resp. c').



Consider the three labellings $L_1=(\{b,c'\},\{a,b',c\},\{\})$, $L_2=(\{b',c\},\{a,b,c'\},\ \{\})$ and $L_3=(\{a,b',c'\},\{b,c\},\{\})$.

Copyedited by: ES MANUSCRIPT CATEGORY: Corner Article

Assume, towards a contradiction, that there exists an aggregation operator F that satisfies *No-Tie Universal Domain*, *Collective Rationality*, *Weak Systematicity*, *Anonymity* and *Supportiveness*.

By *No-Tie Universal Domain*, we may consider any profile consisting of legal labellings as long as it does not cause a tie. We consider here three agents, but the same proof can be shown for any set of agents that is divisible by three. Consider the three profiles $\mathcal{L}=(L_1,L_2,L_3)$, $\mathcal{L}'=(L_1',L_2',L_3')=(L_3,L_1,L_2)$ and $\mathcal{L}''=(L_1'',L_2'',L_3'')=(L_2,L_3,L_1)$.

Since $\forall i, L_i(a)=L_i'(c)$, then by *Weak Systematicity* we know:

$$[F(\mathcal{L})](a)=[F(\mathcal{L}')](c). \tag{3a}$$

But since $\mathcal{L}$ and $\mathcal{L}'$ are permutations of each other we know $[F(\mathcal{L})]=[F(\mathcal{L}')]$ by *Anonymity* and so we obtain

$$[F(\mathcal{L})](c)=[F(\mathcal{L}')](c). \tag{3b}$$

From (3a and 3b):

$$[F(\mathcal{L})](a)=[F(\mathcal{L})](c). \tag{3c}$$

Similarly, since $\forall i, L_i(b)=L_i''(c)$, then by *Weak Systematicity* we know:

$$[F(\mathcal{L})](b)=[F(\mathcal{L}'')](c). \tag{3d}$$

But since $\mathcal{L}$ and $\mathcal{L}''$ are permutations of each other we know $[F(\mathcal{L})]=[F(\mathcal{L}'')]$ by *Anonymity* and so we obtain

$$[F(\mathcal{L})](c)=[F(\mathcal{L}'')](c). \tag{3e}$$

From (3d) and (3e):

$$[F(\mathcal{L})](b)=[F(\mathcal{L})](c). \tag{3f}$$

From (3c) and (3f):

$$[F(\mathcal{L})](a)=[F(\mathcal{L})](b)=[F(\mathcal{L})](c). \tag{3g}$$

The last equation suggests that a, b and c have the same collective labelling. However, by *Collective Rationality*, the only legal labelling that satisfy (3g) is `undec`:

$$[F(\mathcal{L})](a)=[F(\mathcal{L})](b)=[F(\mathcal{L})](c)=\texttt{undec}. \tag{3h}$$

However, F satisfies *Supportiveness* by assumption. Contradiction.

■

One can draw a connection between this result and the previous one. Relaxing *Universal Domain* to *No-Tie Universal Domain*, introduces another impossibility result, in which *Unanimity* is replaced with the stronger postulate *Supportiveness*. Additionally, one can compare this result to the analogue of Arrow's theorem in judgement aggregation [22], which involves *Unanimity*, *Independence*, and *Non-dictatorship*, the weaker versions of *Supportiveness*, *Weak Systematicity*, and *Anonymity* respectively in our theorem. However, their result also involves *completeness*, i.e. no proposition can be collectively undecided, which we do not have as a condition in our result.

The above impossibility results highlight a major barrier to reaching good collective judgement about argument evaluation in general. These establish the limits of aggregation in the context of argumentation, and come in accordance with the similar topics of aggregation such as preference

Copyedited by: ES MANUSCRIPT CATEGORY: Corner Article

aggregation [1] and judgement aggregation [21]. Unfortunately, there is no escape from violating the involved conditions or accepting irrational aggregate argument labellings without somewhat lowering our standards in terms of desirable criteria.

7 Collective rationality postulates

In this section, we characterize *Collective Rationality* in terms of conditions that need to be satisfied by profiles. To do this, we need to go back to the definition of legal (i.e. *complete*) labelling (Definition 4), and break it down into further constituents defined over the outcome of an aggregation operator.

The following condition, which we call *IN-Collective Rationality (IN-CR)*, requires that if an argument a is collectively accepted by the agents, then the agents must collectively reject all counter-arguments against a.

IN-Collective Rationality (IN-CR) For any profile $\mathcal{L}$ and $a \in \mathcal{A}$, if $[F(\mathcal{L})](a) = \texttt{in}$ then:

$$\nexists b \in \mathcal{A}, \text{ s.t. } (b \rightharpoonup a \wedge [F(\mathcal{L})](b) = \texttt{in}) \qquad \textbf{(IN-CR1)}$$

and

$$\nexists b \in \mathcal{A}, \text{ s.t. } (b \rightharpoonup a \wedge [F(\mathcal{L})](b) = \texttt{undec}) \qquad \textbf{(IN-CR2)}$$

Note that IN-CR1, the first part of IN-CR, represents the condition of *conflict-freeness* applied on the output. The condition of *conflict-freeness* is usually agreed on as a minimal reasonable condition in argument evaluation.

We present now the *OUT-Collective Rationality (OUT-CR)* condition. Intuitively, this condition means that if an argument a is collectively rejected by the agents, then the agents must also collectively agree on accepting at least one of the counter-arguments against a.

OUT-Collective Rationality (OUT-CR) For any profile $\mathcal{L}$ and $a \in \mathcal{A}$, if $[F(\mathcal{L})](a) = \texttt{out}$ then $\exists b \in \mathcal{A}$, such that $b \rightharpoonup a$ and $[F(\mathcal{L})](b) = \texttt{in}$.

We present now the *UNDEC-Collective Rationality (UNDEC-CR)* condition. An argument must be labelled `undec` if and only if: (i) it is not the case that all of its defeaters are `out`, i.e. at least one of its defeaters is `undec`; and (ii) none of its defeaters is `in`.

UNDEC-Collective Rationality (UNDEC-CR) For any profile $\mathcal{L}$ and $a \in \mathcal{A}$, if $[F(\mathcal{L})](a) = \texttt{undec}$ then:

$$\nexists b \in \mathcal{A}, \text{ s.t. } (b \rightharpoonup a \wedge [F(\mathcal{L})](b) = \texttt{in}) \qquad \textbf{(UNDEC-CR1)}$$

and

$$\exists b \in \mathcal{A}, \text{ s.t. } (b \rightharpoonup a \wedge [F(\mathcal{L})](b) = \texttt{undec}) \qquad \textbf{(UNDEC-CR2)}$$

The following result follows immediately from the definitions.

PROPOSITION 3
An argument aggregation operator F satisfies *Collective Rationality* if and only if for each profile $\mathcal{L} = (L_1, \ldots, L_n)$ in its domain, it satisfies the *IN-CR*, *OUT-CR*, and *UNDEC-CR* conditions.

8 Plurality rule with classical semantics

In this section, we analyse the performance of AWPR with respect to *Collective Rationality* when agents labellings are restricted to some classical semantics (i.e. *complete*, *grounded*, *stable*, *semi-stable* and *preferred*). This investigation gives a novel meaning to classical semantics in social choice

Copyedited by: ES MANUSCRIPT CATEGORY: Corner Article

settings. Rather than simply being compared by their logical rigour from the perspective of a single agent, semantics are compared based on the extent to which they facilitate collectively rational agreement among agents.

Our strategy will be based on the following approach. Since, by Proposition 3, *Collective Rationality* arises iff *IN-CR*, *OUT-CR* and *UNDEC-CR* are satisfied, it is enough to check whether AWPR satisfies those properties.

8.1 Complete semantics

Since the complete semantics generalizes other classical semantics, we provide analysis for it first. Every property that is satisfied by AWPR when individuals' labellings are *complete* labellings would be also satisfied by AWPR when individuals' labellings are restricted to the other classical semantics that we consider.

It is very interesting to see that, as the proposition below shows, when agents collectively accept an argument, the structure of the AWPR will ensure that they will not collectively accept any of its defeaters:

PROPOSITION 4
AWPR satisfies *IN-CR1*. Using the argument-wise plurality rule, given any profile $\mathcal{L}=(L_1,\ldots,L_n)$, if an argument a is collectively accepted, none of its defeaters will be collectively accepted. Formally, if $[M(\mathcal{L})](a)=\texttt{in}$ for some arbitrary $a\in\mathcal{A}$ then $\nexists b\in\mathcal{A}$, such that $b\rightharpoonup a$ and $[M(\mathcal{L})](b)=\texttt{in}$.

PROOF. Suppose that $[M(\mathcal{L})](a)=\texttt{in}$ holds. By definition:

$$|\{i:L_i(a)=\texttt{in}\}|>|\{i:L_i(a)=\texttt{out}\}| \tag{4a}$$

Since each L_i is a legal labelling, an agent who votes in for a must also vote out for each defeater of a. Therefore:

$$\forall b\rightharpoonup a \quad |\{i:L_i(b)=\texttt{out}\}|\geq|\{i:L_i(a)=\texttt{in}\}| \tag{4b}$$

We want to show that: $\nexists b\in\mathcal{A}$ such that $b\rightharpoonup a$ and $[M(\mathcal{L})](b)=\texttt{in}$

Assume (towards contradiction) that the contrary holds. That is, $\exists b'\in\mathcal{A}$ such that $b'\rightharpoonup a$ and $[M(\mathcal{L})](b')=\texttt{in}$. Then:

$$|\{i:L_i(b')=\texttt{in}\}|>|\{i:L_i(b')=\texttt{out}\}| \tag{4c}$$

Since every agent who voted in for b' would have voted out for a, we have:

$$|\{i:L_i(a)=\texttt{out}\}|\geq|\{i:L_i(b')=\texttt{in}\}| \tag{4d}$$

By (4c) and (4d):

$$|\{i:L_i(a)=\texttt{out}\}|>|\{i:L_i(b')=\texttt{out}\}| \tag{4e}$$

while from (4b) and (4e) we have that:

$$|\{i:L_i(a)=\texttt{out}\}|>|\{i:L_i(a)=\texttt{in}\}| \tag{4f}$$

But this contradicts (4a) and the assumption that $[M(\mathcal{L})](a)=\texttt{in}$. ■

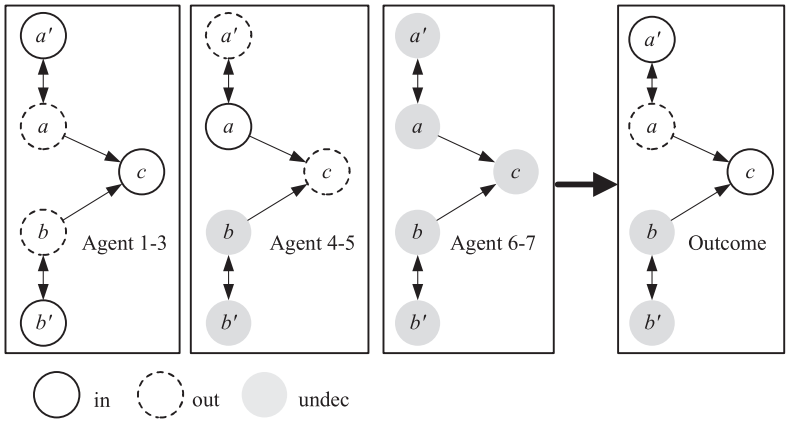


FIGURE 7. Seven votes collectively accepting c, without collectively rejecting b.

It is important to recognize that Proposition 4 is a non-trivial result. It shows that, with AWPR, the postulate IN-CR1 is satisfied. This means, as we mentioned earlier, that AWPR satisfies the 'collective' version of *conflict-freeness*, a condition that is usually agreed on as a minimal reasonable condition in argument evaluation. This comes 'for free' as a result of the intrinsic structure of the individual labellings, leading to coordinated votes. Note, however, that the *IN-CR* postulate is not fully satisfied. Although Proposition 4 guarantees that a collectively accepted argument will never have a collectively accepted defeater, it does not guarantee *IN-CR2*, that none of its defeaters will be collectively undecided. This is demonstrated in the following remark.

REMARK 1
AWPR violates IN-CR2. If an argument is collectively accepted, some of its defeaters might be collectively undecided.

PROOF. Suppose argument c has two defeaters, a and b. Suppose we have 7 agents, with votes as shown in Figure 7. Clearly, while c is collectively accepted because $[M(\mathcal{L})](c) = \texttt{in}$, one of its defeaters is not collectively rejected because $[M(\mathcal{L})](b) = \texttt{undec}$. ■

As we saw earlier in Example 6, *OUT-CR* is violated by AWPR.

REMARK 2
AWPR violates *OUT-CR*. If an argument is collectively rejected, it is not guaranteed that one of its defeaters will be collectively accepted.

PROOF. See Example 6 for a counterexample. ■

The following remark shows that there are no intrinsic guarantees for satisfying *UNDEC-CR1*.

REMARK 3
AWPR violates *UNDEC-CR1*. If an argument is collectively undecided, it is possible that one of its defeaters will be collectively accepted.

Copyedited by: ES MANUSCRIPT CATEGORY: Corner Article

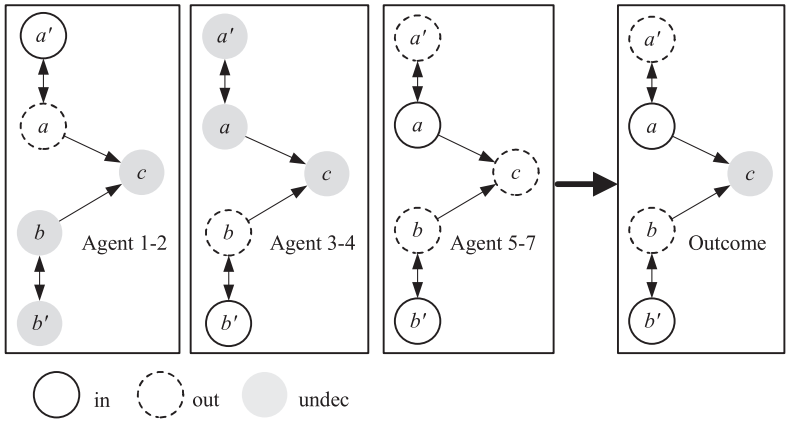


FIGURE 8. Seven agents collectively undecided on c, but collective accepting a.

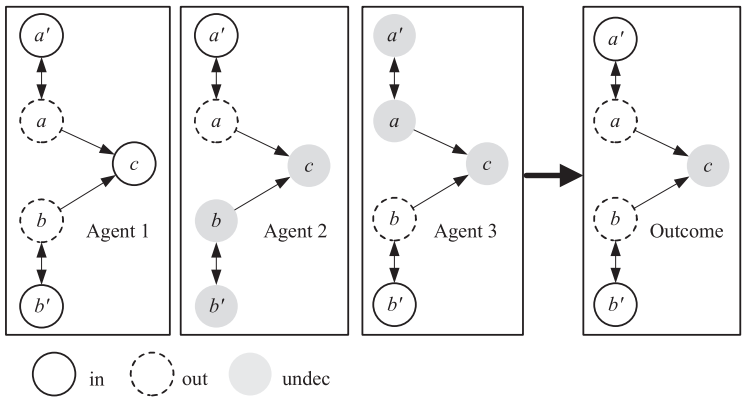


FIGURE 9. Three votes collectively undecided about c, but not collectively undecided about any of its defeaters a or b.

PROOF. Suppose argument c has two defeaters, a and b. Suppose we have 7 agents. Suppose the votes are as shown in Figure 8. We have $[M(\mathcal{L})](c) = \texttt{undec}$ with 4 votes, but we have $[M(\mathcal{L})](a) = \texttt{in}$ with 3 votes, thus violating the postulate. ■

Similarly, the remark below shows that *UNDEC-CR2* is not intrinsically guaranteed.

REMARK 4
AWPR violates *UNDEC-CR2*. If an argument is collectively undecided, it is possible that none of its defeaters will be collectively undecided.

PROOF. Suppose argument c has two defeaters, a and b. Suppose we have 3 agents, with votes as shown in Figure 9. Clearly, we have $[M(\mathcal{L})](c) = \texttt{undec}$, but we have $[M(\mathcal{L})](a) = \texttt{out}$ and $[M(\mathcal{L})](b) = \texttt{out}$, which would have required c to be $\texttt{in}$. ■

Copyedited by: ES MANUSCRIPT CATEGORY: Corner Article

8.2 Other classical semantics

As we noted before, each possible *complete* labelling represents a valid self-defending viewpoint, therefore restricting votes to *complete* labellings is akin to requiring that each vote in judgement aggregation is consistent, or that each preference in preference aggregation is transitive and complete. Other classical semantics are essentially restrictions (i.e. sub-cases) of complete semantics. For example, restricting votes to *preferred* semantics requires each individual to be more committed, maximizing (w.r.t. set-inclusion) the set of accepted (or the set of rejected) arguments, while restricting votes to *semi-stable* semantics requires each individual to be less conservative, minimizing (w.r.t. set-inclusion) the set of arguments about which they are undecided. It is not clear, a priori, what such requirements, applied on the individual, would have on the collective rationality of the outcome of voting.

In this subsection, we provide an analysis for the *grounded*, *stable*, *semi-stable* and *preferred* semantics as more restricted forms of labellings to choose from. Note that the definition of *Universal Domain*, introduced earlier using *complete* semantics, is now redefined with respect to these semantics, while the definition of *Collective Rationality* is unchanged.

The following proposition looks trivial but, as we will see, it is the most positive result in this subsection.

PROPOSITION 5
If for every argument, agents can only vote for the *grounded* labelling, then M satisfies *IN-CR1*, *IN-CR2*, *OUT-CR*, *UNDEC-CR1* and *UNDEC-CR2*. Equivalently, M satisfies *Collective Rationality*.

PROOF. Trivial since there always exists one grounded labeling [8, 15], and M satisfies *Unanimity*. ■

As a corollary of Proposition 4, when agents votes are restricted to *stable* (respectively *semi-stable* or *preferred*) labellings, AWPR satisfies IN-CR1.

COROLLARY 2
When agents can only vote for *stable* (respectively *semi-stable* or *preferred*) labellings, AWPR satisfies IN-CR1

PROOF. From Proposition 4, if agents can only vote for *complete* labellings, then AWPR satisfies IN-CR1. Since every *stable* (respectively *semi-stable* or *preferred*) labelling is a complete labelling, then when agents votes are restricted to these semantics, AWPR satisfies IN-CR1. ■

LEMMA 1
When agents can only vote for a *stable* labelling, AWPR satisfies *IN-CR2*. If an argument is collectively accepted, none of its defeaters is collectively undecided.

PROOF. Suppose, towards a contradiction, that there exists an argument that is collectively accepted and one of its defeaters is collectively undecided. Then, by *Supportiveness*, there exists one submitted labelling (by some agent) in which this argument is undecided. However, agents are only allowed to submit a *stable* labelling, and *stable* labellings have no argument labelled undecided. Contradiction. ■

REMARK 5
When agents can only vote for *stable* (respectively *semi-stable* or *preferred*) labellings, AWPR violates *OUT-CR*. If an argument is collectively rejected, it is possible that none of its defeaters is collectively accepted.

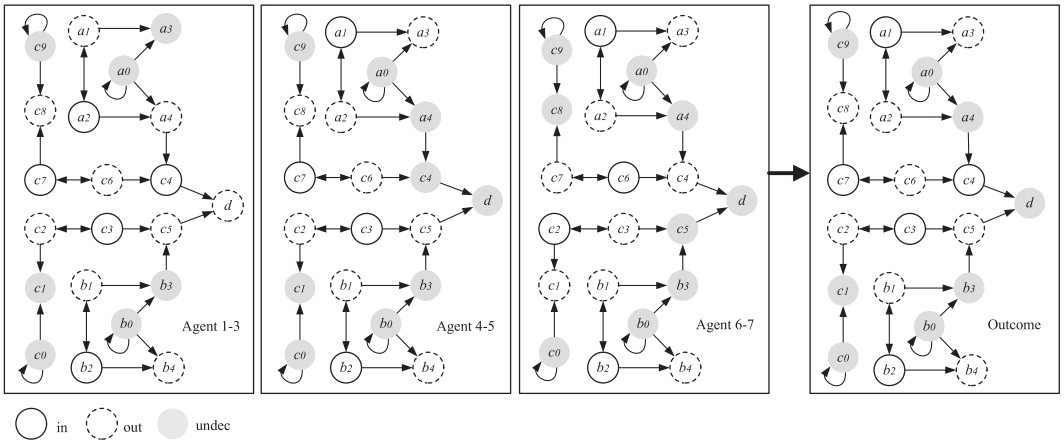


FIGURE 10. A counterexample shows how, given semi-stable (respectively preferred) semantics, AWPR violates IN-CR2 and UNDEC-CR1.

PROOF. See Example 6 for a counterexample. ■

LEMMA 2
When agents can only vote for a *stable* labelling, AWPR satisfies *UNDEC-CR* (i.e. it satisfies both *UNDEC-CR1* and *UNDEC-CR2*). If an argument is collectively undecided, none of its defeaters is collectively accepted, and at least one of its defeaters is collectively undecided.

PROOF. Since in *stable* labelling no argument is labelled undecided, by *Supportiveness*, there is no argument that is collectively undecided. Then, this lemma holds. ■

We continue with the *semi-stable* and *preferred* semantics.

REMARK 6
When agents can only vote for a *semi-stable* (respectively *preferred*) labelling, AWPR violates *IN-CR2*. If an argument is collectively accepted, it is possible that one of its defeaters is collectively undecided.

PROOF. Suppose argument c_4 has two defeaters, a_4 and c_6. Suppose we have 7 agents, with votes as shown in Figure 10. Clearly, while c_4 is collectively accepted because $[M(\mathcal{L})](c_4) = \texttt{in}$, one of its defeaters, namely a_4, is collectively undecided because $[M(\mathcal{L})](a_4) = \texttt{undec}$. ■

REMARK 7
When agents can only vote for a *semi-stable* (respectively *preferred*) labelling, AWPR violates *UNDEC-CR1*. If an argument is collectively undecided, it is possible that one of its defeaters is collectively accepted.

PROOF. Suppose argument d has two defeaters, c_4 and c_5. Suppose we have 7 agents, with votes as shown in Figure 10. Clearly, while d is collectively undecided because $[M(\mathcal{L})](d) = \texttt{undec}$, one of its defeaters, namely c_4, is collectively accepted because $[M(\mathcal{L})](c_4) = \texttt{in}$. ■

Copyedited by: ES MANUSCRIPT CATEGORY: Corner Article

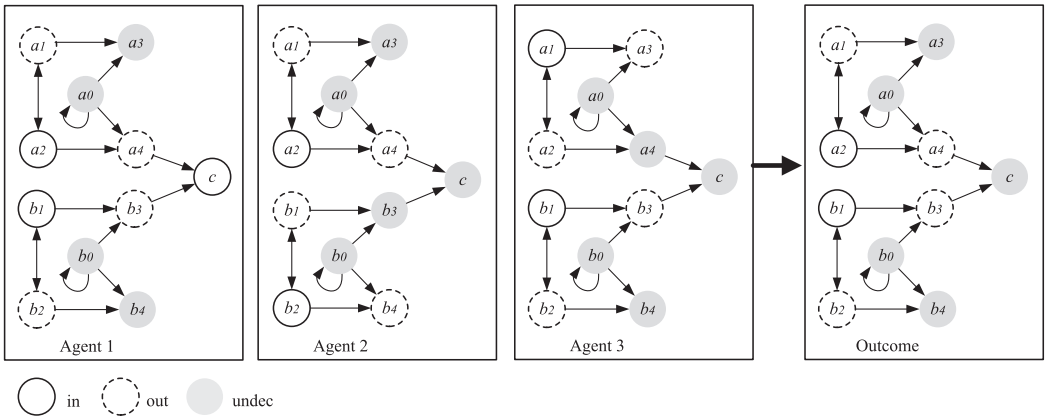


FIGURE 11. A counterexample shows how, given semi-stable (respectively preferred) semantics, AWPR violates UNDEC-CR2.

TABLE 1. The *Collective Rationality* properties that are satisfied/violated by AWPR given different semantics.

Semantics	**IN-CR**		**OUT-CR**	**UNDEC-CR**	
	IN-CR1	**IN-CR2**		**UND-CR1**	**UND-CR2**
Grounded	Yes (Prop. 5)	Yes (Prop. 5)	Yes (Prop. 5)	Yes (Prop. 5)	Yes (Prop. 5)
Stable	Yes (Cor. 2)	Yes (Lem. 1)	No (Rem. 5)	Yes (Lem. 2)	Yes (Lem. 2)
Semi-stable	Yes (Cor. 2)	No (Rem. 6)	No (Rem. 5)	No (Rem. 7)	No (Rem. 8)
Preferred	Yes (Cor. 2)	No (Rem. 6)	No (Rem. 5)	No (Rem. 7)	No (Rem. 8)
Complete	Yes (Prop. 4)	No (Rem. 1)	No (Rem. 2)	No (Rem. 3)	No (Rem. 4)

REMARK 8
When agents can only vote for a *semi-stable* (respectively *preferred*) labelling, AWPR violates *UNDEC-CR2*. If an argument is collectively undecided, it is possible that none of its defeaters is collectively undecided.

PROOF. Suppose argument c has two defeaters, a_4 and b_3. Suppose we have 3 agents, with votes as shown in Figure 11. Clearly, while c is collectively undecided because $[M(\mathcal{L})](c) = \texttt{undec}$, none of its defeaters is collectively undecided. ■

To sum up, the only restriction that would satisfy the *Collective Rationality* is the *grounded* semantics (Proposition 5). This is trivially true because only one *grounded* labelling exists. However, *stable* semantics violates *Collective Rationality* only because it violates *OUT-CR*. As for the *semi-stable* and *preferred* semantics, they only satisfy *IN-CR1*, a property they inherit from the *complete* semantics. Refer to Table 1 for a summary of the results we have found.

Copyedited by: ES MANUSCRIPT CATEGORY: Corner Article

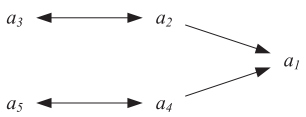


FIGURE 12. An example about issues.

9 Restricting the domain of argumentation graphs to satisfy collective rationality

In an earlier section, we showed that, AWPR violates *Universal Domain* and *Collective Rationality*. In this section, we investigate whether AWPR can satisfy *Collective Rationality* by restricting the argumentation framework to graphs with certain graph-theoretical properties. We show that graphs consisting of disconnected issues (a notion we define below) and graphs in which arguments have limited defeaters (in some sense) guarantee collectively rational outcomes when the AWPR is used.

9.1 Disconnected issues

The notion of 'issue' was defined in [7] in order to quantify disagreement between graph labellings. In this section, we use this notion to provide a possibility result.

Crucial to the definition of the 'issue' is the concept of 'in-sync'. Two arguments a and b are said to be *in-sync* if the (*complete*) label of one cannot be changed without causing a change of equal magnitude to the label of the other.

DEFINITION 10 (in-Sync $\equiv$ [7])
Let $Comp(AF)$ be the set of all complete labellings for argumentation framework $AF = \langle \mathcal{A}, \rightharpoonup \rangle$. We say that two arguments $a, b \in \mathcal{A}$ are in-sync ($a \equiv b$):

$$a \equiv b \text{ iff } (a \equiv_1 b \vee a \equiv_2 b) \tag{5}$$

where:

- $a \equiv_1 b$ iff $\forall L \in Comp(AF) : L(a) = L(b)$.
- $a \equiv_2 b$ iff $\forall L \in Comp(AF) : (L(a) = \texttt{in} \Leftrightarrow L(b) = \texttt{out}) \wedge (L(a) = \texttt{out} \Leftrightarrow L(b) = \texttt{in})$

This relation forms an equivalence relation over the arguments, and the equivalence classes are called "issues".

DEFINITION 11 (Issue [7])
Given the argumentation framework $AF = \langle \mathcal{A}, \rightharpoonup \rangle$, a set of arguments $\mathcal{B} \subseteq \mathcal{A}$ is called an issue iff it forms an equivalence class of the relation *in-Sync* ($\equiv$).

For example, in Figure 12, the graph consists of 3 issues, namely $\{a_1\}$, $\{a_2, a_3\}$ and $\{a_4, a_5\}$.

The following lemma is crucial in showing the main result of this subsection. We show that if the defeaters of an argument belong to the same issue as the argument, then the collective labelling of this argument chosen by AWPR is always a legal labelling.

LEMMA 3
Let $AF=\langle\mathcal{A},\rightharpoonup\rangle$ be an argumentation framework. Let $a\in\mathcal{A}$ be an argument in this framework. If every defeater of a (call it b) belongs to the same issue of a (i.e. $\forall b\in a^{-}$: $b\equiv a$), then AWPR would always produce a legal collective labelling for argument a.

PROOF. Let $b_1,\ldots,b_m\in\mathcal{A}$ such that $b_j\in a^{-}$ and $a\equiv b_j\ \forall j=1,\ldots,m$. Then, for every complete labelling L:

$$L(a)=\texttt{out}\Leftrightarrow L(b_1)=\texttt{in}\Leftrightarrow\ldots\Leftrightarrow L(b_m)=\texttt{in} \tag{6a}$$

$$L(a)=\texttt{in}\Leftrightarrow L(b_1)=\texttt{out}\Leftrightarrow\ldots\Leftrightarrow L(b_m)=\texttt{out} \tag{6b}$$

$$L(a)=\texttt{undec}\Leftrightarrow L(b_1)=\texttt{undec}\Leftrightarrow\ldots\Leftrightarrow L(b_m)=\texttt{undec} \tag{6c}$$

From (6a), (6b) and (6c), for every labelling profile $\mathcal{L}=(L_1,\ldots,L_n)$:

$$|\{i:L_i(a)=\texttt{out}\}|=|\{i:L_i(b_1)=\texttt{in}\}|=\ldots=|\{i:L_i(b_m)=\texttt{in}\}| \tag{6d}$$

$$|\{i:L_i(a)=\texttt{in}\}|=|\{i:L_i(b_1)=\texttt{out}\}|=\ldots=|\{i:L_i(b_m)=\texttt{out}\}| \tag{6e}$$

$$|\{i:L_i(a)=\texttt{undec}\}|=|\{i:L_i(b_1)=\texttt{undec}\}|=\ldots=|\{i:L_i(b_m)=\texttt{undec}\}| \tag{6f}$$

From (6d), (6e) and (6f):

$$[M(\mathcal{L})](a)=\texttt{out}\Leftrightarrow[M(\mathcal{L})](b_1)=\texttt{in}\Leftrightarrow\ldots\Leftrightarrow[M(\mathcal{L})](b_m)=\texttt{in} \tag{6g}$$

$$[M(\mathcal{L})](a)=\texttt{in}\Leftrightarrow[M(\mathcal{L})](b_1)=\texttt{out}\Leftrightarrow\ldots\Leftrightarrow[M(\mathcal{L})](b_m)=\texttt{out} \tag{6h}$$

$$[M(\mathcal{L})](a)=\texttt{undec}\Leftrightarrow[M(\mathcal{L})](b_1)=\texttt{undec}\Leftrightarrow\ldots\Leftrightarrow[M(\mathcal{L})](b_m)=\texttt{undec} \tag{6i}$$

From (6g), (6h) and (6i), AWPR satisfies *IN-CR*, *OUT-CR*, and *UNDEC-CR* with respect to a in this case. Then, a is always legally collectively labelled by AWPR if every defeater of it is in the same issue as a. ■

Given the previous lemma, we show that if the argumentation framework consists of a set of disconnected issues, then AWPR satisfies *Collective Rationality* for this framework.

THEOREM 5
For every $AF=\langle\mathcal{A},\rightharpoonup\rangle$ that consists of a set of disconnected components (i.e. disconnected subgraphs), each of which forms an issue, the argument-wise plurality rule would always produce collectively rational outcomes.

PROOF. Since AF consists of a set of disconnected issues, then $\forall a\in\mathcal{A}$, a has the following property: $\forall b\in\mathcal{A}$ such that $b\in a^{-}$ then $b\equiv a$. From Lemma 3, a is always legally collectively labelled by AWPR. Then AWPR satisfies *Collective Rationality* for this AF. ■

This result shows that under argumentation frameworks that consist of disconnected issues, AWPR always satisfies *collective rationality*. Indeed, as long as all arguments in every connected component are 'in-sync', the labelling of one argument fully specifies the labelling of all those connected to it. Then, one can think of these disconnected components/issues as a set of independent propositions, and voting is done issue-wise.

Copyedited by: ES MANUSCRIPT CATEGORY: Corner Article

9.2 *Limited defeaters*

Now we move to another condition. It simply states that the defeaters of any argument are limited by the flexibility of labelling of these defeaters. To illustrate the latter term, we use a concept called the 'justification status', which is defined in [34]. Intuitively, the justification status of an argument is the set of possible labellings that this argument can take.

DEFINITION 12 (Justification Status [34])
Let $AF = \langle \mathcal{A}, \rightharpoonup \rangle$ be an argumentation framework, and $a \in \mathcal{A}$ some argument. The justification status of a is the outcome yielded by the function $\mathcal{JS} : \mathcal{A} \to 2^{\{\mathtt{in},\mathtt{out},\mathtt{undec}\}}$ such that $\mathcal{JS}(a) = \{L(a) | L \in Comp(AF)\}$.

There are six possible justification statuses. Neither $\emptyset$ nor $\{\mathtt{in}, \mathtt{out}\}$ is a possible justification status. The former is because each argumentation framework has at least one *complete* labelling. The later is because of the following theorem.

THEOREM 6 ([34, Theorem 2])
Let $AF = \langle \mathcal{A}, \rightharpoonup \rangle$ be an argumentation framework, and $a \in \mathcal{A}$ some argument. If AF has two complete labellings L_1 and L_2 such that $L_1(a) = \mathtt{in}$ and $L_2(a) = \mathtt{out}$, then there exists a labelling L_3 such that $L_3(a) = \mathtt{undec}$.

The following lemma shows that an argument with one of its defeaters belong to the same issue as long as all the other defeaters of this argument have the justification status of $\{\mathtt{out}\}$.

LEMMA 4
Let $AF = \langle \mathcal{A}, \rightharpoonup \rangle$ be an argumentation framework, and $a, b \in \mathcal{A}$ two arguments such that $b \rightharpoonup a$. If the following holds:

$$\forall c \neq b : (c \rightharpoonup a \Rightarrow \mathcal{JS}(c) = \{\mathtt{out}\}) \tag{7a}$$

Then a and b belong to the same issue (i.e. $a \equiv b$). Moreover, a is always legally collectively labelled by AWPR.

PROOF. One can show that:

$$L(a) = \mathtt{out} \Leftrightarrow L(b) = \mathtt{in} \tag{7b}$$

$$L(a) = \mathtt{in} \Leftrightarrow L(b) = \mathtt{out} \tag{7c}$$

$$L(a) = \mathtt{undec} \Leftrightarrow L(b) = \mathtt{undec} \tag{7d}$$

Hence, $a \equiv b$.

Moreover, in a similar way to Lemma 3, one can show that, for every possible profile $\mathcal{L} = (L_1, \ldots, L_n)$, the following holds:

- If $[M(\mathcal{L})](a) = \mathtt{out}$ then $[M(\mathcal{L})](b) = \mathtt{in}$ $(b \in a^-)$.
- If $[M(\mathcal{L})](a) = \mathtt{in}$ then $[M(\mathcal{L})](b) = \mathtt{out}$, and by *Unanimity*, $\forall c \neq b : (c \rightharpoonup a \Rightarrow [M(\mathcal{L})](c) = \mathtt{out})$.
- If $[M(\mathcal{L})](a) = \mathtt{undec}$ then $[M(\mathcal{L})](b) = \mathtt{undec}$ $(b \in a^-)$, and by *Supportiveness*, $\forall c \neq b : (c \rightharpoonup a \Rightarrow [M(\mathcal{L})](c) \neq \mathtt{in})$.

Hence, a is always legally collectively labelled by AWPR. ■

COROLLARY 3
Let $AF = \langle \mathcal{A}, \rightharpoonup \rangle$ be an argumentation framework, and $a \in \mathcal{A}$ an argument. If $|a^-| = 1$ then a is always legally collectively labelled.

PROOF. From Lemma 4, a is always legally collectively labelled by AWPR. ■

Now we present the main theorem for this subsection. It says if all arguments have limited defeaters then AWPR always produces legally collective labellings. The limitation of the defeaters is characterized in both their number and their justification statuses.

THEOREM 7
Let $AF = \langle \mathcal{A}, \rightharpoonup \rangle$ be an argumentation framework. If each argument in $\mathcal{A}$ has at most one defeater that can be labelled `undec` then AWPR satisfies *Collective Rationality*.

PROOF. Suppose we have an $AF = \langle \mathcal{A}, \rightharpoonup \rangle$ s.t. each argument $a \in \mathcal{A}$ has at most one defeater $b \in a^-$ s.t. $\texttt{undec} \in \mathcal{JS}(b)$. Then, using Theorem 6:

$$\forall c \neq b : (c \rightharpoonup a \Rightarrow \mathcal{JS}(c) = \{\texttt{in}\} \vee \mathcal{JS}(c) = \{\texttt{out}\})$$

Now for each argument $a \in \mathcal{A}$, all defeaters c with $\mathcal{JS}(c) = \{\texttt{out}\}$ have no effect on the label of a, so one can remove these defeaters. Additionally, if one of the defeaters c (other than b) has $\mathcal{JS}(c) = \{\texttt{in}\}$, then all other defeaters (including b) will also have no effect on the label of a, so one can also remove those defeaters. As a result, for each argument a we will end up with one of the following:

- a has only one defeater b and $\texttt{undec} \in \mathcal{JS}(b)$, or
- a has only one defeater c and $\mathcal{JS}(c) = \{\texttt{in}\}$, or
- a has no defeaters

Note that in the last case, we would have $\mathcal{JS}(a) = \{\texttt{in}\}$, and since AWPR satisfies *Unanimity*, a would be legally collectively labeled `in`. As for the first two cases, using Corollary 3, a would be legally collectively labelled. ■

9.3 *Relating the two restrictions*

In this section, we proposed classes of argumentation frameworks that guarantee collective rationality for AWPR. Note that neither of the two classes (given in Theorems 5 and 7) is a generalization or a special case of the other. Example 7 shows an AF that satisfies the condition in Theorem 5 (i.e. disconnected issues), but violates the condition in Theorem 7 (i.e. limited defeaters), while Example 8 shows an AF that satisfies the condition in Theorem 7 (i.e. limited defeaters), but violates the condition in Theorem 5 (i.e. disconnected issues).

EXAMPLE 7
Note that the argumentation framework in Figure 13 satisfies the condition in Theorem 5. All the arguments a, b, c, d and e are in the same issue, so this AF consists of disconnected issues (only one issue in this case). However, this AF violates the condition in Theorem 7, since argument a is defeated by two arguments b and c, each of these defeaters has a justification status of $\{\texttt{in}, \texttt{out}, \texttt{undec}\}$, and so their justification statuses share `undec`.

Copyedited by: ES MANUSCRIPT CATEGORY: Corner Article

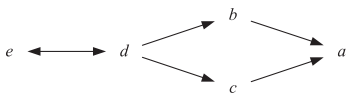


FIGURE 13. An AF that satisfies the condition in Theorem 5, but violates the condition in Theorem 7.

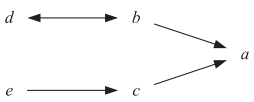


FIGURE 14. An AF that satisfies the condition in Theorem 7, but violates the condition in Theorem 5.

EXAMPLE 8
Note that the argumentation framework in Figure 14 satisfies the condition in Theorem 7. The only argument that is defeated by more than one argument is argument a which has two defeaters b and c. Moreover, $\texttt{undec} \notin \mathcal{JS}(c)$, so $\texttt{undec} \notin \mathcal{JS}(b) \cap \mathcal{JS}(c)$. However, this AF violates the condition in Theorem 5, since it contains two connected issues. The first issue is $\{a, b, d\}$ and the second issue is $\{c, e\}$.

Note that although the two proposed conditions can seem strong, constructing weaker conditions is not an easy task. Consider the graph in Figure 6. It violates both conditions (it consist of three connected issues $\{a, a'\}$, $\{b, b'\}$ and $\{c\}$, and both defeaters of c i.e. a and b have $\texttt{undec}$ in their justification status). However, this framework is very close to frameworks that satisfy one of the two conditions. For example, in the framework in Figure 6, if we remove the defeat $b \rightharpoonup b'$ we get a graph similar to the one in Figure 14, which satisfies the limited defeaters condition. On the other hand, if we remove the defeat $b \rightharpoonup c$ instead, we get a framework consisting of two disconnected issues, namely $\{a, a', c\}$ and $\{b, b'\}$. This suggests the difficulty of finding weaker conditions.

10 Discussion

In this article, we presented an extensive analysis of social-choice-theoretic aspects of Dung's highly-influential argumentation semantics. Argumentation-based semantics have mainly been compared on the basis of how they deal with specific benchmark problems that reflect specific logical structures from the point of view of a single omniscient observer (e.g. argument graph structures with odd-cycles, floating defeaters, etc.). Recently, it has been argued that argumentation semantics must be evaluated based on more general intuitive principles [3]. Our work can be seen as a contribution in this direction, focusing on issues relating to multi-agent preferences.

The closest work to the present article is Caminada and Pigozzi [10]. In their work, they propose three aggregation operators, namely sceptical, credulous and super credulous. Although the operators satisfy *Collective Rationality*, they violate *Independence*. These operators are also more applicable to scenarios where the compatibility of the collective labelling with each individual's labelling is appreciated or needed. Argument-wise plurality rule, on the other hand, can be applied to classical scenarios where some individuals might naturally disagree with the opinion of the group. Additionally, unlike our work, their work focuses on the proposed operators with only little attention to the general aggregation problem. Only four postulates are proposed, namely *Universal Domain*, *Collective*

Rationality, *Anonymity* and *Independence*, and there are no general impossibility results that holds for *any* operator.

Our results on the aggregation of different argument evaluations by multiple agents provide a new approach for conflict-resolution in multi-agent systems. While this work combines both arguing and voting, two processes that employ different procedures, we assume these two processes are done independently and by different groups of individuals. For example, a jury can vote on the evaluation of arguments that were laid down by the lawyers of two opposing sides. Thus, the arguing part, which happens between the lawyers occurs in an independent step before the voting step, on which our analysis focuses.

Our results contribute to research on aggregation in the context of argumentation. The social choice theoretic Arrovian properties have been analyzed in the context of social argument justification in [32]. An *extended* argumentation framework $AF^n = \langle \mathcal{A}, \rightharpoonup_1, \ldots, \rightharpoonup_n \rangle$ is defined, where each $\rightharpoonup_i$, $1 \leq i \leq n$, is a particular attack relation among the arguments in $\mathcal{A}$, representing different attack criteria. Then, the authors define an *aggregate* argumentation framework $AF^* = \langle \mathcal{A}, \mathcal{F}(\rightharpoonup_1, \ldots, \rightharpoonup_n) \rangle$, where $\mathcal{F}(\rightharpoonup_1, \ldots, \rightharpoonup_n)$ is an attack relation obtained by the aggregation of the individual attack criteria $\rightharpoonup_1, \ldots, \rightharpoonup_n$, via different kinds of mechanisms (e.g. majority voting, qualified voting and mechanisms that can be described by classes of decisive sets). The aggregation of individual attack criteria cannot be assimilated to the kind of mechanisms proposed here. In [32] an individual may sanction an attack between two given arguments while another individual may not, which in terms of labellings means that for the same pair of arguments there may exist the following two labellings: (`in`, `out`) and (`in`, `in`). This is impossible in our setting. Hence, the Arrovian properties (e.g. *Collective Rationality*) are conceived differently.

In [6] the authors analyse the problem of aggregating different individual argumentation frameworks over a common set of arguments in order to obtain a unique socially justified set of arguments. One of the procedures considered there is one in which each individual proposes a set of justified arguments and then the aggregation leads to a unique set of socially justified arguments. The AWPR mechanism proposed here fits this procedure for the special case in which individually justified arguments are simply the sets of arguments labelled `in` for each individual.

There is much work on using an individual agent's preferences to help evaluate arguments (e.g. based on given priorities over rules [26]). But this line of work does not address the preferences of *multiple* agents and how they may be aggregated. In other related work, Bench-Capon [4] associates arguments with values they promote or demote, and considers different audiences with different preferences over those values. Such preferences determine whether particular defeats among arguments succeed. Thus, one gets different argument graphs, one for each audience. Bench-Capon uses this to distinguish between an argument's *subjective acceptance* with respect to a particular audience, and its *objective acceptance* in case it is acceptable with respect to all possible audiences. Our work differs in two important ways. First, in our framework, an agent (or equivalently, an audience) does not have preferences over individual arguments, but rather preferences over how to evaluate all arguments collectively (i.e. over labellings). Secondly, our concern here is not with how individual agents (or audiences) accept an argument, but rather on the possibility of achieving important social-choice properties in the final aggregated labelling.

In relation to aggregation, Coste-Marquis *et al.* explored the problem of aggregating multiple argumentation frameworks [11]. Each agent's judgement consists of a different argument graph altogether. This contrasts significantly with our work, in which agents do not dispute the argument graph, but rather how it must be evaluated/labelled. Our setting is more akin to a jury situation, in which all arguments have been presented by the prosecution and defence team, and are visible to the

Copyedited by: ES MANUSCRIPT CATEGORY: Corner Article

jury members. The jury members themselves do not introduce new arguments, but are tasked with aggregating their individual judgements about the arguments presented to them.

Finally, we refer to the work of Rahwan and Larson [27] on strategic behaviour when arguments are distributed among agents, and where these agents may choose to show or hide arguments. Thus, their interest is in how agents contribute to the *construction* of the argument graph itself, which is then evaluated centrally by the mechanism (e.g. a judge). In contrast, our work is concerned with how agents individually cast votes on how to evaluate each argument in a *given* fixed graph.

Our work opens new research problems for the computational social-choice community. As is the case with other aggregation domains, the aggregation paradox in argument evaluation is an example of a fundamental barrier. Thus the impossibility results are important because they give conclusive answers and focus research in more constructive directions (e.g. weakening the desired properties in order to avoid the paradox). An algorithmic agenda would complement this research by providing efficient algorithms for such problems. Strategic manipulation, by mis-reporting one's true vote, is also an important area of investigation, especially when such manipulations are exercised by coalitions of agents.

Acknowledgements

We are grateful to Andrew Clausen for his insightful comments. We are especially grateful to Martin Caminada for discussions surrounding the second impossibility result. Much of this article was written during a visit of Edmond Awad to the University of Luxembourg which was generously supported by SINTELNET. Richard Booth's work was carried out while working at the University of Luxembourg, and was supported by the FNR (DYNGBaT project).

References

[1] K. Arrow. *Social Choice and Individual Values*. Wiley, 1951.

[2] K. Arrow, A. Sen and K. Suzumura, eds. *Handbook of Social Choice and Welfare*. Vol 1. Elsevier Science Publishers, 2002.

[3] P. Baroni and M. Giacomin. On principle-based evaluation of extension-based argumentation semantics. *Artificial Intelligence*, **171**, 675–700, 2007.

[4] T. Bench-Capon. Persuasion in practical argument using value-based argumentation frameworks. *Journal of Logic & Computation*, **13**, 429–448, 2003.

[5] T. Bench-Capon and P. Dunne. Argumentation in artificial intelligence. *Artificial Intelligence*, **171**, 619–641, 2007.

[6] G. Bodanza and M. Auday. Social argument justification: some mechanisms and conditions for their coincidence. In *Symbolic and Quantitative Approaches to Reasoning with Uncertainty, 10th European Conference, ECSQARU*. Vol. 5590 of *Lecture Notes in Computer Science*, C. Sossai and G. Chemello, eds, pp. 95–106. Springer, 2009.

[7] R. Booth, M. Caminada, M. Podlaszewski and I. Rahwan. Quantifying disagreement in argument-based reasoning. In *Proceedings of the 11th International Conference on Autonomous Agents and Multiagent Systems-Volume 1*, pp. 493–500. International Foundation for Autonomous Agents and Multiagent Systems, 2012.

[8] M. Caminada. On the issue of reinstatement in argumentation. In *Proceedings of the 10th European Conference on Logics in Artificial Intelligence (JELIA)*. Vol. 4160 of *Lecture Notes in Computer Science*, M. Fisher, W. van der Hoek, B. Konev and A. Lisitsa, eds, pp. 111–123. Springer, 2006.

Copyedited by: ES MANUSCRIPT CATEGORY: Corner Article

[9] M. Caminada and D. M. Gabbay. A logical account of formal argumentation. *Studia Logica*, **93**, 109–145, 2009.
[10] M. Caminada and G. Pigozzi. On judgement aggregation in abstract argumentation. *Autonomous Agents and Multi-Agent Systems*, **22**, 64–102, 2011.
[11] S. Coste-Marquis, C. Devred, S. Konieczny, M.-C. Lagasquie-Schiex and P. Marquis. On the merging of Dung's argumentation systems. *Artificial Intelligence*, **171**, 730–753, 2007.
[12] F. Dietrich. A generalised model of judgement aggregation. *Social Choice and Welfare*, **28**, 529–565, 2007.
[13] E. Dokow and R. Holzman. Aggregation of binary evaluations with abstentions. *Journal of Economic Theory*, **145**, 544–561, 2010.
[14] E. Dokow and R. Holzman. Aggregation of non-binary evaluations. *Advances in Applied Mathematics*, **45**, 487–504, 2010.
[15] P. M. Dung. On the acceptability of arguments and its fundamental role in nonmonotonic reasoning, logic programming and n-person games. *Artificial Intelligence*, **77**, 321–358, 1995.
[16] W. Gärtner. *A Primer on Social Choice Theory*. Oxford University Press, 2006.
[17] A. Gibbard. Manipulation of voting schemes. *Econometrica*, **41**, 587–601, 1973.
[18] U. Grandi. *Binary Aggregation with Integrity Constraints*. PhD Thesis, 2012.
[19] D. Grossi and G. Pigozzi. *Judgment Aggregation: A Primer*. Morgan & Claypool, 2014.
[20] C. List. The theory of judgement aggregation: An introductory review. *Synthese*, **187**, 179–207, 2012.
[21] C. List and P. Pettit. Aggregating sets of judgements: An impossibility result. *Economics and Philosophy*, **18**, 89–110, 2002.
[22] C. List and B. Polak. Introduction to judgement aggregation. *Journal of economic theory*, **145**, 441–466, 2010.
[23] C. List and C. Puppe. Judgment aggregation: a survey. In *The Handbook of Rational and Social Choice*, P. Anand, C. Puppe and P. Pattanaik, eds, Oxford University Press, 2009.
[24] H. Moulin. *The Strategy of Social Choice*, Vol. 18. Elsevier, 2014.
[25] E. Muller and M. A. Satterthwaite. The equivalence of strong positive association and strategy-proofness. *Journal of Economic Theory*, **14**, 412–418, 1977.
[26] H. Prakken and G. Sartor. Argument-based logic programming with defeasible priorities. *Journal of Applied Non-classical Logics*, **7**, 25–75, 1997.
[27] I. Rahwan and K. Larson. Mechanism design for abstract argumentation. In *7th International Joint Conference on Autonomous Agents & Multi Agent Systems, AAMAS'2008, Estoril, Portugal*, L. Padgham, D. Parkes, J. Mueller and S. Parsons, eds, pp. 1031–1038, 2008.
[28] I. Rahwan and G. R. Simari, eds. *Argumentation in Artificial Intelligence*. Springer, 2009.
[29] I. Rahwan and F. Tohmé. Collective argument evaluation as judgement aggregation. In *Proceedings of the 9th International Conference on Autonomous Agents and Multiagent Systems: volume 1-Volume 1*, pp. 417–424. International Foundation for Autonomous Agents and Multiagent Systems, 2010.
[30] M. Satterthwaite. Strategy-proofness and Arrow's conditions: Existence and correspondence theorems for voting procedures and social welfare functions. *Journal of Economic Theory*, **10**, 187–217, 1975.
[31] A. Sen. The impossibility of a paretian liberal. *The journal of political economy*, **78**, 152–157, 1970.
[32] F. Tohmé, G. Bodanza and G. Simari. Aggregation of attack relations: a social-choice theoretical analysis of defeasibility criteria. In *Foundations of Information and Knowledge Systems, 5th*

Copyedited by: ES MANUSCRIPT CATEGORY: Corner Article

International Symposium, FoIKS, Vol. 4932 of *Lecture Notes in Computer Science*, S. Hartmann and G. Kern-Isberner, eds, pp. 8–23. Springer, 2008.

[33] P. Vincke. Aggregation of preferences: a review. *European Journal of Operational Research*, **9**, 17–22, 1982.

[34] Y. Wu, M. Caminada and M. Podlaszewski. A labelling-based justification status of arguments. *Studies in Logic*, **3**, 12–29, 2010.

Received 14 March 2015