Journal of Artificial Intelligence Research x (2021) x-x Submitted 5/21; published x/xx

A model to support collective reasoning: Formalization, analysis and computational assessment

Jordi Ganzer JORDI.GANZER_ RIPOLL@KCL.AC.UK
Department of Informatics, King's College London, UK

Natalia Criado NCRIADO@UPV.ES
Universitat Politecnica de Valencia, Spain

Maite Lopez-Sanchez MAITE_LOPEZ@UB.EDU
Faculty of Mathematics and Computer Science, University of Barcelona, Spain

Simon Parsons SPARSONS@LINCOLN.AC.UK
School of Computer Science, University of Lincoln, UK

Juan A. Rodriguez-Aguilar JAR@IIIA.CSIC.ES
Artificial Intelligence Research Institute (IIIA-CSIC), Spain

Abstract

In this paper we propose a new model to represent human debates and methods to obtain collective conclusions from them. This model overcomes two drawbacks of existing approaches. First, our model does not assume that participants agree on the structure of the debate. It does this by allowing participants to express their opinion about all aspects of the debate. Second, our model does not assume that participants' opinions are rational, an assumption that significantly limits current approaches. Instead, we define a weaker notion of rationality that characterises coherent opinions, and we consider different scenarios based on the coherence of individual opinions and the level of consensus. We provide a formal analysis of different opinion aggregation functions that compute a collective decision based on the individual opinions and the debate structure. In particular, we demonstrate that aggregated opinions can be coherent even if there is a lack of consensus and individual opinions are not coherent. We conclude with an empirical evaluation demonstrating that collective opinions can be computed efficiently for real-sized debates.

1. Introduction

This paper is concerned with collective reasoning. By "collective reasoning", we mean the process by which a group of individuals — humans, software agents, or a combination of the two — reach a consensus through a process of debate. In particular, we are interested in the situation in which participants put forward statements, express their opinion about statements put forward by others, and then all of these views are aggregated to identify an opinion that summarises the collective view. This general view of the collective reasoning process can be applied to e-participation systems and, in particular, to participatory democracy, where citizens get involved in decision making about public problems (Noveck, 2009). In fact, the deployment of these systems is nowadays gaining momentum around the globe[1], and therefore, many platforms have been developed (Consul, 2021; Decidim.Org, 2016).

In studying collective reasoning, we draw on work from three main areas: social choice theory, judgement aggregation and argumentation. *Social choice theory* (Aziz et al., 2017; List, 2018) studies approaches for establishing how a group, facing a choice between many alternatives, can make that choice. Given a set of alternatives and a set of agents who possess preference relations over the alternatives, social choice theory focuses on how to yield a collective choice that appropriately reflects the agents' individual preferences. A related line of work, but one which focuses on the acceptability of a single issue, is *judgement aggregation*. This tackles the problem of whether to collectively accept a single issue once the participants have put forward their opinion on it (Endriss & Moulin, 2016; List & Pettit, 2002). *Computational argumentation* (Rahwan & Simari, 2009) focuses more on resolving conflicts in opinions. Given a set of arguments for particular options, and a set of relations (typically conflicts, but also support) between the arguments, argumentation is concerned with identifying those arguments that might be accepted by a rational agent, for different ideas of what makes an argument acceptable. Combining judgement aggregation and social choice theory with argumentation we find several proposals that structure debates using arguments and attack relationships (Awad et al., 2017b; Leite & Martins, 2011), or attack and defence (or support) relationships (Ganzer-Ripoll et al., 2019). These allow participants to put forward arguments, relations between arguments, and opinions about these arguments. They then produce an output that is intended to reflect the collective opinion of the participants on the the debate.

This paper proposes and analyses a new formal model, which we call the "relational reasoning model", that provides an alternative approach to collective reasoning. The specific contributions of this work are:

- *A new formal model.* We present our *relational reasoning model*, a formal model which extends previous frameworks for collective reasoning. It extends argumentation-based approaches by: providing a more flexible notion of relationships between statements, not being restricted to attack and support (Awad et al., 2017b; Ganzer-Ripoll et al., 2019; Leite & Martins, 2011); allowing the structure of opinions to be expressed unlike approaches based on abstract argumentation (Awad et al., 2017b; Coste-Marquis et al., 2007; Leite & Martins, 2011); and introducing "coherence" as a less rigid requirement for the relationship between opinions than the ideas of acceptability used in argumentation. Our approach is also more flexible in allowing the expression of opinions about both statements, as the work of Ganzer *et al.* (2019) and Leite and Martins (2011), and about the relationships between statements, as in the work by Dunne *et al.* (2011)[2]. No existing model allows both. In a further extension to approaches like that of Awad *et al.* (2017b), we allow opinions to be real-valued rather than discrete valued.

1. For example, regarding participatory democracy https://www.direct-democracy-navigator.org/ reports, as for October 2022, 1995 legal designs in 106 countries.
2. The paper by Dunne *et al.* (2011) is not about combining collective opinions on relationships between arguments, but it provides the groundwork for such a system by studying argumentation where the relationships between arguments have different weights.

- *Families of aggregation functions.* We propose a number of opinion aggregation functions that use the participants' opinions about a debate to compute a collective view. These functions make different use of the dependencies between opinions. We provide two families of functions with members that ignore dependencies at all, and other members that use dependencies in different ways. These families of functions together span all the ways in which the dependencies can be taken into account.
- *Formal and computational analysis.* We assess these families of functions against a wide-ranging set of properties adapted from the social choice literature (List & Pettit, 2002). The assessment makes use of four scenarios that make different assumptions about the opinions of participants. We follow the formal analysis with a computational analysis. This first computes the computational complexity of the aggregation functions, and then provides an empirical analysis of those functions when computing collective opinions for a range of scenarios that are larger (in terms of statements and number of opinions) than any online debates that we are aware of. This analysis shows that collective opinions can be computed in real time on quite modest hardware.

The structure of this paper is as follows. Section 2 provides an introduction to the relational reasoning model; and Section 3 and Section 4 provide a formal definition of the model. Then, Section 5 defines the problem of computing collective coherence and introduces the properties that will be used to assess aggregation functions, while Section 6 defines a family of aggregation functions and Section 7 uses the properties to analyse the functions in different scenarios (the proofs can be found in Appendix A). Section 8 provides the computational assessment of the model; Section 9 relates the work presented in this paper to other relevant work in the literature; and Section 10 summarises our conclusions.

2. Introducing the relational reasoning model

The *Relational Reasoning Model*, RRM for short, is a model designed to represent a debate where participants discuss a proposal by putting forward information and giving their opinions. In this section we introduce the main elements of the model, and in the following section we present the full formalisation. As depicted at the top of Figure 2, the RRM is composed of two main parts: the structural part, representing the relationships between the information expressed in (or *content* of) a debate; and the interpretative part, representing the participants' *opinions* about the content of a debate.

Content of a debate. The RRM has two main abstract elements that capture the structure of a debate: *statements* and the *relationships* among them. *Statements* represent plain sentences that describe facts such as, for example (in an urban context), s_0= "Building a modern sports centre" or s_1="Diminishing the historical character of the neighbourhood". *Relationships* represent the reasoning that connects statements. For instance, we may consider the reasoning r_1= "Building a modern sports centre will imply diminishing the historical character of the neighbourhood" connecting s_0 to s_1. In general, each relationship connects a set of source statements to some destination statement. We can think of relationships as logical inferences that relate the statements in the debate.

In the RRM we consider debates that discuss a particular subject or proposal, and refer to it as the *target* τ. The target initiates the debate and thus acts as root of the structure of the statements and their relationships. We call this structure a *Directed Relational Framework*, or DRF for short. Figure 1 a) illustrates the DRF that results from considering s_0 to be the target τ and relating it to s_1 through r_1. We graphically represent the DRF as a graph where nodes correspond to statements and relations to arcs. Relations are directed to reflect the direction of reasoning (i.e., from premises to conclusions) in the debate structure.

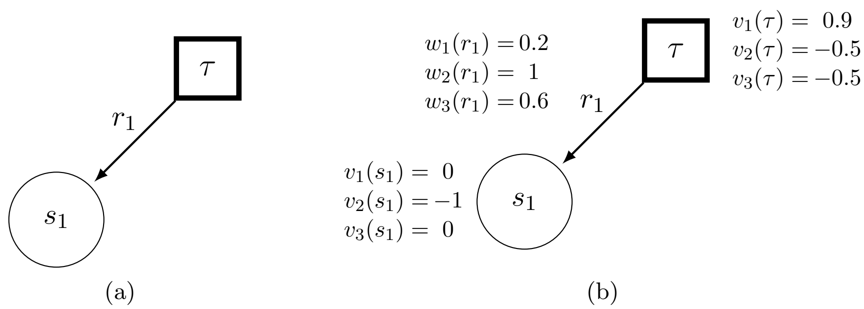

Figure 1: Graphical representation of: a) DRF, the relationship r_1 between proposal τ ("building a modern sports centre") and statement s_1 ("diminishing the historical character of the neighbourhood"); and b) Opinions over the DRF in a).

Participants' opinions. Participants provide their opinions about the elements in the DRF structure. We encode these subjective opinions by means of two functions: the *valuation* function, which assigns values to statements; and the *acceptance* function, that assigns values to the relationships. The valuation function represents the participants' judgement about the statements in the debate. As for the acceptance function, it represents the truth participants see in the statements' relationships (i.e., the reasoning that connects statements). Thus, the desirability or undesirability that each participant feels about each statement of the debate is represented by a positive or negative value assigned with the valuation function, and the conformity that each participant relates to the connections between the statements is represented by an acceptance value assigned by the acceptance function. Following the DRF example in Figure 1 a), Figure 1 b) depicts the opinions of three participants: participant 1 is an indoor sport practitioner that values the target τ very positively ($v_1(\tau) = 1$), is neutral towards s_1 ($v_1(s_1) = 0$), and thinks the implication is somehow weak ($w_1(r_1) = 0.2$); Participant 2 loves history (and values it much more than sports) and fully agrees with relationship r_1 ($w_2(r_1) = 1$, $v_2(s_1) = -1$, and $v_2(\tau) = -0.5$); and participant 3 likes horse riding (rather than indoor sports) and partially agrees with the reasoning ($v_3(\tau) = -0.5$, $w_3(r_1) = 0.6$, and $v_3(s_1) = 0$). In order to compute a collective opinion on the proposal, these individual participants' opinions will be aggregated, as explained in a subsequent paragraph, into a collective opinion by means of an aggregation function.

Direct and Indirect Opinion Considering the influence of opinion (i.e., how opinions about conclusion statements affect the one on their premises), we differentiate direct from indirect opinions: *direct opinion* refers to the value directly given to a statement by a participant; and *indirect opinion* represents those values given to the related statements (i.e., conclusions) and its relationships. We assume opinions on conclusions influence those on their premises. Although it may seem natural to expect participants to be rational and, thus, to provide consistent direct and indirect opinions, inconsistencies may arise. We argue that assuming rationality is too demanding for modelling human debates. In fact, in our previous urban example, one may well envision that a participant hopes for a new sports center (valuing τ positively) even if they are concerned with the historical character of the neighbourhood (valuing s_1 negatively) and accepting the reasoning r_1 relating both statements. Thus, by comparing the direct opinion with the indirect opinion about each statement we can provide a notion of *coherence* that accounts for consistency between these opinions.

Opinion Aggregation Once all users have expressed their opinions about statements and relationships, opinions must be aggregated to calculate a collective opinion. This aggregation can take into account direct opinions, indirect opinions, or a combination of both. In establishing suitable aggregation functions, we have to take into account that individual opinions may be incoherent. Nonetheless, we aim to design aggregation functions that can combine these "imperfect" individual opinions into a "reasonable" collective opinion.

As mentioned in the introduction, the relational reasoning model is more expressive than existing, related, frameworks (Awad et al., 2017b; Ganzer-Ripoll et al., 2016; Leite & Martins, 2011) because it can model situations where participants do not agree on the relationships between different facts. This is captured in the model by the acceptance function. We argue that this acceptance function makes it possible to be explicit about the subjectivity that may be associated to those relationships. Thus, subjectivity may not only be expressed by means of the opinions about statements — i.e., the valuation function — but also through the acceptance function.

3. Formalising the relational reasoning model

From the informal introduction of debates in previous section, we are now ready to formally describe its components and associated processes. This section is devoted to specify both the debate structure and the participants' opinions. In what follows, Figure 2 serves as a reference to follow this formalisation.

3.1 Formalising the structure

First, we introduce the formal notion of a *relational framework* to capture the relationships between statements. Our notion of relationship will consider a non-empty set of source statements and a destination statement. In general, by relating a set of (source) statements to a (destination) statement, we indicate that the source statements support inferring the destination statement, though the framework is agnostic about the form that the support and the inference mechanism takes. For instance, in our example in Figure 3, statements s_2 and s_3 support inferring s_4. Formally:

Definition 3.1. A *relational framework* RF is a pair $\langle \mathcal{S}, \mathcal{R} \rangle$, where $\mathcal{S}$ is a set of statements and $\mathcal{R} \subset \mathcal{P}(\mathcal{S}) \times \mathcal{S} \times \mathbb{N}$ is a relation that does not contain cycles, namely there is no subset of relationships $\{(\Sigma_0, s_1, c_1), \ldots, (\Sigma_{n-1}, s_n, c_n)\} \subseteq \mathcal{R}$ such that $s_i \in \Sigma_i$, $i \in \{1, \ldots, n-1\}$, and $s_n \in \Sigma_0$.

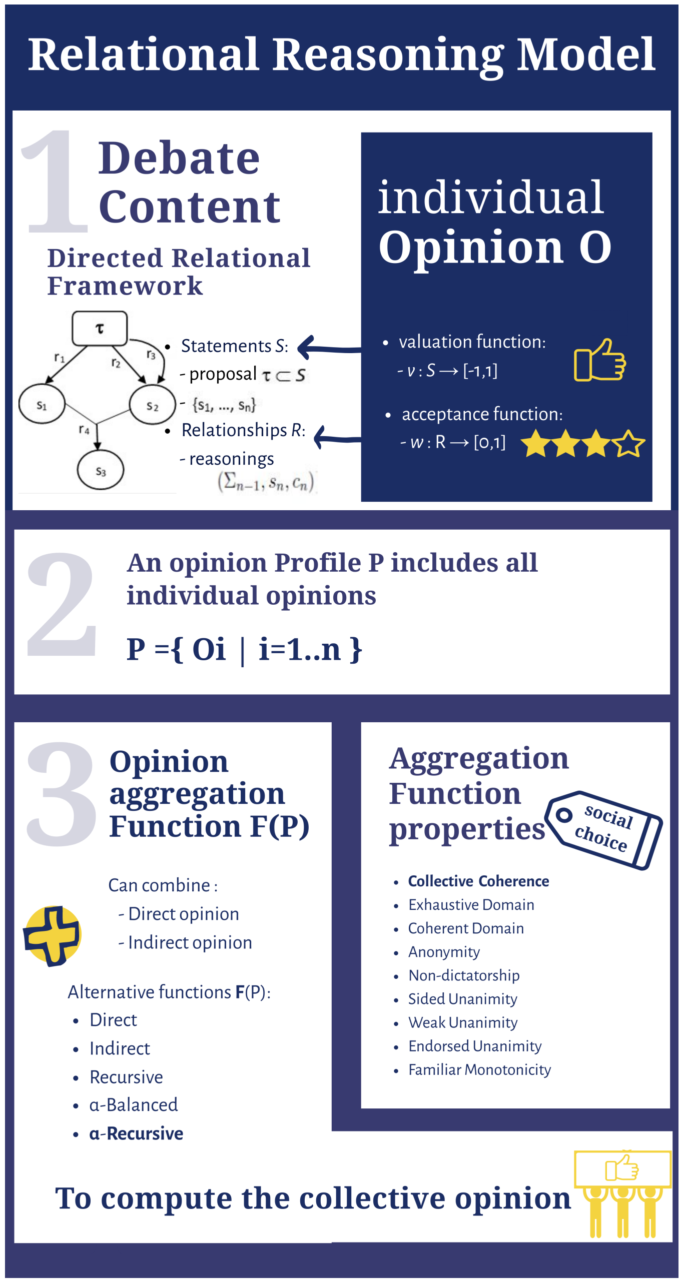

Figure 2: The basic elements of our relational reasoning model.

Since the relation $\mathcal{R}$ is acyclic, it follows that $\mathcal{R}$ is neither *reflexive* ($\forall s \in \mathcal{S}$, $(\Sigma \cup \{s\}, s, c) \notin \mathcal{R}$) nor *symmetric* ($\forall s_1, s_2 \in \mathcal{S}$, if $(\Sigma_1 \cup \{s_1\}, s_2, c_2) \in \mathcal{R}$ then $(\Sigma_2 \cup \{s_2\}, s_1, c_1) \notin \mathcal{R}$). Note that we do not impose any restriction on the transitivity of the relation $\mathcal{R}$. The fact that relational frameworks do not contain cycles is a limitation on what can be represented using our approach, but we do not think it is a serious limitation — we discuss this further in Section 10.

Notice also that we include a natural number within the relation in order to differentiate relationships between the same set of statements Σ and s. From a practical perspective, this allows to signal that alternative relationships can bear on the very same statements[3] (as shown in Figure 3, where target τ is related to statement s_1 through relationships r_1 and r_6).

Now, since debates are aimed at achieving a collective decision on target topics, we extend our definition above to incorporate the notion of *target* statements as follows:

Definition 3.2. A *directed relational framework* (DRF) is a tuple $\langle \mathcal{S}, \mathcal{R}, T \rangle$ such that:

(i) $\langle \mathcal{S}, \mathcal{R} \rangle$ is a relational framework;

(ii) $T \subset \mathcal{S}$ is a set of target statements;

(iii) target statements in T can only be the source of relationships, namely for any relationship $(\Sigma, s, c) \in \mathcal{R}$, $s \notin T$;

(iv) all non-target statements are connected to targets so that for any statement $s \in S$, $s \notin T$, there is a path $\{(\Sigma_0, s_1, c_1), \ldots, (\Sigma_{n-1}, s, c_n)\} \subset \mathcal{R}$ such that $T \cap \Sigma_0 \neq \emptyset$; and

(v) every target statement shares some common descendant with any other target statement, namely for every pair of targets $\tau, \tau' \in T$ there is a statement $s \in S$ such that there is a path from τ to s and another path from τ' to s.

Note that a DRF is constrained to be a *connected acyclic* graph, albeit one that can have several targets. This reflects the idea that, since a DRF represents a single debate, every statement in that debate should have some connection to the rest of the debate.

In what follows we slightly extend our urban example briefly introduced in Section 2 to produce a graphical representation of a DRF that will help us visualise the information in a debate. Recall that our example considered statements τ ("Building a modern sports centre") and s_1 ("Diminishing the historical character of the neighbourhood"), as well as relationship r_1 ("Building a modern sports centre will imply diminishing the historical character of the neighbourhood") connecting both. Besides that, next we consider further statements and relationships as listed in tables 1 and 2 respectively. Finally, figure 3 depicts the connections between statements through relationships. Note that r_4 is a hyperedge, connecting three statements[4].

3. The inclusion of a natural number into the specification of a relationship does not affect the formal contributions of the paper, since, as it will be shown later, relationships are grouped into (and subsequently

Statement	Description
τ	Building a modern sports centre
s_1	Diminishing the historical character of the neighbourhood
s_2	Attraction of more affluent residents to the neighbourhood
s_3	Attraction of new business to the neighbourhood
s_4	Crime reduction in the neighbourhood
s_5	Property values rise in the neighbourhood

Table 1: Statements for the sports centre example.

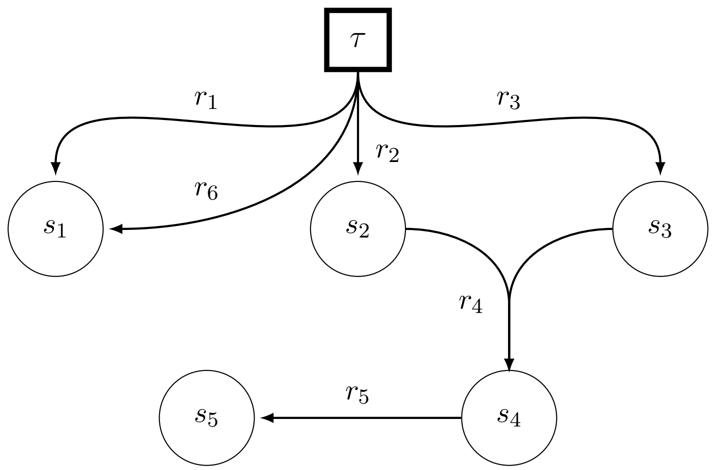

Figure 3: The DRF for the sports centre example.

3.2 Formalising opinions

Now we address the formalisation of the opinions put forward by participants in a debate. We consider that opinions can be held both about statements and relationships. We therefore define two functions that capture the opinions of individuals: (i) a *valuation function* over statements; and (ii) an *acceptance function* over relationships. On the one hand, a valuation function conveys the subjective value that an individual places on each statement. On the other hand, an acceptance function expresses the subjective plausibility that an individual assigns to each relationship, representing the reasoning that connects statements. Formally:

Definition 3.3 (Valuation function)**.** Given a DRF $\langle \mathcal{S}, \mathcal{R}, T \rangle$, a *valuation function* $v : \mathcal{S} \longrightarrow I$ maps each statement to a value in $I = [-1, 1]$.

traversed through) the sets R^+ defined in equation 1 (page 12), without considering any restrictions on the statements they relate.

4. Notice that one participant could introduce an extra relationship from τ to s_5, representing the reasoning "A new community center will give more relevance to the neighbourhood, that will increase the house price", which is not the sum of r_2, r_3 and r_4, but represents a whole new way to connect τ to s_5. This shows that the transitivity allowed in the model, connecting statements from non-consecutive levels of the debate via a single relationship, is not intended to represent the combined reasoning formed by the reasoning steps in between.

Relationship	**Reasoning**	**Connection**
r_1	Building a modern sports centre will imply diminishing the historical character of the neighbourhood.	τ to s_1
r_2	The new sport centre will make the neighbourhood more attractive for wealthy residents because they are more interested in leisure activities.	τ to s_2
r_3	A new community centre will attract more businesses to the surrounding area.	τ to s_3
r_4	Having richer residents and more businesses will increase the security measures around the neighbourhood and therefore, reduce criminal activities.	$\{s_2, s_3\}$ to s_4
r_5	The reduction of crime will increase the price of the houses in the neighbourhood.	s_4 to s_5
r_6	To build a new sport centre, the existing listed building in that location will be demolished.	τ to s_1

Table 2: Reasoning for the sports centre example.

Given a statement $s \in \mathcal{S}$: if $v(s) = 1$ we say that s counts on *full positive valuation*; if $v(s) = -1$ we say that s counts on *full negative valuation*; and if $v(s) = 0$ we say that s has *neutral valuation*.

Definition 3.4 (Acceptance function)**.** Given a DRF $\langle \mathcal{S}, \mathcal{R}, T \rangle$, an *acceptance function* maps each relationship to a value in I^+, $w : \mathcal{R} \longrightarrow I^+ = [0, 1]$.

Given a relationship $r \in \mathcal{R}$ and an acceptance function w, we will refer to the value $w(r)$ as the *acceptance degree* of r. If $w(r) = 1$ we say that the acceptance function expresses *full agreement* with the relationship, whereas if $w(r) = 0$ we say that it expresses *full disagreement*.

Considering our running example, graphically represented in figure 3, figures 4 and 5 show the valuation functions and acceptance functions of agents 1, 2, and 3: v_1, v_2 and v_3 encode agents' valuations for statements, while w_1, w_2 and w_3 encode agents' acceptances of relationships. We consider now the description of agents' opinions in Section 2 to exemplify how they translate into valuations and acceptances. Thus, for instance, agent one is "highly positive" about the target τ ($v_1(\tau) = 0.9$), but neutral regarding statement s_1 ($v_1(s_1) = 0$). Furthermore, agent one considers that the plausibility of relationship r_1 is "little" ($w_1(r_1) = 0.2$).

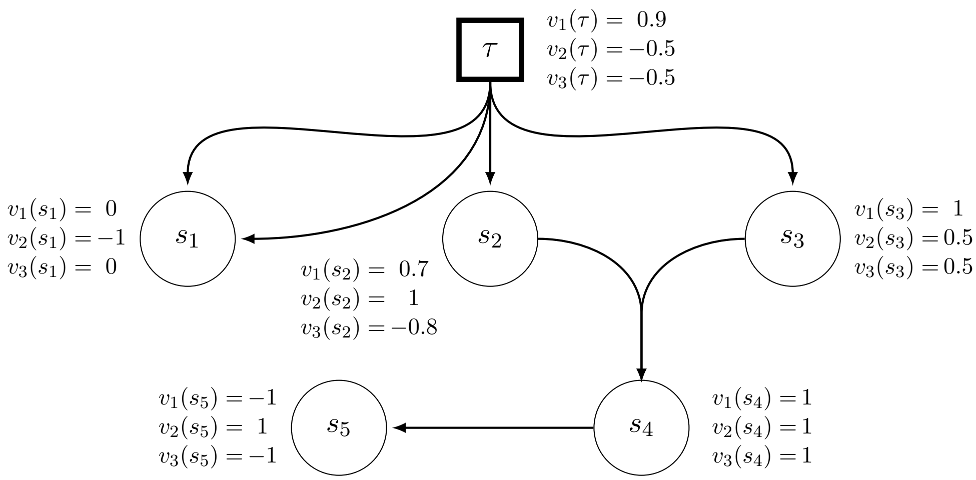

Figure 4: Agents' valuation functions.

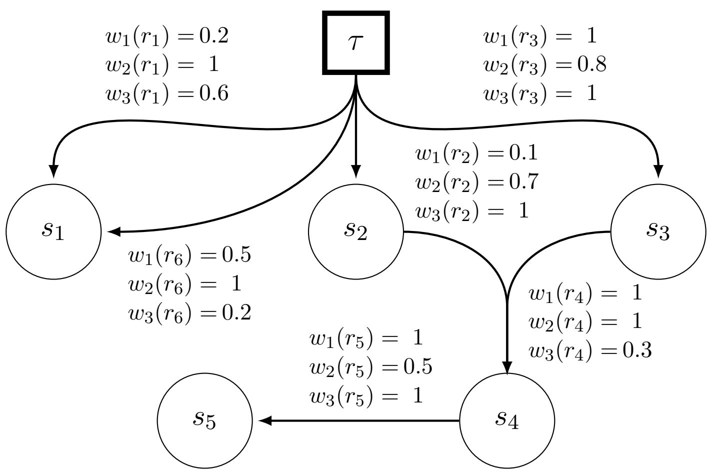

Figure 5: Agents' acceptance functions.

Now we are ready to formally introduce the notion of individual opinion over a DRF.

Definition 3.5 (Opinion)**.** Given a $DRF = \langle \mathcal{S}, \mathcal{R}, T \rangle$, an *opinion* over the DRF is a pair $O = (v, w)$ such that v is a valuation function and w is an acceptance degree.

For practical purposes, we assume that for each relationship, there is at least one individual that provides a non-zero acceptance value for that relationship, otherwise, it would not influence the debate. Henceforth, we shall note the class of all opinions over a DRF as $\mathbb{O}(DRF)$.

As depicted in figures 4 and 5, each agent i involved in a debate will have its individual opinion $O_i = (v_i, w_i)$.

Next, we define our notion of an opinion profile, which brings together the opinions of the individuals involved in a debate. From hereon we use the term "agent" along with the term "individual" to refer to the participants in the debate.

Definition 3.6 (Opinion profile)**.** Let $Ag = \{1, \ldots, n\}$ be a set of n agents and a $DRF = \langle \mathcal{S}, \mathcal{R}, T \rangle$. An *opinion profile* is a collection of opinions $(O_1 = (v_1, w_1), \ldots, O_n = (v_n, w_n)) \in \mathbb{O}(DRF)^n$ over the DRF such that $O_i = (v_i, w_i)$ stands for the opinion of agent i.

Now, our goal will be to compute a collective opinion from the opinions in an opinion profile, which contains the opinions issued by the participants in a debate

4. Characterising coherent opinions

Previous work on the formal modelling of debates has placed restrictions on the opinions that individuals can put forward. For example, Awad et al. (2017b) interprets the opinions expressed by individuals as labels, in the sense of Baroni et al. (2011), for the arguments that they are expressing opinions about. Thus, an argument can be labelled `in`, meaning that the individual thinks that it holds, `out`, meaning that the individual thinks it does not hold, or `undec`, meaning that the individual is not sure whether it holds or not. These labellings are restricted to be *complete* labellings (Baroni et al., 2011), broadly meaning that they conform to a notion of rationality where arguments are `out` if they are attacked by arguments that have been established to be `in`, and are `in` if they are only attacked by arguments that are `out`. We believe that the restrictions imposed by these notions are too restrictive for modelling human debates, as humans may express opinions that are far from rational.

Instead, we impose weaker conditions for an individual opinion to be classified as reasonable or *coherent*, along the lines of our former work (Ganzer-Ripoll et al., 2019). Hence, given a statement, we contrast opinions expressed about that statement, the *direct opinion*, with the opinions expressed about the immediate descendants of the statement, what we call the *indirect opinion*, and look for ways in which these may be made somewhat consistent.

Informally, what we do is the following. First, we compute an estimated opinion for a statement based on the indirect opinion about it. Then, we say that a set of opinions about a statement are *coherent* if the estimated opinion for the statement aligns with the direct opinion about the statement. This will be the case when the opinion (valuations) about the descendants is close to the overall opinion (valuation) about the statement. Considering our example in figure 4 again, consider statement τ, its descendants (s_1, s_2, and s_3), and the opinion of agent 2 (v_2). We observe that although the direct opinion about τ is rather negative ($v_2(\tau) = -0.5$), the valuations for its descendants are diverse: while the valuation for s_1 is also rather negative ($v_2(s_1) = -1$), and hence in line with τ, the valuations on the other descendants are rather positive ($v_2(s_2) = 1$ and $v_2(s_3) = 0.5$), and hence not in line with τ. Thus, at first sight[5] it would seem that the overall estimated opinion is not in line with the direct opinion.

5. Note that we are not considering acceptances at this point.

In what follows, we first formalise our notion of *estimation* as an aggregated measure formed from the indirect opinion about a statement — i.e., the collection of values for the descendants and their relationships. This will consider valuations and acceptance degrees related to a statement and its relationships so that the more accepted a relation between a statement and its descendants, the more important the opinion about that descendant. Thereafter, we will formalise our notion of coherence by measuring how close the direct opinion about a statement is to the estimated opinion about that statement.

First, we introduce some concepts and notations that will aid us in later steps. Given a $DRF = \langle \mathcal{S}, \mathcal{R}, T \rangle$, we define the *set of relationships from* $s \in \mathcal{S}$ as the set of relationships having s in the set of initial statements. Formally,

$$R^+(s) = \{r = (\Sigma, s', c) \in \mathcal{R} \mid s \in \Sigma\}. \tag{1}$$

We will use the term *descendants of a statement s, denoted by $D(s)$, to refer to* any statement s_r connected to s by a relationship r that has s as one of the initial statements and s_r as final statement. Formally,

$$D(s) = \{s_r \in \mathcal{S} \mid r = (\Sigma, s_r, c) \in R^+(s)\}.$$

Next, we define the concept that connects direct and indirect opinion in order to characterise our notion of coherence. We will use an *estimation function* to compute an estimate of the direct opinion using the values gathered for the indirect opinion.

Definition 4.1. Given a $DRF = \langle \mathcal{S}, \mathcal{R}, T \rangle$ and $O = (v, w)$ and opinion over the DRF, the *estimation function* is a valuation function mapping each statement to a value in the set I following the next schematic:

$$\begin{array}{rccc} e: & \mathcal{S} & \longrightarrow & I \\ & s & \longmapsto & e(s) = f(IO(s)) \end{array}$$

where $IO(s) = \{(v(s_r), w(r)) | \ r \in R^+(s)$ and s_r is the descendant attached to $r\}$, and, if $|R^+(s)| = 0$, then $e(s) = v(s)$. Otherwise (if $|R^+(s)| > 0$):

$$f: \quad D = (I \times I^+)^* \quad \longrightarrow \quad I$$

such that

- for any s such that $|R^+(s)| > 0$, D is any $(I \times I^+)^{|R^+(s)|}$.
- f is monotonic with respect to the domain D. I.e. f increases or decreases accordingly to the changes on its inputs, with respect the order endowed in the domain.

In other words, the estimation function computes an estimated value for a statement using the valuations and acceptance degrees for the indirect opinion about that statement. For instance, we can estimate the target $e(\tau)$ in figure 3 by considering the valuations of descendant statements s_1, s_2, and s_3 in figure 4 together with the acceptances of their corresponding relations (r_1, r_2, r_3) in figure 5. However, if the statement has no descendants — as is the case for s_5 — then we simply take its valuation: $e(s_5) = v(s_5)$. This definition is designed to be an abstract definition that allows for many estimation functions to be defined to compute different approximations for the direct opinion. Henceforth, we will use the weighted average of the valuations on the descendants, where the weights are the acceptance degrees on the relations leading to each descendant. In this manner, the more accepted a relation, the more valued the opinion on the descendant. Formally,

$$e(s) = \begin{cases} v(s) & \text{if } R^+(s) = \emptyset \text{ or } \sum_{r \in R^+(s)} w(r) = 0, \\ \frac{\sum_{r \in R^+(s)} w(r) v(s_r)}{\sum_{r \in R^+(s)} w(r)} & \text{otherwise.} \end{cases}$$

Informally, an opinion is characterised as coherent for a given statement when the value assigned by the participant (issuing the opinion) to the statement (i.e., its direct opinion) is aligned with the values and plausibility assigned to its descendants (i.e., its estimate opinion). Furthermore, given the continuous values allowed on the opinion, we can choose the degree of coherence by using a parameter ϵ. Formally:

Definition 4.2 (Coherence)**.** Consider a $DRF = \langle \mathcal{S}, \mathcal{R}, T \rangle$ and an $\epsilon \in (0,1)$ difference[6]. We say that opinion $O = (v, w)$ is ϵ-*coherent* on $s \in \mathcal{S}$ when

$$|v(s) - e(s)| < \epsilon.$$

In general, an opinion O will be ϵ-coherent if it is ϵ-coherent for every statement in $\mathcal{S}$. From the definition of coherence, this means that for every statement its direct and indirect opinions are very close, and hence there is an agreement between direct and indirect opinions.

We will notate as $\mathbb{C}_\epsilon(DRF)$ the class of all the ϵ-coherent opinions. Thus, if O is an ϵ-coherent opinion then $O \in \mathbb{C}_\epsilon(DRF)$.

Example 4.1. Following the example, now we can compare the values from the expectation function and the actual value given by participant 1 to each statement, its direct opinion, see figure 6. We can see that if $\epsilon \in (0.3, 1)$ then the opinion of participant 1 for the statements s_1, s_2, s_3 and s_5 is ϵ-coherent but not for statement s_4 due to the difference between direct opinion and estimated value, which is the maximum possible. Because of this statement s_4, the opinion of participant 1 cannot be classified as ϵ-coherent for any $\epsilon \in (0,1)$.

6. We choose the interval (0,1) for the value of ϵ as the minimum interval that guarantees that if the direct opinion is 1 (or -1) then an opinion cannot be classified as coherent when the estimation function value is of the opposite sign, i.e., $e(s) \not\lesssim 0$ (or $e(s) \not\gtrsim 0$ respectively).

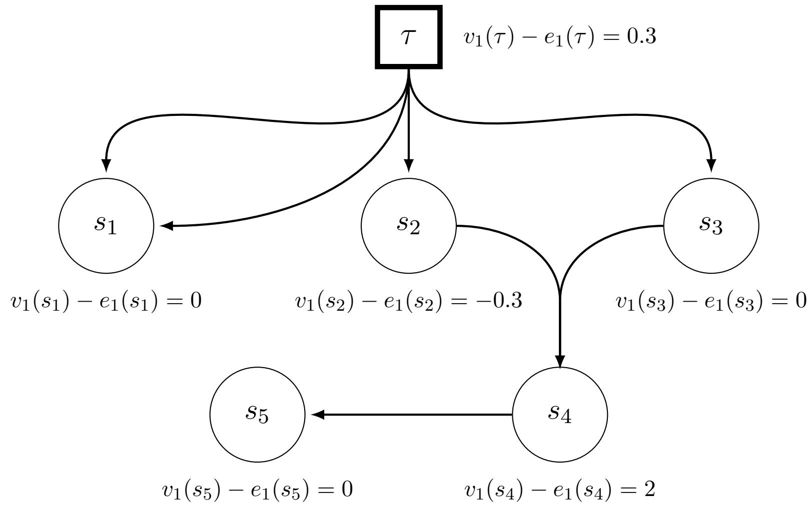

Figure 6: Coherence of Agent 1.

5. Formalising the collective decision making problem

As stated above, our goal is to help agents reach a collective decision on target statements. This corresponds to the third stage in Figure 2. In Section 5.1 we cast our goal as an opinion aggregation problem. We propose to solve such problem using an aggregation function that synthesises a single opinion out of all agents' opinions. Although opinions can be aggregated in different ways, here we follow our previous work (Ganzer-Ripoll et al., 2019) in requiring that the outcome of an aggregation must satisfy desirable social choice properties. In particular, Section 5.2 introduces desirable social choice properties to help analyse and compare opinion aggregation functions.

5.1 The opinion aggregation problem

The problem at hand is how to aggregate the opinions in an opinion profile to produce a single opinion so that single opinion is a reasonable summary of the opinions in the opinion profile. If the opinion profile represents the views expressed by individuals in a debate, the combination should represent the collective opinion of all the individuals. The opinion aggregation function, which we formalise below, is the mechanism for establishing this collective opinion.

Definition 5.1 (Opinion aggregation function)**.** Given a DRF and a set of n agents, a function $F : \mathcal{D} \subset \mathbb{O}(DRF)^n \longrightarrow \mathbb{O}(DRF)$ mapping an opinion profile to a single opinion is called an *opinion aggregation function*. Given an opinion profile P in the domain $\mathcal{D}$, $F(P)$ is called the *collective opinion* by F and it will be noted as $F(P) = (v_{F(P)}, w_{F(P)})$.

In Section 6 we define specific opinion aggregation functions that compute a collective opinion. Before that, we introduce the properties that we will use to analyse aggregation functions.

5.2 Social choice properties

Social choice theory provides formal properties to characterise aggregation methods in terms of outcome fairness (Dietrich, 2007). In what follows, we formally adapt some of the desirable social properties of an aggregation function that were introduced by Awad *et al.* (2017b) and in our previous work (Ganzer-Ripoll et al., 2019). Besides adapting properties, we define some novel properties that characterise aggregation functions motivated by the fact that here we are considering opinions to be continuous-valued in contrast to the discrete-valued opinions used previously (Awad et al., 2017b; Ganzer-Ripoll et al., 2019).

First, we characterise aggregation functions in terms of the opinion profiles that they can take as input. Thus, we adapt from Awad et al. (2017b) the notion of *exhaustive domain* to characterise opinion aggregation functions that are defined for any opinion profile. Thereafter, we modify this property to limit an opinion aggregation function to operate with ϵ-coherent opinion profiles.

Exhaustive Domain (ED). An opinion aggregation function F satisfies *exhaustive domain* if its domain is $\mathcal{D} = \mathbb{O}(DRF)^n$, namely if the function can operate over all profiles.

ϵ-Coherent Domain (ϵ-CD). An opinion aggregation function F satisfies *ϵ-coherent domain* if its domain $\mathcal{D}$ contains all ϵ-coherent opinion profiles, namely $\mathbb{C}_\epsilon(DRF)^n \subseteq \mathcal{D}$.

We will sometimes refer to ϵ-Coherent Domain as "coherent domain". Note an opinion aggregation function satisfying exhaustive domain also satisfies ϵ-coherent domain. In this paper we will just analyse the most general property (e.g., exhaustive domain) being satisfied. We also define *collective ϵ-coherence* as a property characterising opinion aggregation functions that produce ϵ-coherent collective opinions. Therefore, our notion of collective ϵ-coherence here is more relaxed than the crisp notion of coherence we used before (Ganzer-Ripoll et al., 2019).

Collective ϵ-coherence (ϵ-CC). An opinion aggregation function F satisfies *ϵ-collective coherence* if for every opinion profile $P \in \mathcal{D}$, the opinion produced by aggregation function F is ϵ-coherent, namely $F(P) \in \mathbb{C}_\epsilon(DRF)$.

In accordance with our previous work (Ganzer-Ripoll et al., 2019), here we consider ϵ-CC as the most desirable property that can be satisfied by an aggregation function, since collective coherence is the foundation of the acceptability of collective decisions (Thagard, 2002). Notice also that, as in (Ganzer-Ripoll et al., 2019), collective coherence can be regarded as the counterpart of the notion of Awad *et al.*'s *collective rationality* (Awad et al., 2017b), which states that output of the aggregation should be a *complete* labelling. It is also worth noticing the difference between the notions of coherence for an opinion and collective coherence for an aggregation function. On the one hand, deciding on whether a single opinion is coherent or not (definition 4.2) is based on assessing whether the direct opinions and indirect opinions *agree* (are close) for every statement. On the other hand, the notion of collective coherence refers to the capability of an aggregation function to output a coherent opinion when aggregating the opinions in a profile.

Next, anonymity and non-dictatorship characterise the importance of the agents involved in a debate that yields a collective opinion. On the one hand, anonymity is a social choice property requiring that the opinions of all the agents involved in a debate are considered to be equally significant. On the other hand, non-dictatorship requires that no agent overrules the opinions of the rest of agents.

Anonymity (A) Let $P = (O_1, \ldots, O_n)$ be an opinion profile in $\mathcal{D}$, σ a permutation over Ag, and $P' = (O_{\sigma(1)}, \ldots, O_{\sigma(n)})$ the opinion profile resulting from applying σ over P. An opinion aggregation function F satisfies *anonymity* if $F(P) = F(P')$.

Non-Dictatorship (ND). An opinion aggregation function F satisfies *non-dictatorship* if no agent $i \in Ag$ satisfies that $F(P) = O_i$ for every opinion profile $P \in \mathcal{D}$.

Notice that non-dictatorship is a weaker version of anonymity since it follows directly from it — any aggregation function that satisfies anonymity will satisfy non-dictatorship. Again, we will just analyse anonymity for the different aggregation functions.

Now we consider how an opinion aggregation function behaves when agents agree on their opinions about statements. Unanimity is the social choice property that characterises the behaviour of aggregation functions when there is agreement among agents' opinions. The classical notion of unanimity defines unanimity as a situation in which all agents share the very same opinion. While this is quite possible in settings where agents only have a few discrete possibilities for expressing their opinion (Awad et al., 2017b; Ganzer-Ripoll et al., 2019), it is not likely to occur in the setting we are studying here, where opinions can take a wide range of values. As a result, we propose some relaxed variations which are more useful for the setting we consider. First, we say that *sided unanimity* will hold when, for each statement, either all opinions on it are positive or negative. Formally,

Sided Unanimity (SU). Let $P = (O_1, \ldots, O_n)$ be an opinion profile, where $P \in \mathcal{D}$. An opinion aggregation profile F satisfies *sided-unanimity* if for every $s \in S$:

- if $v_i(s) > 0$ for all $i \in Ag$ then $v_{F(P)}(s) > 0$;
- if $v_i(s) < 0$ for all $i \in Ag$ then $v_{F(P)}(s) < 0$.

We also find a weaker version of sided unanimity to be worth distinguishing:

Weak Unanimity (WU). Let $P = (O_1, \ldots, O_n)$ be an opinion profile, where $P \in \mathcal{D}$. An aggregation profile F satisfies *Weak unanimity* if, for every $s \in S$:

- if $v_i(s) = 1$ for all $i \in Ag$ then $v_{F(P)}(s) > 0$;
- if $v_i(s) = -1$ for all $i \in Ag$ then $v_{F(P)}(s) < 0$.

Although WU requires that all agents agree on fully positive (1) or fully negative (-1) valuations on statements, it does not require that the output of the opinion aggregation function takes one of those values, as the usual version of unanimity would (hence the name). Weak unanimity has value when translating valuations expressed in a discrete model (Awad et al., 2017b; Ganzer-Ripoll et al., 2019) into our model, and so has value in allowing us to relate our model to those which came before.

From the definitions above, it follows that the two notions of unanimity are related.

Proposition 5.1 (Unanimity relationships)**.** If an opinion aggregation function satisfies Sided Unanimity then it satisfies Weak Unanimity.

Proof. Clearly, if an aggregation function cannot hold the sign of the aggregation when the assumptions of Weak Unanimity are satisfied, then it is straightforward to see that will fail to satisfy Sided Unanimity. □

As a final unanimity property, we adapt the notion of endorsed unanimity from Ganzer et al. (2019) to consider unanimity based on indirect opinions. In short, an opinion aggregation function satisfies *endorsed unanimity* if, for each statement, the collective opinion on the statement is in line with the unanimous indirect opinion on it. Formally,

Endorsed Unanimity (EU). Let $P = (O_1, \dots, O_n)$ be an opinion profile such that $P \in \mathcal{D}$. An aggregation profile F satisfies *endorsed unanimity* if for every $s \in \mathcal{S}$:

(i) if $v_i(s_d) = 1$ for any $i \in Ag$ and $s_d \in D(s)$ (called *full positive support*), then $v_{F(P)}(s) > 0$; and

(ii) if $v_i(s_d) = -1$ for any $i \in Ag$ and $s_d \in D(s)$ (called *full negative support*), then $v_{F(P)}(s) < 0$.

Next, we introduce monotonicity properties to study how the result of an opinion aggregation function changes as opinions change. First, we adapt the notion of monotonicity from Awad et al. (2017b). In particular, Awad et al. (2017b)'s notion of monotonicity states that if some of the direct opinions about a statement increase (or decrease) the collective opinion should increase (or decrease) accordingly. That notion only takes into account the direct opinion about each statement. Since we aim at handling opinion aggregation functions that merge both direct and indirect opinions, we adapt the notion of *familiar monotonicity*[7] from (Ganzer-Ripoll et al., 2019). In our case, familiar monotonicity requires that when the direct opinion on a statement increases, the collective opinion does not decrease provided that the opinions on the descendants of the statement do not change either. Formally:

Familiar Monotonicity (FM). Let $s \in S$ be a statement, and $P = (O_1, \dots, O_n)$ and $P' = (O'_1, \dots, O'_n)$ such that for every opinion i satisfies that $v_i(s) \leq v'_i(s)$, and, $w_i(r) = w'_i(r)$ and $v_i(s_r) = v'_i(s_r)$ for every relationship $r \in R(s)$ and its associated descendant $s_r \in D(s)$. We say that an opinion aggregation function F satisfies FM if $v_{F(P)}(s) \leq v_{F(P')}(s)$.

Having listed these properties, we note that they are not all equally important. For a multi-party discussion, we believe that the most important property is collective coherence. If an aggregation function is collectively coherent, the resulting combined opinion will be coherent regardless of the coherence of the initial opinions that are being merged. In other words, an aggregation function that satisfies collective coherence will always discover a coherent overall opinion no matter how incoherent are the opinions on which it is based. Along with collective coherence, the property that we would like to see for an aggregation function is exhaustive domain (and, hence, coherent domain) because it allows for broad

7. The name derives from the fact that this form of monotonicity takes into account opinion about the descendents of a statement which make up its family.

applicability of the aggregation function. Finally, we regard the usual social choice property of anonymity (and, by extension, non-dictatorship) as essential.

Among the unanimity properties, we find sided, weak and endorsed unanimity, which allow some maneuverability when using dependencies to build the collective opinion, to be more desirable than classical unanimity. Though it is natural to require some form of monotonicity, we do not consider the classical monotonicity property to be desirable because it entails discarding indirect opinions. Thus, in its place we prefer familiar monotonicity, which takes into consideration indirect opinion. Finally, since we focus on the design of aggregation functions towards the use of both direct and indirect opinions, we do not include independence which, although commonly used in social choice, just depends on direct opinions and disregards indirect opinions. Any reader interested in an analysis of the classical notions of unanimity, monotonicity and independence for the relational reasoning model can find it in an extended version of this paper (Ganzer et al., 2020).

6. Aggregation functions to enact collective decision making

Since there are dependencies between statements in a relational reasoning model, the question when designing an aggregation function is how to exploit dependencies, which fundamentally amounts to deciding how to exploit indirect opinions as well as direct opinions. In this section we design two families of aggregation functions for the relational reasoning model that allow us to compute collective decisions. Both families use some some combination of direct and indirect opinions, as introduced in Section 2. On the one hand, the family of *balanced* aggregation functions proposes a linear combination of direct opinions and indirect opinions (from *immediate* descendants). On the other hand, the family of *recursive* aggregation functions proposes to further exploit dependencies. Thus, it proposes a linear combination of direct opinions and indirect opinions, but this time indirect opinions come from considering *all* descendants.

Our aim, later on (in Section 7), is to study both families in terms of the social choice properties that they satisfy. Both families are defined in terms of some *base* aggregation functions that we introduce first.

We start by defining a function that only aggregates direct opinions, hence disregarding indirect opinions. This function obtains a collective opinion by averaging valuations per statement and acceptance degrees per relation from the individual opinions in an opinion profile. Formally:

Definition 6.1 (Direct aggregation)**.** Let $\langle \mathcal{S}, \mathcal{R}, T \rangle$ be a DRF and $P = (O_1, \ldots, O_n)$ an opinion profile over the DRF. The *direct aggregation* of P over the DRF is defined as a function $D(P) = (v_{D(P)}, w_{D(P)})$, where $v_{D(P)}(s) = \frac{1}{n}\sum_{i=1}^{n} v_i(s)$ and $w_{D(P)}(r) = \frac{1}{n}\sum_{i=1}^{n} w_i(r)$ for any statement $s \in \mathcal{S}$ and relationship $r \in \mathcal{R}$.

Next, we define a function that only aggregates indirect opinions, disregarding direct opinions. Therefore, it is the converse of the direct aggregation function. The aggregation of indirect opinions employs the aggregation estimations, hence employing the estimation functions from definition 4.1. Formally:

Definition 6.2 (Indirect aggregation)**.** $\langle \mathcal{S}, \mathcal{R}, T \rangle$ be a DRF and $P = (O_1, \ldots, O_n)$ an opinion profile over the DRF. The *indirect aggregation* of P over the DRF is defined as a function $I(P) = (v_{I(P)}, w_{I(P)})$, where $v_{I(P)}(s) = \frac{1}{n}\sum_{i=1}^{n} e_i(s)$, where e_i is an estimation function, and $w_{I(P)}(r) = \frac{1}{n}\sum_{i=1}^{n} w_i(r)$ for any statement $s \in \mathcal{S}$ and relationship $r \in \mathcal{R}$.

While the direct aggregation function computes the average of individuals' direct opinions, the indirect aggregation function computes the average of individuals' indirect opinions as estimated opinions. Observe that both functions calculate the aggregation of acceptance degrees likewise. This is the case for all the aggregation functions defined in this section, and hence the difference between them lies in the aggregation of valuations.

At this point, we are ready to introduce our first family of aggregation functions, which is based on a linear combination of the direct and indirect aggregation functions.

Definition 6.3 (α-Balanced aggregation)**.** Let $\langle \mathcal{S}, \mathcal{R}, T\rangle$ be a DRF and $P = (O_1, \ldots, O_n)$ an opinion profile over the DRF. Given the direct aggregation $D(P) = (v_{D(P)}, w_{D(P)})$, the indirect aggregation $I(P) = (v_{I(P)}, w_{I(P)})$, we define the α-balanced aggregation function $B_\alpha(P) = (v_{B_\alpha(P)}, w_{B_\alpha(P)})$, with $\alpha \in [0, 1]$, for any statement $s \in \mathcal{S}$ and relationship $r \in \mathcal{R}$, such that:

$$v_{B_\alpha(P)} = \alpha \cdot v_{D(P)} + (1 - \alpha) \cdot v_{I(P)}$$
$$w_{B_\alpha(P)}(r) = \frac{1}{n}\sum_{i=1}^{n} w_i(r).$$

By changing the value of α we vary the importance of the direct opinion with respect to the indirect opinion. The functions resulting from definition 6.3 form a family of balanced aggregation functions: $\{B_\alpha\}_{\alpha\in[0,1]}$. In particular, by setting α to 0 we obtain the indirect aggregation function, and by setting it to 1 we obtain the direct aggregation function. Figure 7 shows the valuations of the direct, indirect, and α-balanced aggregation functions for the opinion profile of our running example in figure 4.

Next, we define another base aggregation function that exploits indirect opinions differently. For a given statement, the so-called *recursive aggregation* function calculates its aggregated valuation by using the collective opinion on its descendants, which in turn is recursively computed from their descendants, and so on. This recursive computing ends up when reaching statements without descendants whose indirect opinion is empty. Therefore, the recursive aggregation, unlike balanced aggregations, disregards individual valuations in the indirect opinion, and employs their collective opinions instead.

Definition 6.4 (Recursive aggregation)**.** Let $\langle \mathcal{S}, \mathcal{R}, T\rangle$ be a DRF and $P = (O_1, \ldots, O_n)$ an opinion profile over the DRF. The *recursive aggregation* of P over the DRF is defined as a function $R(P) = (v_{R(P)}, w_{R(P)})$ for any statement $s \in \mathcal{S}$ and relationship $r \in \mathcal{R}$ such that:

$$v_{R(P)}(s) = \begin{cases} \frac{1}{\sum_{r\in R^+(s)} w_{R(P)}(r)} \sum_{r\in R^+(s)} v_{R(P)}(s_r) \cdot w_{R(P)}(r) & \text{if } R^+(s) \neq \emptyset \\ v_{D(P)}(s) & \text{otherwise} \end{cases}$$

and $w_{R(P)}(r) = \frac{1}{n}\sum_{i=1}^{n} w_i(r)$.

Recall that $R^+(s)$ stands for the relationships connecting s to a descendant s_r of s through the relationship r. Thus, the recursive function computes the average of the indirect

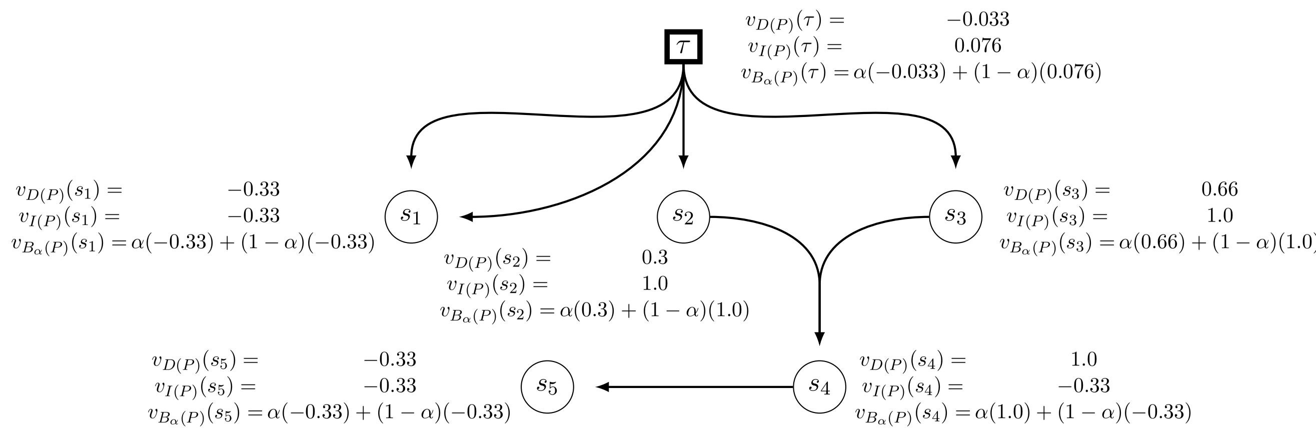

Figure 7: Direct (D), Indirect (I) and α-Balanced (B_α) aggregation functions.

collective opinion computed so far. In fact, we could say that, due to its recursive character, the function computes the estimated opinion for each statement in a bottom-up manner. The aggregation of opinions starts considering the direct opinions at the “leaves” of the debate, namely at the statements with no descendants, and moves up until reaching the targets.

At this point, we are ready to define our second family of aggregation functions by combining the direct and recursive aggregation functions.

Definition 6.5 (α-recursive aggregation)**.** Let $\langle \mathcal{S}, \mathcal{R}, T\rangle$ be a DRF and $P = (O_1, \ldots, O_n)$ an opinion profile over the DRF. Given the direct aggregation $D(P) = (v_{D(P)}, w_{D(P)})$, the recursive aggregation $R(P) = (v_{R(P)}, w_{R(P)})$, we define the α-recursive aggregation function $R_\alpha(P) = (v_{R_\alpha(P)}, w_{R_\alpha(P)})$, with $\alpha \in [0, 1]$, for any statement $s \in \mathcal{S}$ and relationship $r \in \mathcal{R}$ such that:

$$v_{R_\alpha(P)} = \alpha \cdot v_{D(P)} + (1-\alpha) \cdot v_{R(P)}$$

$$w_{R_\alpha(P)}(r) = \frac{1}{n}\sum_{i=1}^{n} w_i(r).$$

Figure 8 shows the valuations of the direct, recursive, and α-recursive aggregation functions for the opinion profile of our running example in figure 4.

7. Analysing opinion aggregation functions

In this section we compare the aggregation functions introduced in Section 6 in terms of their satisfaction, or otherwise, of the social choice properties introduced in Section 5. Our analysis will run along two dimensions: (1) the *coherence* of an opinion profile; and (2) the *consensus* on the acceptance degrees of an opinion profile. Thus, we will consider whether agents’ opinions are constrained to be coherent (the opinion profile is coherent) or not, and whether agents agree on acceptance degrees (there is consensus on acceptance degrees) or not. This results in four debate scenarios to analyse:

1. Unconstrained opinion profiles;
2. Constrained opinion profiles: assuming consensus on acceptance degrees;
3. Constrained opinion profiles: assuming coherent profiles; and
4. Constrained opinion profiles: assuming consensus on acceptance degrees and coherent profiles.

The analysis of these scenarios will help us assess the price to pay, in terms of social choice properties, if the opinions stated by participating agents are not necessarily coherent. The scenarios will also help us assess the price to pay when the relationships between statements are open for discussion by means of acceptance degrees.

Scenario 2 takes inspiration from multiple debate systems that are already in use, which do not allow participants to value the relationships between sentences in a debate differently.. In such systems the relationships, which provide the structure of the debate, cannot

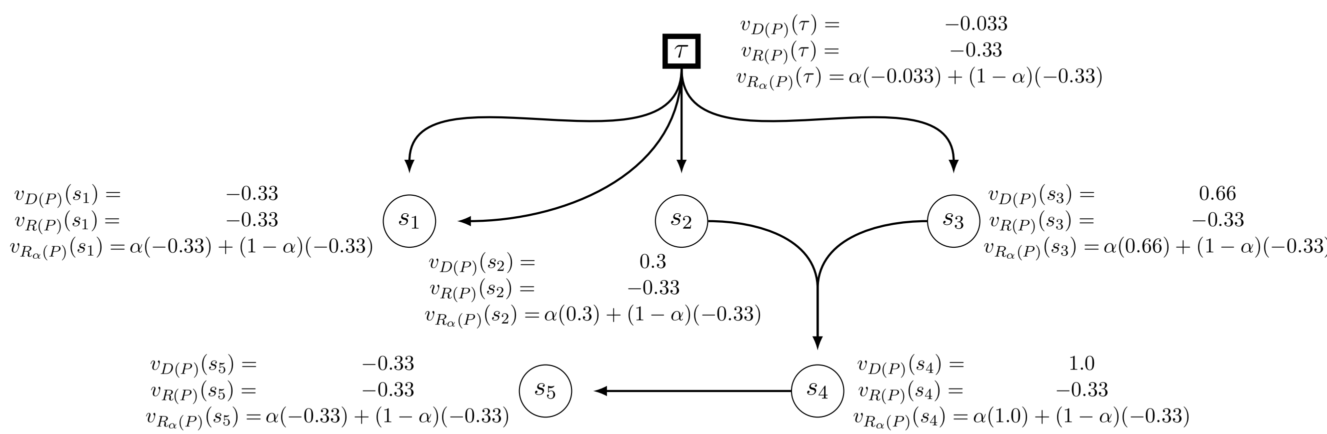

Figure 8: Direct (D), Recursive (R) and α-Recursive (R_α) aggregation functions.

be questioned. Scenarios 3 and 4 capture debates in which participants act rationally, hence exhibiting coherence. Finally, scenario 1 represents the most open scenario, where participants are not expected to act rationally and where even the relationships between sentences are open to discussion.

For the sake of readability, we do not include the formal analysis in the body of the paper. Instead, we present the main results of our analysis here while Appendix A details our formal analysis, including the proofs of the results reported hereafter.

7.1 Unconstrained opinion profiles

This is the most general scenario that we consider. We assume unconstrained opinion profiles, which means that *any* opinion profile is deemed to be possible input for the aggregation functions introduced in Section 6. In other words, the domain of our aggregation functions is the class $\mathbb{O}(DRF)^n$ itself, and hence opinions need not be coherent nor agree on acceptance degrees.

Table 3 displays the social choice properties fulfilled by the functions defined in Section 6 in this general case. There is one column per aggregation function and one row per social choice property. In the table, a green square (with a tick) indicates that a property is fulfilled, while a red square (with a cross) indicates that a property is not fulfilled. As to the more general aggregation functions, α-B(alanced), and α-R(ecursive), in some cases we specify the values α for which a given property holds. Notice that, for both families, we show the results considering $\alpha \in (0, 1)$, not considering 0 or 1. The cases for the extreme values (0 and 1) represent aggregations functions displayed in other columns.

Desirable properties	D	I	R	α-B	α-R
Collective coherence ☆	✗	✗	✓	✗	$\alpha < \epsilon/2$
Exhaustive domain	✓	✓	✓	✓	✓
Anonymity	✓	✓	✓	✓	✓
Sided Unanimity	✓	✗	✗	✗	✗
Weak Unanimity	✓	✗	✗	$\alpha > 1/2$	$\alpha > 1/2$
Endorsed Unanimity	✗	✓	✗	$\alpha < 1/2$	✗
Familiar monotonicity	✓	✓	✗	✓	✗

Table 3: Social choice properties satisfied by aggregation functions D(irect), I(ndirect), R(ecursive), α-B(alanced), and α-R(ecursive) for: (i) a general scenario considering unconstrained opinion profiles; (ii) a scenario considering constrained opinion profiles: consensus on acceptance degrees.

Domain and anonymity. Table 3 shows that Exhaustive Domain (ED) (and by extension Coherent Domain), and Anonymity (and Non-dictatorship) are fulfilled by all the proposed opinion aggregation functions. Since no constraints are imposed on opinion profiles received as input, ED is satisfied, and since no agent in an opinion profile receives a special treatment, Anonymity holds.

Collective coherence. The (D)irect, (I)ndirect and α-Balanced functions do not satisfy collective coherence. The result of such aggregation methods largely depends on the coherence of the opinion profile at hand, and in this scenario that can be incoherent. More positively, the (R)ecursive aggregation function does satisfy Collective Coherence (CC). Out of the family of recursive aggregation functions (α-R), which rely on D and R, those for which $\alpha < \epsilon/2$, where ϵ is set to assess the coherence of the output, also satisfy CC. This tells us that the closer is α to 0 (the less the use of the direct opinion), the more coherent the collective opinion obtained by an α-R function will be. The closer α is to $\epsilon/2$, the less coherent the collective opinion obtained by an α-R function will be. When α goes beyond $\epsilon/2$, the α-R function depends too much on the direct opinion (which does not satisfy CC) and CC does not hold.

Unanimity. Sided and Weak Unanimity are not satisfied by the Indirect and Recursive aggregation functions. This is because the indirect opinion, employed by all these aggregation functions, ignores unanimity on the direct opinion of a statement and in some cases these functions can produce a result in the opposite direction. On the other hand, the Direct function, which only depends on the direct opinions of a statement, does satisfy all the unanimity properties. This benefits the Balanced and Recursive families, which satisfy Weak unanimity for some values of α. Notice that only Balanced and Recursive aggregation functions for which α is greater than 1/2 satisfy Weak unanimity. Such values of α lessen the influence of the indirect opinion and sway the result towards the Direct aggregation function, which does satisfy the property. Regarding Sided Unanimity, not even the influence of the Direct function is enough to guarantee that unanimity is preserved, and therefore no aggregation function in the Balanced or Recursive families fulfil it for any value of $\alpha \in (0, 1)$.

On the other hand, regarding Endorsed unanimity, the situation changes for the Direct and Indirect functions. They flip sides so that the Direct function does not fulfill Endorsed unanimity, but the Indirect function does. This is because unanimity in this case resides in indirect opinions, and hence it is in line with the Indirect function, which only depends on indirect opinions. However, this goes against the Direct function, which disregards indirect opinions, and hence unanimity on its values. Conversely to the Weak unanimity case for the Balanced family, now we need that the values of α are less than 1/2 to sway the balanced aggregation towards the Indirect function, and hence, satisfy Endorsed unanimity. Next, although it might seem reasonable that aggregation functions in the Recursive family also fulfil Endorsed unanimity, they do not. This is caused by the recursive behaviour of these aggregation functions, which can overlook unanimity on indirect opinions to use instead opinions deep in the debate on which there might be no unanimity. And last, due to the failure of the Direct and Recursive aggregation functions to fulfil Endorsed unanimity, so do all the aggregation functions in the Recursive family, no matter the value of α.

Monotonicity. The Familiar Monotonicity property is fulfilled by the Direct function (as a consequence of fulfilling Monotonicity), the Indirect function, and therefore by the whole family of Balanced functions that are combinations of the Direct and Indirect functions. The Recursive function, and hence the Recursive family, fails to satisfy Familiar Monotonicity because, given a statement, the aggregated opinion about its descendants does not solely depend on the valuations on these descendants alone. Instead, the aggregated opinion about its descendants recursively depends on descendants down the relational framework. Thus, changes of opinion on "grandchildren" statements can cause a change of opinion independently of any change of the direct opinion.

7.2 Constrained opinions: assuming consensus on acceptance degrees

As previously mentioned, multiple debate systems that are already in use do not allow participants to value the relationships between sentences differently because these relationships, which provide the structure of the debate, cannot be questioned. Therefore, we take inspiration from these systems and consider this second scenario that assumes consensus on acceptance degrees. Recalling the example introduced in Section 3 (see tables 1 and 2 and Figure 3), this means that participants can value s_1 ("diminishing of the historical character of the neighbourhood") differently but we assume they all agree in its implication relationship with τ ("building a modern sports centre"). In this manner, Figure 4 still holds, but Figure 5 will not be considered.

Assuming consensus on acceptances does not lead to any gain in the fulfilment of social choice properties with respect to those already claimed in Section 7.1, so previous Table 3 show the results for this scenario as well.

Since this scenario is the closest to be used in current practice, it may be worth noting that current e-participatory systems have a specific user interface for the representation of the opinions in a debate. For example, Decidim (2016) allows participants to issue and relate comments in a forum-like format, and to value those comments through a like/dislike-sort-of system. Thus, in order to apply our aggregation functions, first we need to transform its representation into our DRF structure (see Subsection 3.1) and to map likes and dislikes into our valuation function (see Subsection 3.2). Later on, once all the opinions have been aggregated with the social choice properties from Table 3, the user interface can also be adapted to display the overall assessment over the target. This is particularly useful for both participants and the decision maker, as the aggregated valuation of the target τ provides a clear guidance for deciding whether the proposal in the target should be accepted. In fact, the system designers could even establish an automatic acceptance criteria by, for example, defining a threshold $\theta \geq 0$ so that τ gets accepted if $v_{F(P)}(\tau) \geq \theta$.

7.3 Constrained opinions: assuming coherent profiles

In the following, we assume that the opinion profile in a debate is constrained to be coherent at some degree (according to some value $\epsilon \in (0, 1)$), so that each of the opinions in the profile is always coherent. Recall that we consider that coherence occurs when the direct and indirect opinions are in line. Therefore, assuming coherence is expected to have a positive impact on aggregation functions that exploit indirect opinions to compute a collective opinion. Here we assess the gain.

Table 4 shows the properties satisfied by our aggregation functions when assuming coherence. The light green squares with check marks identify properties that are now satisfied, but were not (in Table 3) when not imposing coherence. Therefore, assuming coherence yields new positive results. More precisely, Table 4 shows that assuming coherence leads to the satisfaction of desirable unanimity properties for several of the functions. First, given the coherence assumption, the unanimity on the direct opinion drags the indirect opinion to become more similar to it, and therefore the Indirect function gains Weak unanimity. Now, since the Direct function also satisfies it, it follows that all α-Balanced functions now fulfil it too. Furthermore, thanks to the alignment that the coherence assumption brings between the direct and indirect opinions, the Direct function fulfils the Endorsed unanimity property. Therefore, having Endorsed unanimity fulfilled now by the Indirect and Direct functions, the aggregation functions in the Balanced family also fulfil it for any α.

Observe that unanimity and the coherence assumption work well together. Coherence on one sentence brings together its direct and indirect opinions, making it impossible for both to be far apart, and therefore allowing the Direct and Indirect functions to fulfil more unanimity properties.

Finally, the family of Recursive function now fulfils Endorsed unanimity, though, not for any α. Depending on the degree of coherence allowed in the opinion profile, i.e. the value of ϵ, the interval of α values allowing R_α to fulfil Endorsed unanimity will change. In this case, α has to be greater than $1/(2-\epsilon)$, representing the need to overcome the bad result obtained by the Recursive function with respect to the Endorsed unanimity property.

Desirable properties	D	I	R	α-B	α-**R**
Collective coherence ☆	✗	✗	✓	✗	$\alpha < \epsilon/2$
Exhaustive domain	✓	✓	✓	✓	✓
Anonymity	✓	✓	✓	✓	✓
Sided Unanimity	✓	✗	✗	✗	✗
Weak Unanimity	✓	✓	✗	✓	$\alpha > 1/2$
Endorsed Unanimity	✓	✓	✗	✓	$\alpha > \frac{1}{2-\epsilon}$
Familiar monotonicity	✓	✓	✗	✓	✗

Table 4: Social choice properties fulfilled when assuming coherent opinions. The light colours highlight those properties that are fulfilled in addition to those in Table 3.

7.4 Constrained opinions: assuming consensus on acceptance degrees and coherent profiles

In what follows we assume both previous constraints on the opinion profiles: coherence on the opinions and consensus on acceptances degrees. First, consensus on acceptance degrees on relationships represents an even more simplified debate than that in previous sections where participants only provide their opinions on sentences. Second, the coherence assumed on opinions aligns direct and indirect opinions. Overall, both assumptions yield major benefits in terms of the satisfaction of desired social choice properties, as we discuss next.

Table 5 shows the gain in fulfilment of desirable properties with respect to Table 4. Now, all our aggregation functions can guarantee ϵ-coherent aggregated opinions. This major improvement is because the consensus on acceptance degrees forbids the participants to value a relationship as 0, which is key to ensure collective coherence for the Direct and Indirect functions when the opinion profiles are coherent. We assume that, for each relationship, at least one agent has valued it other than 0, because otherwise it would be as if the relationship did not exist, and this forces all the participants to have a positive value too.

7.5 Summary

From the analysis for each debate scenario above, we can draw the following general observations:

Desirable properties	D	I	R	α-B	α-**R**
Collective coherence ☆	✓	✓	✓	✓	✓
Exhaustive domain	✓	✓	✓	✓	✓
Anonymity	✓	✓	✓	✓	✓
Sided Unanimity	✓	✗	✗	✗	✗
Weak Unanimity	✓	✓	✗	✓	$\alpha > 1/2$
Endorsed Unanimity	✓	✓	✗	✓	$\alpha > \frac{1}{2-\epsilon}$
Familiar monotonicity	✓	✓	✗	✓	✗

Table 5: Social choice properties fulfilled when assuming coherent opinions and consensus on acceptance degrees. The light colours highlight those properties that are fulfilled in addition to those in Table 4.

- The aggregation functions of the recursive family achieve collective coherence provided that they place little weight on direct opinions (or opinions are coherent and there is consensus on acceptance degrees).
- Coherence in opinion profiles favours unanimity (specifically, WU and EU), though in different ways. I and α-Balanced are fully satisfied, while the family of recursive functions leans on the direct aggregation function to fulfil some unanimity properties with restrictions. As a result, the α-Recursive family only satisfy WU and EU under strong conditions on α, because the R function never satisfies them.
- Coherent opinion profiles are not enough for D, I, and α-Balanced functions to achieve collective coherence. They also require consensus on acceptance degrees. Recursive functions do not require such consensus (nor even coherent opinion profiles), and hence, they are *robust* to the divergence of opinions on the relations between statements in a debate.
- While the D, I, B_α functions manage to achieve familiar monotonicity in all scenarios, the aggregation functions in the recursive family cannot achieve this even when counting on coherent opinion profiles and consensus on acceptance degrees. This is because the aggregated opinion on descendants recursively depends on descendants down the relational reasoning model. Thus, changes of opinion on "grandchildren" or deeper statements can cause a change of opinion independently of any change of the direct opinion.

Based on these general observations above, it is the task of the decision maker to decide the aggregation operator to choose considering: (1) the features of the debate scenario at hand; and (2) the desirable properties to guarantee. As a rule of thumb, since in real-world debates we cannot assume individual rationality (coherence), we believe that recursive aggregation functions are the best choice to achieve collective rationality, though we would pay the price of losing some other valuable properties, in particular unanimity for values of α that promote a large use of the direct opinion. Otherwise, if we do not value the coherence of the collective output, or we can guarantee somehow that the opinions of participants are coherent and the participatory system at hand does not allow for divergence on acceptance degrees, then, the Direct function becomes the aggregation function of choice. Within such constrained settings, the Direct aggregation function fulfils almost every property considered, even all of them in the debate scenario in Section 7.4. We conclude that it seems a good trade-off to consider the Recursive family, which can behave as similar to the Direct or to the Recursive function as wanted, and set the value of α depending on the features and goals in hand.

8. Computational Analysis

The purpose of this section is twofold. First, given the opinion aggregation problem in Section 5.1, we explain the complexity of the different algorithms for computing a collective decision on its target. In particular, we provide an algorithm for computing the recursive aggregation function. Thereafter, in Section 8.2, we empirically analyse the use of that algorithm to solve real-world collective decision problems.

8.1 Computing Aggregation Functions

Algorithm 1 Compute recursive aggregation

function ComputeRecursiveAggregation($\langle \mathcal{S}, \mathcal{R}, T \rangle, (O_1, \dots, O_n)$)
  for each relationship $r \in \mathcal{R}$ **do** ▷ Compute averaged acceptances
    aggregated_acceptance[r]← average_acceptances($w_1(r), \dots, w_n(r)$)
  $\mathcal{H}(\langle \mathcal{S}, \mathcal{R}, T \rangle)$ ←DRF_to_B-hypergraph($\langle \mathcal{S}, \mathcal{R}, T \rangle$) ▷ Generate $\mathcal{H}$: B-hypergraph reprentation of DRF
  sorted_statements← reverse(topological_sorting($\mathcal{H}(\langle \mathcal{S}, \mathcal{R}, T \rangle)$)) ▷ Compute topological sorting of $\mathcal{H}$
  for s in sorted_statements **do** ▷ Compute aggregated valuations
    valuation[s] ← 0 ▷ To accumulate aggregated valuations over descendants
    normaliser[s] ← 0 ▷ To normalise aggregated valuations over descendants
    compute relationships $R(s)$ to descendants
    if $R(s) \neq \emptyset$ **then** ▷ if s has descendants
      for each relationship $r \in R(s)$ **do**
        s_r ← descendant from relationship r
        valuation[s] ← valuation[s] + aggregated_valuation[s_r] · aggregated_acceptance[r]
        normaliser[s] ← normaliser[s] + aggregated_acceptance[r]
      valuation[s] ← valuation[s] / normaliser[s]
    else ▷ s has no descendants
      valuation[s] ← average_valuations($v_1(s), \dots, v_n(s)$)
    aggregated_valuation[s] ← valuation[s]
  return aggregated_valuation,aggregated_acceptance

All of the aggregation functions proposed in Section 6 can be calculated by rather efficient algorithms. For example, the direct function calculates the average for all statements and relationships in a DRF $\langle \mathcal{S}, \mathcal{R}, T \rangle$ considering the direct opinions in an opinion profile $P = (O_1, \dots, O_n)$. Hence, its complexity is given by $\mathbf{O}((|\mathcal{R}| + |\mathcal{S}|) \times |P|)$, where $|\mathcal{R}|, |\mathcal{S}|$ are the number of relationships and statements, respectively; and $|P|$ is the number of opinions in an opinion profile. Computing the indirect and balanced functions can be done by calculating the aggregated acceptance of each relationship as an average and by calculating the aggregated valuation of each statement as the average of the estimation function, which in turn is an average of the indirect opinions for that statement. Hence, their complexity is given by $\mathbf{O}(|\mathcal{R}| \times |\mathcal{S}| \times |P|)$. The calculation of the recursive function can be done by calculating the aggregated acceptance of each relationship as an average and calculating the aggregated valuation of each statement by starting with statements with no descendants and using these results to calculate the aggregated valuation of the statements directly connected to them. Algorithm 1 contains the pseudocode for the recursive function. In particular, the algorithm starts by computing aggregated acceptances ($w_{R(P)}$) as a weighted average (lines 2-3), which has a complexity of $\mathbf{O}(|\mathcal{R}| \times |P|)$. Then, the algorithm computes aggregated valuations ($v_{R(P)}$) starting from the statements with no descendants. In order to do that, we first perform a topological sorting of the DRF. Starting from the statements without descendants, the algorithm computes aggregated valuations until reaching the statements in T (lines 5-18). As we will show below, the topological sorting of the DRF has a complexity of $\mathbf{O}(|\mathcal{R}| \times |\mathcal{S}|)$. Hence, the calculation of the aggregated valuations has a complexity of $\mathbf{O}(|\mathcal{S}| \times max(|\mathcal{R}|, |P|))$. In online debates the number of opinions is usually higher than the number of relationships, hence, the complexity of the calculation of aggregated valuations can be typically given by $\mathbf{O}(|\mathcal{S}| \times |P|)$ and the total complexity of the recursive functions is $\mathbf{O}(|\mathcal{R}| \times |P|)$.

Topological sorting of a DRF. To calculate the topological sorting of a DRF we take advantage of well known results from hypergraph theory:

1. First we transform the graph associated to the DRF into an acyclic B-hypergraph (as defined in (Gallo et al., 1993)), which is a directed hypergraph where the head of all hyperedges has only one node. That transformation is performed in line 4 and the resulting hypergraph is denoted by $\mathcal{H}(\langle S, \mathcal{R}, T \rangle)$.

   Obtaining an acyclic B-hypergraph from a DRF is straightforward. In fact, the graph associated with a DRF is an acyclic B-hypergraph with the exception of the relationships that connect the very same statements. For instance, consider relationships r_1 and r_6 in figure 3 linking τ to s_1. Since in a hypergraph there cannot be two or more hyperedges over the very same nodes, we will only consider one single hyperedge. In our example, it suffices to consider either r_1 or r_6. We do not lose anything by doing this simplification because we want to obtain the topological sorting of a DRF, and hence considering one of the relationships connecting the very same statements is enough.

2. Performing the topological sorting over the B-hypergraph (line 5). Gallo et al. (1993) provide an algorithm to calculate the inverse topological sorting in an F-hypergraph with a complexity of $\mathbf{O}(|\mathcal{R}| \times |\mathcal{S}|)$. An F-hypergraph is a directed hypergraph where the tail of all hyperedges has only one node, and, hence, any given B-hypergraph can be transformed into a symmetric F-hypergraph by changing the direction of the hyperedges. Note that the inverse topological sorting of the symmetric F-hypergraph coincides with the topological sorting in the original B-hypergraph.

In (Ganzer-Ripoll et al., 2020) we provide a publicly-available implementation of algorithm 1 together with all the aggregation functions defined in this paper. Furthermore, we also provide guidelines on how to reproduce the experiments reported in Section 8.2 below.

8.2 Empirical Analysis

In what follows we empirically analyse the time required by our implementation of algorithm 1 to compute collective decisions. Based on the analysis above, we generated synthetic debates composed of DRFs and opinion profiles. On the one hand, we artificially generated

DRFs whose statements are the nodes of a directed acyclic B-hypergraph and whose hyperedges represent the relationships between statements. We chose the number of statements in our synthetic DRFs from $\{100, 150, 200\}$ to represent small, medium and large scenarios. Regarding the relationships between statements, we considered two parameters:

- *Density of relationships.* Given a relationship $r = (\Sigma, s)$, we say that the number of statements in Σ is the density of r. Since each relationship is represented as a hyperedge in a B-hypergraph, the density of relationships amounts to the size of the tails of hyperedges in the hypergraph. The (average) density of relationships in our artificial DRFs took values from $\{1, 2, 3\}$. We set the density value to 1 to generate DRFs for which there is a one-to-one connection between statements, and so each DRF is in fact a DAG. As to the other two density values (2 and 3), they allow us to generate DRFs where each relationship has two statements connected to one statement, and three statements connected to one statement respectively.

- *Density of number of relationships*, namely the average number of statements to which each statement is connected to through relationships. This corresponds to the average out degree of each statement in the DRF. We chose values for this parameter within $\{1, 2.5, 5\}$ to generate DRFs with low, medium, and high density of relationships.

To finish generating a debate, we must generate opinion profiles. We generated profiles with number of opinions from $\{10^6, 3 \cdot 10^6, 5 \cdot 10^6\}$, to represent the largest known actual-world scenarios[8]. The values for valuations and acceptances were randomly generated within $[-1, 1]$ and $[0, 1]$ respectively.

All the computations of collective decisions for our artificially generated debates were performed on an Ubuntu 16.04 box with an Intel(R) Core(TM) i7-8700K CPU @ 3.70GHz, with 31GiB system memory, and 8th Gen Core Processor Host Bridge/DRAM R. Furthermore, our experiments only considered the recursive aggregation function (specified in algorithm 1) because it is the most computationally expensive of those introduced in Section 5.

We performed three types of analysis:

- *Sensitivity to number of participants.* Figure 9 shows that the time to compute collective decisions increases as the number of participants increases. The figure shows the results for a medium density of number of relationships and a low density of relationships (which is the most expensive case as we discuss below).

- *Sensitivity to density of number of relationships.* Figure 10 shows that the time to compute collective decisions increases as the density of number of relationships grows.

8. To the best of our knowledge, the Brexit discussion on UK (Petitions, 2019) constitutes the largest such discussion: news outlets reported when the number of supporters passed 2 million (BBC, 2019) and the numbers kept growing during the 6 month period that the discussion was open. By the time it closed, there were 6,103,056 participants. Contrasting numbers of participants can be found for other popular initiatives such as an environmental proposal in Parlement et Citoyens which had 51,493 votes (Parlement & Citoyens, 2015) and in the participatory budgeting process in Helsinki (City of Helsinki, 2019) with 54,246 registered people, which represents 10% of the city voters. Note that the Parlement et Citoyens and Helsinki debates are probably more representative of real online debates than the Brexit example, where participants were, in effect, just voting on a specific proposal.

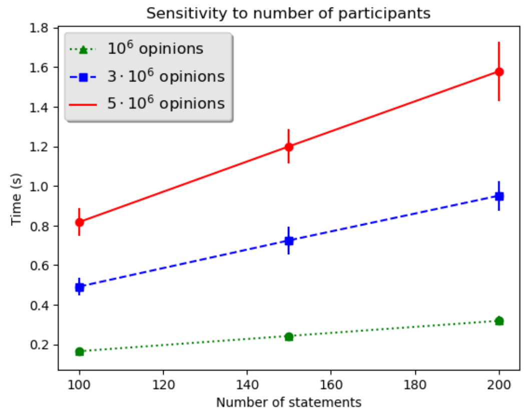

Figure 9: Sensitivity to number of participants in a DRF. Computational time required by the recursive aggregation function as the number of statements grows.

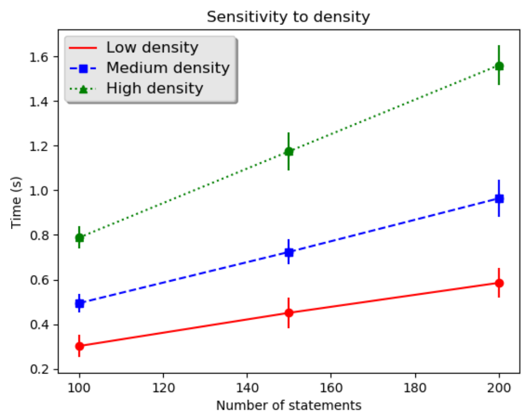

Figure 10: Sensitivity to density of relationships in a DRF. Density is considered in terms of the average out degree of statements: low (2.0), medium (5.0), high (10.0). We show the computational time required by the recursive aggregation function for $3 \cdot 10^6$ opinions as the number of statements grows.

Notice though most actual-world scenarios would lie between the low and medium cases, and hence it would take less than one second to solve even the largest debate.

- *Sensitivity to density of relationships.* Figure 11 shows that the time to compute collective decisions decreases as the density of relationships increases. The figure shows the results for $3 \cdot 10^6$ opinions and a medium density of number of relationships. This indicates that, surprisingly, our algorithm needs more time when relationships between sentences are one-to-one.

Overall, notice that computing collective decisions in all the artificially generated debates took less than 1.6 seconds. Therefore, we can conclude that our opinion aggregation functions can be employed to cope with large-scale debates in real time.

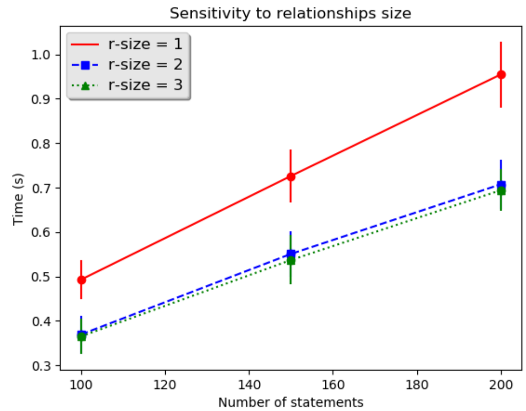

Figure 11: Sensitivity to the number of statements in the source of a relationship (size of the tail of hypergraph edges). We show the computational time required by the recursive aggregation function as the number of statements grows and as the size of hyperedges (tails) grows.

9. Related Work

In this section we review the work from the literature that is closest to the work in this paper. That includes work on approaches for computing the outcome of a set of arguments (Section 9.1), on approaches for analysing the behaviour of discussions from the standpoint of social choice theory (Section 9.2), and on systems for supporting online discussions (Section 9.3), acknowledging that the allocation of some work into a specific section is a somewhat arbitrary since it could validly be considered under more than one heading. Indeed we discuss different aspects of some of the most relevant work in more than one section. Note that the authors believe that their work is most closely related to work on social choice, and so Section 9.2 is the most substantial of the three.

9.1 Computational argumentation

We start with work on computing the outcome of a set of competing arguments, an area of research known as *computational argumentation*. Computational argumentation (Rahwan & Simari, 2009) has a lengthy history within artificial intelligence, going back at least as far as the work of Fox *et al.* (1980) and McGuire *et al.* (1981).

At the time of writing, work in computational argumentation is split into two broad groups. First, historically, is work which is concerned with the internal structure of arguments — what arguments are constructed from, and how this construction takes place — as well as how to compute the outcome. This line of work has reached its current endpoint with *structured argumentation* systems like logic-based argumentation (Besnard & Hunter, 2001), assumption-based argumentation (Dung et al., 2006) and structured argumentation systems such as ASPIC+ (Modgil & Prakken, 2013), and DeLP (García & Simari, 2004). Second is the line of work on *abstract argumentation*, begun by Dung (1995), which focuses much less on the internal structure of arguments, and instead is mainly concerned with the relationships between arguments. This has led to a large body of work expanding on

Dung's work (Dung, 1995), for example, work by Baroni and Giacomin (2009), Modgil and Caminada (2009), and Vreeswijk (1997).

In Dung's original work, (Dung, 1995), the focus is solely on "attack" relations, where arguments are in conflict, and the only way that one argument can express support for another is by attacking any attackers of that second argument. However, as in our sports centre example, in representing human discourse, it is common to find situations in which one argument is expressed which is in direct support of either the conclusion of another argument, or some element of that other argument. *Bipolar argumentation* provides formal models which capture this kind of support (Villata et al., 2012; Polberg & Hunter, 2018; Prakken, 2020; Lagasquie-Schiex, 2023). Key work developing an account of bipolar argumentation with abstract arguments was carried out by Amgoud *et al.* (2008), and Cayrol and Lagasquie-Schiex (2005a, 2005b). Since this initial work on bipolar argumentation can capture all of Dung-style argumentation, it can be viewed as a generalization, and subsequent work has generalized it further. For example, Brewka et al. (Brewka & Woltran, 2014; Brewka, Pührer, & Woltran, 2019) introduced the GRAPPA framework — itself based on abstract dialectical frameworks (Brewka & Woltran, 2010) which are another generalization of Dung-style systems — which can express a range of classes of argument that expand on just "support" and "attack"[9]. Most recently, Escañuela Gonzalez *et al.* (2021) have provided a general means of expressing information about support and attack that allows features such as numerical labels, and elements that allow the combination of strengths of arguments to propagate to those arguments that are supported or attacked. Both GRAPPA and the work of Escañuela Gonzalez *et al.* clearly have strong similarities with the mechanisms at the heart of our DRFs.

There is another way to, broadly, classify work on argumentation into two groups. One line of work, again exemplified by Dung (1995), focuses on argumentation as a mechanism for extracting consistent points of view from an inconsistent knowledge base. The other line of work deals with how arguments combine, or accrue, in favour of, or against some conclusion. This distinction cuts across the structured/abstract distinction with, for example the work of Baroni and Giacomin (2009) being concerned with consistency in abstract argumentation, and that of Modgil and Prakeen (2013) dealing with consistency in structured argumentation. On the other hand, the work of Besnard and Hunter (2001), Prakken (2005), and Verheij (1995) discusses accrual in structured argumentation, while that of Cayrol and Lagasquie-Schiex (2005a) looks at accrual in abstract argumentation.

All the work mentioned above uses argumentation as a mechanism for a single entity to come to a conclusion. However, as Sycara (1990), Walton and Krabbe (1995) and others have pointed out, argumentation is also a natural mechanism for multiple entities to use to reach consensus on some topic.

Our work connects to several of these themes in argumentation. First, since we are interested in aggregating the opinion from a number of participants, our work is clearly related to the use of argumentation in multiagent interaction. Second, the fact that individual steps in the participants' reasoning process are represented in our approach means that our work is connected to work on structured argumentation. (We would argue that it

9. The classes in one example in (Brewka & Woltran, 2010) are "strong support", "support", "attack" and "strong attack", which are very similar in denotation and in interpretation to the dictionaries in the work of Fox, for example that used by Krause *et al.* (1995).

is more abstract, since the relationships that connect statements are not restricted to be rules.) Third, our work connects with the idea of argumentation as a means of extracting a coherent view from a number of conflicting opinions. The fact that this coherent view takes into account the votes of participants also gives our work a fourth connection with the argumentation literature, in its relation to work on accrual. Fifth, the fact that we are concerned with opinions for and against target statements clearly connects our work to bipolar argumentation, and the kinds of labels that we use connects it most closely to the work of Escañuela Gonzalez *et al.* (2021) within work on bipolar argumentation. However, our work also differs from all of that listed so far in that the latter is ultimately focused on applying some form of the Dung semantics to resolve the conflicts between arguments whereas our work does not consider the Dung semantics at all. (This latter is part of the reason why we do not consider our work to be a contribution to the literature on argumentation.)

The voting aspect of DRFs also places our work in close relation to that of *social argumentation*, in which Dung-style argumentation has been adapted to capture elements of online debates. In particular, our work is related to the main line of work on social argumentation (Correia et al., 2014; Eğilmez et al., 2013; Leite & Martins, 2011), and previous work on collective argumentation[10] (Awad et al., 2017b; Caminada & Pigozzi, 2011; Ganzer-Ripoll et al., 2019). The work on collective argumentation, like our work here, is heavily influenced by social choice theory, and we will discuss more below when we outline the connections between our work and other work on social choice. We distinguish our work from that on social argumentation in that (i) our work deals with structured reasoning, rather than abstract arguments; (ii) it is not tied to notions of attack between arguments; and (iii) the underlying mechanism is analysed in terms of properties derived from social choice theory rather than from argumentation theory. Finally, one might view our work as being about the combination of different sets of arguments, one for each person who votes on the arguments or the relationships between them. From that perspective, our work also connects with that of Coste-Marquis *et al.*(2007), which takes as input different sets of arguments and relationships between them, and outputs consistent sets of arguments, thus "merging" the input sets. See the work of Bodanza *et al.* (2017) for a survey of work on this topic, and the work of Chen and Endriss (2019) for an excellent overview of developments over recent years.

9.2 Social choice theory

Given a set of alternatives and a set of agents who possess preference relations over the alternatives, social choice theory focuses on how to yield a collective choice that appropriately reflects the agents' individual preferences (Aziz et al., 2017). With this aim, social choice theory has extensively explored many ways of aggregating agents' individual preferences (Gaertner, 2009). Since there is a consensus in the literature on the desirable properties that a "fair" way of aggregating preferences should satisfy (e.g. no single agent can impose their view on the aggregate; if all agents agree, the aggregate must reflect the agreement; and so on (Gaertner, 2009)), aggregation functions can be characterised and compared in

10. Note that our work, and other work on combining the arguments from a group (Awad et al., 2017b; Caminada & Pigozzi, 2011; Ganzer-Ripoll et al., 2019) has little commonality with the "collective argumentation" studied by Bochman (2003), which is concerned with argumentation in which relationships exist between sets of arguments.

terms of the desirable properties they satisfy. Notice though that social choice theory counts on multiple *negative* results, namely impossibility results showing the incompatibility of certain sets of desirable properties such as Arrow's famous impossibility theorem (Arrow & Maskin, 2012).

Much of the work in social choice theory has placed little emphasis on the structure of the objects over which agents are expressing their preferences. However, there is a growing body of research that considers preferences on arguments in some form or other. This is the work that we referred to above as "collective argumentation". Here the foundational work was that of Rahwan and Thome (2010), later developed by Awad *et al.* (2017b), which considered the same problem that we tackle here: given a topic of discussion and a set of agents expressing their individual opinions about the statements made in the discussion, how can the agents reach a collectively rational decision? The way that this is tackled by Awad *et al* (2017b) is as a version of the "merging" problem mentioned above (Bodanza et al., 2017). That is, Awad *et al.* (2017b) consider each participant to have a set of arguments, and the relationships between them, and an opinion about which arguments are labelled as being acceptable and which are not. The problem they then solve is how to compute a set of labels for the arguments that reflect the opinions of all the participants such that the aggregation of opinions satisfies desirable social choice properties The same problem was also considered by Caminada and Pigozzi (2011), and more recently by us (Ganzer-Ripoll et al., 2019). As Awad *et al.* (Awad et al., 2017a) points out, their earlier work (Awad et al., 2017b) and that of Caminada and Pigozzi (2011) take different approaches, with Awad *et al.* (2017b) considering the opinions as votes, resolved by taking the plurality for individual arguments, while Caminada and Pigozzi (2011) offers a range of operators that yield a labelling which confirms to the constraints of argumentation semantics. The recent work by Chen and Endriss (2019) can be viewed as an extension of the line of research explored by Awad *et a.* (2017b) and Caminada and Pigozzi (2011). Like their forebears, Chen and Endriss (2019) propose methods for aggregating a collection of individual argumentation frameworks, corresponding to participants in a debate, into a single argumentation framework that appropriately reflects the views of the group as a whole. Like us, Chen and Endriss (2019) use techniques from social choice theory to investigate the properties of the aggregation rules. However, the aim is different. Chen and Endriss (2019) analyse aggregation rules in terms of their preservation of *semantic properties* of argumentation framework, while our focus in this paper is on *social choice properties* of the aggregation operators.

From the perspective of argumentation, the major difference between this line of work (Awad et al., 2017b; Caminada & Pigozzi, 2011; Chen & Endriss, 2019) and ours, is that we do not start from a set of opinions that are well-formed in an argumentation sense, that is a a "legal" labelling (Baroni et al., 2011). This is because we want to represent human opinions that may not be rational in an argumentation-theoretic sense (exactly as argued by Leite and Martins (2011)).

The next difference between the work in this paper and what has been done before is the richness of the representation. Here our work provides three main extensions. First, previous work on collective argumentation (Awad et al., 2017b; Caminada & Pigozzi, 2011; Chen & Endriss, 2019; Ganzer-Ripoll et al., 2019) all deals with abstract arguments. Here we deal with structured objects, and, as already mentioned, these are objects that are more general than the usual object studied in structured argumentation since we place no real constraints on the kind of reasoning captured by the relationships that hold between statements. Second, unlike previous work, we allow opinions to be expressed both about individual statements and the relations between them. This combines what has been studied in previous work on collective argumentation (Awad et al., 2017b; Caminada & Pigozzi, 2011; Chen & Endriss, 2019; Ganzer-Ripoll et al., 2019) and social argumentation (Leite & Martins, 2011), where opinions are expressed about individual arguments, but not the relations between them (these are assumed to be fixed), and the work of Eğilmez (Eğilmez et al., 2013), which extends previous work on social argumentation (Leite & Martins, 2011) to allow opinions to be expressed about the relationships between arguments. Third, we allow opinions to be real-valued. In this we move away from the labellings studied in previous work on collective argumentation (Awad et al., 2017b; Caminada & Pigozzi, 2011; Ganzer-Ripoll et al., 2019), and their grounding in argumentation semantics, and towards the kind of representation allowed by the work of Leite and Martins (2011)[11].

Finally, we neither assume independence between arguments as a fundamental postulate as is the case in (Awad et al., 2017b; Caminada & Pigozzi, 2011; Chen & Endriss, 2019), nor do we require the resulting aggregation to agree with an argumentation semantics. We will deal with these differences in turn.

Dropping the independence of arguments should come as no surprise since Awad *et al.* (Awad et al., 2017b) questions the necessity of assuming independence because of the dependencies between arguments that come already encoded in the form of relationships such as attack. Despite the importance of independence as a fundamental property in the judgement aggregation literature because of its theoretical value in proving strategy-proofness and strategic manipulation[12], we are not alone in regarding independence as too strong a property. This is because, together with mild further conditions, it implies dictatorship (Lang et al., 2016) and because it is also considered as not very plausible (Mongin, 2008). This explains why relaxing independence has been subject of much research (Lang et al., 2016). This paper goes beyond relaxing independence. Rather, in this paper we introduce several opinion aggregation functions that use the participants' opinions to compute a collective opinion while considering *dependencies* between statements. This is in line with our former work (Ganzer-Ripoll et al., 2019), but here we allow to express dependencies between multiple statements per relationship, while our previous work (Ganzer-Ripoll et al., 2019) (and all the work on collective argumentation which starts with abstract argumentation frameworks (Awad et al., 2017b; Caminada & Pigozzi, 2011; Chen & Endriss, 2019)) constrain relationships, and thus dependencies, to exist between pairs of arguments, hence limiting expressiveness.

11. As this discussion hopefully makes clear, the line of work on social argumentation of Leite and Martins (2011) and Eğilmez *et al.* (2013) — which also includes a study of efficient computation for the approach in Correia *et al.* (2014) — is close to what we have pursued in this paper. However, it remains work on an argumentation framework, where the focus is on abstract arguments, and on combining votes to establish the "winning" argument while we claim our work both deals with more general structures and captures a broader class of relations between those structures. However, our work cannot be a generalization of the work on social argumentation because that, like much work on argumentation frameworks, can handle cycles in the argument graph whereas our DRFs are explicitly acyclic.
12. If the independence criterion is not satisfied, then the function aggregating judgements is not immune to strategic manipulation (Dietrich & List, 2007).

Turning to the fact that the result of our aggregations do not match an standard argumentation semantics, we substitute the notion of “coherence of opinions” for the form of rationality embodied by those semantics. We do this for reasons that we have already touched on above with respect to the input opinions — we feel that insisting on an output that is rational in an argumentation theoretic sense is not necessarily realistic given that we start from opinions that are put forward by human participants who may not be consistent in their views. Instead of forcing the output of aggregation to be rational in an argumentation theoretic sense, we instead compute a measure which assesses how much concordance there is between related opinions, and assess our novel aggregation functions by the degree to which they can assure that their output ensures collective coherence.

From a pure social choice (not a combined argumentation and social choice) perspective, notice that it is common in the literature on judgement aggregation and preference aggregation to impose properties on the objects under aggregation in order that aggregation operators can guarantee desirable properties. For instance, in the case of distance-based aggregators, the *Kemeny* rule (Endriss & Moulin, 2016) only considers consistent judgement sets, and hence disregards those which are not, whereas premise-based aggregators (Endriss & Moulin, 2016) typically make assumptions on the agenda to guarantee consistency and completeness. Our work does not rely on this structuring of the target objects of the aggregation operators. Instead, we have introduced aggregation operators capable of guaranteeing collective coherence when opinions are unconstrained. This is motivated by the need for disregarding rationality when humans are involved in debates, since their opinions may eventually contain contradictions and inconsistencies.

Finally, we would like to highlight that there are further interesting connections between social choice and computational argumentation. This is the case for the work by (Maly & Wallner, 2021), where the authors draw on social choice theory to study sets of postulates for lifting operators in structured argumentation. Lifting operators capture the way that arguments are ranked based on the ranking of their respective sets of defeasible elements, and so relate to our work in their treatment of arguments of different strengths.

9.3 Tools for online discussion

As previously mentioned, public online discussion forums such as Decidim Barcelona (2016), Better Reykjavík (2021), New York City Participatory Budgeting (PBNYC, 2021), and Parlement et Citoyens (Purpoz, 2015) inspire our work. These particular sites allow participants to carry out a structured discussion by offering arguments for and against a (project or policy) proposal, and vote(/support) for arguments. Other participatory tools have also proliferated outside the context of public institutions. There, we find: `consider.it` (Con, 2022), Appgree (2021), Baoqu (2020) Loomio (Jackson & Kuehn, 2016; Loo, 2021), and Kialo (2021). These tools present varying levels of structure of the discussions they support. What distinguishes our work from all of these approaches is that we provide a much more expressive framework as we also allow the expression of opinions about relationships between statements and support relationships involving multiple statements, hence going beyond the limiting pairwise relationships supported by current e-participation systems.

There is also a long-standing line of work which develops tools to map the structure of arguments on some topic, for example (Carr, 2003; Reed & Rowe, 2004; Suthers et al., 1995;

Van Gelder, 2003). This line of work draws from a range of sources, nicely summarized by Shum (2003). The focus of this work is on drawing the relationships between arguments as a means of helping people understand the scenarios rather than to compute an overall outcome of the decision. Thus, our work could be applied in conjunction with any of the tools listed above to support structured discussion. In this line, it is worth noting that Klein's work on the Deliberatorium (Klein, 2012; Klein & Convertino, 2015) allows for the presentation of arguments and their interactions, and aggregates the opinions. However, unlike our work, it fails to check any social choice properties.

10. Conclusions and future work

Existing approaches to modelling human debates suffer from two drawbacks. Firstly, they assume that participants agree on the structure of a debate, that is the pieces of information that are relevant to particular issues and the relationships among them. This assumption is encoded in formal models by not allowing participants to express opinions about the structure. Secondly, existing approaches assume that participants' opinions are rational. This assumption is encoded by enforcing consistency on the individual opinions that are given, and the collective opinions that are computed from the individual ones. To address these limitations, we have proposed a new formal model. Our model allows participants in a debate to express agreement or disagreement with the relationships among statements, thus allowing disagreement with the structure of a debate to be captured. In addition, our model allows participants to express opinions that are inconsistent, defining a weaker notion of rationality to characterise *coherent* participant opinions, and allowing the collective opinions computed from this to be characterized as coherent as well.

Considering the degree of coherence of individual opinions and the level of consensus that participants have about the debate structure, we provide a formal analysis of the outcomes of different opinion aggregation functions in terms of social choice properties. Our analysis demonstrates that the recursive aggregation method is able to compute a coherent collective opinion even if individual opinions are incoherent and there is a lack of consensus on the debate structure. As we impose more restrictions on the coherence of individual opinions and consensus among participants on the debate structure, more aggregation methods also compute coherent collective opinions. We conclude our analysis with a computational evaluation in which we study the computational cost of aggregating collective opinions and experimentally demonstrate that collective opinions can be computed efficiently for real-sized debates.

One limitation of DRFs as a representation of debates is that DRFs are acyclic. We don't see this commitment to acyclic structures to be problematic. Cycles in reasoning are generally considered fallacious (Sinnott-Armstrong, 1999), and reasoning that involves cycles results in situations in which the premises are just as much in need of proof or evidence as the conclusion, and as a consequence, the argument fails to persuade. In this sense, we believe that the restriction to acyclic debates will only exclude the representation of "faulty" or fallacious debates. Indeed, the main participatory platforms such as Decidim do not allow for cycles in their debates, and so our approach arguably does not need the ability to handle cycles in order to be useful. Having said that, looking into the consequences of allowing DRFs to include cycles would be interesting as a way of connecting our work more closely to the literature on argumentation, and that could be an interesting line of future research.

We note that the restriction to being acyclic is in contrast to work on argumentation frameworks where some semantics allow for conclusions to be drawn from cyclic arguments, typically by either extracting conclusions that are unaffected by the contradiction that the cycle of attacks embodies, or identifying a set of extensions which identify the different consistent ways that the contradiction can be resolved. This contrast arises because what we are dealing with is not an argumentation framework (where nodes in the graph are entire arguments) but something closer to an argumentation system like ASPIC+ where the elements are individual steps in a larger structure. Unlike a system like ASPIC+, however, we do not "lift" elements into arguments and then establish the overall conclusions. Rather (from an APSIC+ viewpoint) we build a super-structure from all of the components of all of the arguments (that is what a DRF is, in ASPIC+ terms) and evaluate that directly. Studying the relationship between DRFs and structured argumentation systems like ASPIC+ is another potential line of future work.

There are other interesting possible directions in relating DRFs to argumentation. All of these are complicated by the lack of a direct map from a DRF into an argumentation framework, and so would be predicated on establishing some form of mapping. Given this, there are several directions that one might take the work, and a number of research questions that could be answered. First, there is the question of how the social choice mechanism we have suggested for DRFs maps to argumentation semantics. This would close the loop with work in the tradition of Awad *et al.* (2017b) which brings social choice mechanisms into argumentation. Second, there is the related question of how the results of a DRF relate to those of a social argumentation framework. One way to answer this question would be by adapting Maly and Wallner's (2021) approach to lifting preferences over the components of an argument (adapted to lift the votes for the elements of a DRF) to the preferences between arguments. Third, there is the question of how to map DRFs into argumentation frameworks with collective attacks. Again the mapping is not clear, because of the lack of a clear correspondance between a DRF and an argument, but given that collective arguments are (in argumentation framework terms) hypergraphs (Bikakis et al., 2021) this may be an easier mapping to establish than that to more conventional argumentation frameworks.

Another direction for theoretical exploration is to look more at the social choice properties of DRFs. In this paper we have focused on obtaining positive results regarding our aggregation rules, and we plan to take a complimentary approach in the future and investigate impossibility results by analysing whether some combinations of properties are not possible (in some scenarios).

Finally, we also plan to evaluate the practical impact of using the relational reasoning model in real online debates. For that, we expect to take advantage of our experience (Serramia et al., 2019; Lopez-Sanchez et al., 2021) in evaluating a previous model (Ganzer-Ripoll et al., 2019) using real-world data from Decidim Barcelona (2016). Moreover, and also thinking in practical terms, we plan to investigate methods to analyse the quality of a debate represented in the relational reasoning model. For example, we are interested in studying the use of systematic incoherence in participant opinions as a way to identify structural problems in a debate. Notice that analysing the quality of a debate is a subject of interest (e.g. (Gómez et al., 2008; Gonzalez-Bailon et al., 2010; Aragón, 2019)) when exploiting debates for deliberation (Friess & Eilders, 2015).

Acknowledgements

This work was supported by: the UK Engineering & Physical Sciences Research Council (EPSRC) under grants #EP/P010105/1 and #EP/P010105/2, by projects COREDEM (H2020-785907), CI-SUSTAIN (PID2019-104156GB-I00), and 22S01386-001 from Barcelona City Council through the Fund. Solidaritat UB, 2021 SGR 00313, 2021 SGR 00754, VAE TED2021-131295B-C31, funded by MCIN/AEI/10.13 039/501100011033 and NextGenerationEU/ PRTR, VAE TED2021-131295B-C32, funded by MCIN/AEI/10.13039/501100011033 and the European Union NextGenerationEU/PRTR, VALAWAI (Horizon Europe #101070930), AUTODEMO (SR21-00329) by Fundación La Caixa, Crowd4SDG (H2020-872944), Fairtrans (PID2021-124361OB-C33). Maite Lopez-Sanchez belongs to the WAI research group (University of Barcelona) associated unit to CSIC.

Jordi Ganzer was supported by a PhD studentship from King's College London.

Appendix A. Formal proofs and results

In the following, we prove all the formal results presented in Section 7 regarding the satisfaction of social choice properties by the opinion aggregation functions introduced in Section 6.The section is divided into four parts, one per debate scenario as analysed in Section 7.

1. Unconstrained opinion profiles;
2. Constrained opinion profiles: assuming consensus on acceptance degrees;
3. Constrained opinion profiles: assuming coherent profiles; and
4. Constrained opinion profile: assuming consensus on acceptance degrees and coherent profiles.

Furthermore, for each scenario, our results will be grouped by aggregation function according to the following order: Direct aggregation, Indirect aggregation, Recursive aggregation, Balanced family aggregation and Recursive family aggregation

A.1 Unconstrained opinion profiles

In this section we analyse the social choice properties fulfilled by the aggregation functions introduced in Section 6: assuming unconstrained opinions profiles (*any* opinion profile is deemed to be possible input for the aggregation functions). The results of this section are summarised in Table 3 in Section 7.1.

Proposition A.1. The aggregation function D satisfies the following properties:

(i) Exhaustive Domain and Coherent Domain;

(ii) Anonymity and Non-Dictatorship;

(iii) Familiar Monotonicity;

(iv) Sided Unanimity and Weak Unanimity.

And does not satisfy:

(vi) Collective coherence; and

(vii) Endorsed Unanimity.

Proof. (of Proposition A.1)

(i) Exhaustive Domain is straightforward and Collective Domain follows directly.

(ii) Anonymity and Non-Dictatorship. Let $P = (O_1, \dots, O_n)$ be an opinion profile over a DRF and σ a permutation over a set of agents $Ag = \{1, \dots, n\}$. We must show that D maintains the same collective opinion over the permuted opinion profile $P' = (O_{\sigma(1)}, \dots, O_{\sigma(n)})$, i.e. that $D(P) = D(P')$. This is the case because the next two equalities hold:

$$v_{D(P)}(s) = \frac{1}{n}\sum_{i=1}^{n} v_i(s) = \frac{1}{n}\sum_{i=1}^{n} v_{\sigma(i)}(s) = v_{D(P')}(a);$$

$$w_{D(P)}(r) = \frac{1}{n}\sum_{i=1}^{n} w_i(r) = \frac{1}{n}\sum_{i=1}^{n} w_{\sigma(i)}(r) = w_{D(P')}(r).$$

Therefore, Anonymity holds and Non-Dictatorship follows from it as we discussed in Section 5.2.

(iii) Familiar Monotonicity. Let s be a statement and P and P' two opinion profiles satisfying the assumptions in the definition of the property in Section 9.2, i.e. $P = (O_1, \ldots, O_n)$ and $P = (O'_1, \ldots, O'_n)$ are such that $v_i(s) \leq v'_i(s)$ for every agent $i \in \{1, \ldots, n\}$. Then, from the definition of D, we obtain the aggregated valuation on s is:

$$v_{D(P)}(s) = \frac{1}{n}\sum_{i=1}^{n} v_i(s) \leq \frac{1}{n}\sum_{i=1}^{n} v'_i(s) = v_{D(P')}(s)$$

Therefore, D satisfies Familiar Monotonicity.

(iv) Sided Unanimity and Weak Unanimity. Let $P = (O_1, \ldots, O_n)$ be an opinion profile over a DRF and a statement $s \in \mathcal{S}$ such that $v_i(s) > 0$ for every agent in $Ag = \{1, \ldots, n\}$. The aggregated opinion on s is:

$$v_{D(P)}(s) = \frac{1}{n}\sum_{i=1}^{n} v_i(s) > 0$$

The case when $v_i(s) < 0$ for every agent in Ag can be proven similarly.

Hence, Sided Unanimity is fulfilled by D. According to Proposition 5.1 Weak Unanimity follows from Sided Unanimity.

(v) Collective Coherence. To prove that it does not hold, it suffices to find a DRF and an opinion profile for which there is no collective coherence. Thus, consider the example depicted below in figure 12.

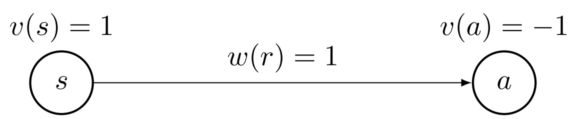

Figure 12

If we check coherence for statement s, we obtain that:

$$|v_{D(P)}(s) - e_{D(P)}(s)| = v(s) - v(a) = 2 > \epsilon.$$

for any $\epsilon \in (0, 1)$, and hence collective coherence does not hold for this profile.

(vi) Endorsed Unanimity. Using the opinion profile depicted in figure 12, we observe that even with full negative support on s (i.e. $v(a) = -1$), the result of the aggregation is the opposite ($v_{D(P)}(s) = 1$). Therefore, this opinion profile also serves as a counterexample to prove that D does not satisfy Endorsed Unanimity.

□

Proposition A.2. The aggregation function I satisfies the following properties:

(i) Exhaustive Domain and Coherent Domain;

(ii) Anonymity and Non-Dictatorship;

(iii) Endorsed Unanimity; and

(iv) Familiar Monotonicity.

And does not satisfy:

(v) Collective coherence;

(vi) Sided Unanimity and Weak Unanimity.

Proof. (of Proposition A.2)

(i) Exhaustive and Coherent domain are straighforward.

(ii) Anonymity and Non-Dictadorship. Let $P = (O_1, \ldots, O_n)$ be an opinion profile over a DRF and σ a permutation over the agent in $Ag = \{1, \ldots, n\}$. We must show that I maintains the same collective opinion over the permuted opinion profile $P' = (O_{\sigma(1)}, \ldots, O_{\sigma(n)})$, i.e. that $I(P) = I(P')$.

For any $i \in \{1, \ldots, n\}$ there is only one $j \in \{1, \ldots, n\}$ such that $\sigma(j) = i$, and hence in terms of expectation functions we know that $e_i = e_{\sigma(j)}$. Using that, we can show that $I(P) = I(P')$ as follows:

$$v_{I(P)}(s) = \frac{1}{n}\sum_{i=1}^{n} e_i(s) = \frac{1}{n}\sum_{i=1}^{n} e_{\sigma(i)}(s) = v_{I(P')}(a);$$

$$w_{I(P)}(r) = \frac{1}{n}\sum_{i=1}^{n} w_i(r) = \frac{1}{n}\sum_{i=1}^{n} w_{\sigma(i)}(r) = w_{I(P')}(r)$$

.

(iii) Endorsed Unanimity. Let s be a sentence and $P = (O_1, \ldots, O_n)$ an opinion profile satisfying that $v_i(s_r) = 1$ for any agent i and descendant $s_r \in D(s)$ of sentence s. Since the expectation over s is:

$$e_i(s) = \frac{1}{\sum_{r \in R^+(s)} w_i(r)} \sum_{r \in R^+(s)} w_i(r) v_i(s_r) = \frac{1}{\sum_{r \in R^+(s)} w_i(r)} \sum_{r \in R^+(s)} w_i(r) = 1,$$

then the aggregated value for s is:

$$v_{I(P)}(s) = \frac{1}{n}\sum_{i} e_i(s) = 1.$$

Analogously, if we assume that $v_i(s_r) = -1$ for any agent i and descendant $s_r \in D(s)$ of sentence s, we would obtain that $v_{I(P)}(s) = -1$. Since $v_{I(P)}(s) > 0$ when there is full positive support (and $v_{I(P)}(s) < 0$ for negative support), Endorsed Unanimity holds.

(iv) Familiar Monotonicity. It is straightforward to see that the output of the Indirect aggregation function, which uses an expectation function, depends only on the values on descendants and their relationships. So, it is clear that a different opinion profile maintaining the same values for descendants and their relationships will not change the output of the function.

(v) Collective Coherence. To prove that it does not hold, it suffices to find a DRF and an opinion profile for which there is no collective coherence. Thus, consider the example depicted below in figure 13. Here $v_{I(P)}(s) = 1$ and $v_{I(P)}(a) = -1 = v_{I(P)}(b)$. Now, if we check coherence for statement s, we obtain that $|v_{I(P)}(s) - e_{I(P)}(s)| = 2 > \epsilon$ for any $\epsilon \in (0, 1)$, and hence collective coherence does not hold for this profile.

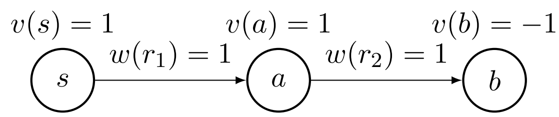

Figure 13

(vi) Sided Unanimity and Weak Unanimity. Next we build a DRF and an opinion profile for which Weak Unanimity does not hold despite satisfying the assumptions. Consider the example in figure 14 with opinion profile $P = (O = (v, w))$. Although $v(s) = 1$, $v_{I(P)}(s) = -1$ instead of greater than 0, and hence I does not satisfy Weak unanimity. As discussed in Section 5.2, an aggregation function satisfying Sided Unanimity also satisfies Weak Unanimity. Thus, since Weak Unanimity does not hold, neither does Sided Unanimity.

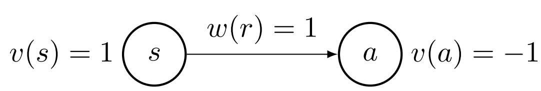

Figure 14

□

Proposition A.3. The aggregation function R satisfies the following properties:

(i) Collective Coherence;

(ii) Exhaustive Domain and Coherent Domain;

(iii) Anonymity and Non-Dictatorship;

And does not satisfy:

(iv) Sided Unanimity and Weak Unanimity;

(v) Endorsed Unanimity;

(vi) Familiar Monotonicity.

Proof. (i) Collective Coherence. Since $v_{R(P)} = e_{R(P)}$, the collective opinion for R is exactly the result of applying the estimation function, and hence collective coherence follows because $|v_{R(P)}(s) - e_{R(P)}(s)| = 0 < \epsilon$ for any $\epsilon \in (0, 1)$ and any sentence $s \in \mathcal{S}$.

(ii) Exhaustive Domain and Coherent Domain. Straightforward.

(iii) Anonymity and Non-Dictatorship. Let $P = (O_1, \ldots, O_n)$ be an opinion profile over a *DRF* and σ a permutation over the agents in $Ag = \{1, \ldots, n\}$. We must show that R maintains the same collective opinion over the permuted opinion profile $P' = (O_{\sigma(1)}, \ldots, O_{\sigma(n)})$, i.e. that $R(P) = R(P')$.

We consider first the sentences $s \in \mathcal{S}$ with no descendants such that $R^+(s) = \emptyset$. Since these have no descendants, R computes the collective opinion on them using D. As shown by Proposition A.1, since D satisfies Anonymity, it will also hold for R when considering sentences with no descendants. Thus, since these sentences, which are used at the beginning of the recursive process run by R, will not change through permutations, the collective opinion over any sentence will be the same after permutations. Therefore, Anonymity holds for R, and from this Non-Dictatorship.

(iv) Weak and Sided Unanimity. The example of *DRF* depicted in figure 15 with opinion profile $P = (O = (v, w))$ will be enough to show that R does not satisfy Weak unanimity. Although $v(s) = 1$, and hence the assumptions for Weak Unanimity hold, $v_{R(P)}(s) = -1$ influenced by the valuation on b. Since $v_{R(P)}(s)$ is not positive, Weak unanimity does not hold for R, and consequently Sided unanimity.

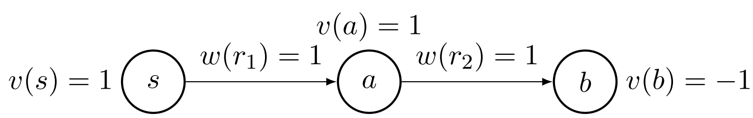

Figure 15

(v) Endorsed Unanimity. Consider again the opinion profile depicted in figure 15. Clearly, since $v(a) = 1$, s has full positive support, but $v_{R(P)}(s) = -1$. Since $v_{R(P)}(s)$ is not positive, Endorsed Unanimity does not hold.

(vi) Familiar Monotonicity. We build a *DRF* and two opinion profiles for which Familiar Monotonicity does not hold despite satisfying the assumptions. Consider the two opinion profiles $P = (O = (v, w))$ and $P' = (O' = (v', w'))$ depicted in figures 16 and 17 respectively. Considering s, these two profiles satisfy the assumptions of Familiar Monotonicity: $v(s) \leq v'(s)$ and the values on the indirect opinion are the same. However, P and P' differ on the value on b: $v(b) = 1$ and $v'(b) = -1$. This leads to a change of value on the aggregated value on s. Thus, $v_{R(P)}(s) \not\leq v_{R(P')}(s)$, and R fails at satisfying Familiar Monotonicity.

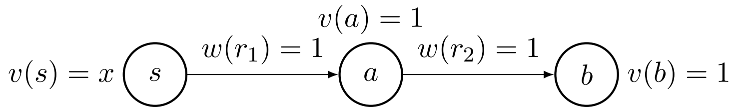

Figure 16

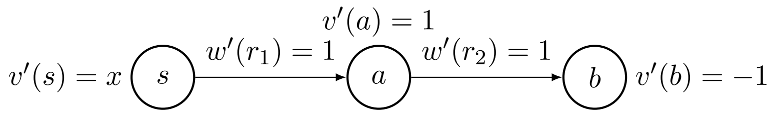

Figure 17

□

Next, we provide the proofs for the analysis of the families of α-balanced aggregation functions $\{B_\alpha\}_{\alpha\in(0,1)}$ and α-recursive aggregation functions $\{R_\alpha\}_{\alpha\in(0,1)}$. Before that, we first introduce some general lemmas that will be useful to build the proofs of the propositions for both families. To ease notation, these general lemmas that follow consider two generic aggregation functions F and G, as well as a generic aggregation function $H = \alpha F + (1-\alpha)G$ instead of $v_H(P) = \alpha v_F(P) + (1-\alpha)v_G(P)$. Hereafter, the following lemmas establish the social properties fulfilled by H.

Lemma A.1. Let F and G be two opinion aggregation functions satisfying Exhaustive domain. For any $\alpha \in (0,1)$, aggregation function $H = \alpha F + (1-\alpha)G$ also satisfies Exhaustive domain.

Proof. Straightforward from the fact that both F and G satisfy Exhaustive domain. □

Lemma A.2. Let F and G two opinion aggregation functions satisfying Anonymity over domain $\mathcal{D}$. For any $\alpha \in (0,1)$, aggregation function $H = \alpha F + (1-\alpha)G$ also satisfies Anonymity over domain $\mathcal{D}$.

Proof. For any given opinion profile P and its permuted profile P', if $F(P) = F(P')$ and $G(P) = G(P')$, then it follows that $H(P) = H(P')$. □

Lemma A.3. Let F and G two opinion aggregation functions satisfying Familiar Monotonicity over domain $\mathcal{D}$. For any $\alpha \in (0,1)$, aggregation function $H = \alpha F + (1-\alpha)G$ also satisfies Familiar Monotonicity on domain $\mathcal{D}$.

Proof. Let $P = (O_1 = (v_1, w_1), \ldots, O_n = (v_n, w_n))$ and $P' = (O'_1 = (v'_1, w'_1), \ldots, O'_n = (v'_n, w'_n))$ be a profile satisfying the assumptions of familiar montonicity for a statement s, i.e. $v_i(s) \leq v'_i(s)$ for any i and for any $r \in R^+(s)$ then $w_i(r) = w'_i(r)$ and $v_i(s_r) = v'_i(s_r)$. Since F and G satisfy familiar monotonicity, then $v_{F(P)}(s) \leq v_{F(P')}(s)$ and $v_{G(P)}(s) \leq v_{G(P')}(s)$. Thus, since $H = \alpha F + (1-\alpha)G$, it follows directly that $v_{H(P)}(s) \leq v_{H(P')}(s)$, so familiar montonicity holds for H. □

Lemma A.4. Let F and G two opinion aggregation functions satisfying Sided unanimity on domain $\mathcal{D}$. For any $\alpha \in (0,1)$, aggregation function $H = \alpha F + (1-\alpha)G$ also satisfies Sided unanimity on $\mathcal{D}$.

Proof. Since Sided unanimity holds for F and G, we know that given any opinion profile P of agents $\{1, \dots, n\}$, i.e. if for any $i \in \{1, \dots, n\}$ $v_i(s) > 0$ then $v_F(s) > 0$ and $v_G(s) > 0$, and since $v_H = \alpha v_F + (1-\alpha)v_G$, it also follows that $v_H(s) > 0$. Likewise for the negative case, so Sided Unanimity holds for H. □

Lemma A.5. Let F and G two opinion aggregation functions satisfying Weak unanimity on the domain $\mathcal{D}$. For any $\alpha \in (0,1)$, aggregation function $H = \alpha F + (1-\alpha)G$ also satisfies Weak unanimity over domain $\mathcal{D}$.

Proof. Since Weak unanimity holds for F and G, we know that given any opinion profile P of agents $\{1, \dots, n\}$, for any $i \in \{1, \dots, n\}$, if $v_i(s) = 1$, then $v_F(s) > 0$ and $v_G(s) > 0$. Since $v_H = \alpha v_F + (1-\alpha)v_G$, it also follows that $v_H(s) > 0$, and hence Weak Unanimity holds for H. Analogously for the negative case. □

Lemma A.6. Let F and G two opinion aggregation functions satisfying Endorsed unanimity on domain $\mathcal{D}$. For any $\alpha \in (0,1)$, aggregation function $H = \alpha F + (1-\alpha)G$ also satisfies Endorsed unanimity on $\mathcal{D}$.

Proof. Since Endorsed unanimity holds for F and G, we know that given any opinion profile P of agents $\{1, \dots, n\}$, for any $i \in \{1, \dots, n\}$ and descendant $s_r \in D(s)$ of sentence s, if $v_i(s_r) > 1$, then $v_F(s) > 0$ and $v_G(s) > 0$. Since $v_H = \alpha v_F + (1-\alpha)v_G$, it also follows that $v_H(s) > 0$, and hence Endorsed Unanimity holds for H. Analogously for the full negative support case. □

We are now ready to prove the results for α-balanced aggregation functions in $\{B_\alpha\}_{\alpha \in (0,1)}$.

Proposition A.4. The family of α-balanced aggregation functions $\{B_\alpha\}_{\alpha \in (0,1)}$ satisfies the following properties:

(i) Exhaustive Domain and Coherent Domain;

(ii) Anonymity and Non-Dictatorship;

(iii) Weak Unanimity for $\alpha \in (\frac{1}{2}, 1)$;

(iv) Endorsed Unanimity for $\alpha \in (0, \frac{1}{2})$;

(v) Familiar Monotonicity;

and does not satisfy:

(vi) Collective coherence;

(vii) Sided Unanimity.

Proof. (i) Exhaustive Domain and Coherent Domain follow from propositions A.1 and A.2, and from Lemma A.1.

(ii) Anonymity and Non-Dictatorship follow from propositions A.1 and A.2, and from Lemma A.2.

(iii) Weak Unanimity. Let $P = (O_1, \dots, O_n)$ be an opinion profile over a DRF for the agents in $Ag = \{1, \dots, n\}$, and $s \in \mathcal{S}$ a sentence such that $v_i(s) = 1$ for any i. By Proposition A.1, we know that Unanimity holds for the Direct aggregation function, and hence $v_{D(P)}(s) = \frac{1}{n}\sum_{i \in Ag} v_i(s) = 1$. Note that in that case the balanced function is given by: $v_{B_\alpha(P)}(s) = \alpha + v_{I(P)}(s) - \alpha v_{I(P)}(s)$. Note $v_{I(P)}(s) \in [-1, 1]$ and that $v_{B_\alpha(P)}(s)$ takes its minimum value when $v_{I(P)}(s) = -1$. The DRF and an opinion profile depicted in figure 18 exemplifies this case.

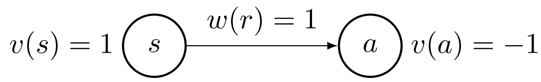

Figure 18

Since $v_{D(P)}(s) = 1$ and $v_{I(P)}(s) = -1$, $v_{B_\alpha(P)}(s) = \alpha - (1 - \alpha) = 2\alpha - 1$. Thus, by choosing any value of α such that $\alpha \in (\frac{1}{2}, 1)$, we ensure that $v_{B_\alpha(P)}(s) > 0$, and Weak unanimity holds. The proof is analogous for the negative case of Weak unanimity.

(iv) Endorsed Unanimity. Let s be a sentence and $P = (O_1, \dots, O_n)$ an opinion profile satisfying that $v_i(s_r) = -1$ for any agent i and descendant $s_r \in D(s)$ of sentence s. In other words, s has full negative support. It follows that $v_{I(P)}(s) = -1$. Likewise for our proof for Weak unanimity above, we consider the case where the function takes its maximum value, which would occur when $v_{D(P)}(s) = 1$. Figure 18 depicts a DRF and single-opinion profile illustrating this case. Since $v_{D(P)}(s) = 1$ and $v_{I(P)}(s) = -1$, $v_{B_\alpha(P)}(s) = \alpha - (1 - \alpha) = 2\alpha - 1$. Thus, by choosing any value of α such that $\alpha \in (0, \frac{1}{2})$, we ensure that $v_{B_\alpha(P)}(s) < 0$, and Endorsed unanimity holds. The proof is analogous for the positive case (full positive support) of Endorsed unanimity.

(v) Familiar Monotonicity follows from propositions A.1 and A.2, and from Lemma A.3.

(vi) Collective coherence.To prove that it does not hold, it suffices to find a DRF and an opinion profile for which there is no collective coherence. Thus, consider the DRF with one-opinion profile depicted below in figure 19. Computing the aggregations for the Direct and Indirect functions, we have that $v_{D(P)}(s) = 1$, $v_{D(P)}(a) = 0$, and, $v_{I(P)}(s) = 0$ and $v_{I(P)}(a) = -1$. Therefore, $v_{B_\alpha(P)}(s) = \alpha$ and $v_{B_\alpha(P)}(a) = (-1)(1 - \alpha)$. And hence, the coherence at sentence s is: $|v_{B_\alpha(P)}(s) - e_{B_\alpha(P)}(s)| = |v_{B_\alpha(P)}(s) - v_{B_\alpha(P)}(a)| = 1$. Thus, we conclude that, for any $\epsilon \in (0, 1)$, ϵ-coherence cannot be satisfied regardless of the value of α. Therefore, B_α does not satisfy ϵ-coherence.

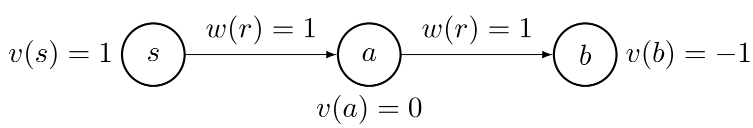

Figure 19

(vii) Sided Unanimity. We will show that for any $\alpha \in (0,1)$ we can find a DRF and an opinion profile for which Sided unanimity does not hold. Consider then the DRF with single-opinion profile in figure 20, where $x \in (0,1)$ is such that $0 < x < \frac{1-\alpha}{\alpha}$. Since $v(s) = x > 0$, the assumptions for Sided unanimity hold at sentence s. Now, since $v_{D(P)}(s) = x$ and $v_{I(P)}(s) = -1$, it follows that $v_{B_\alpha(P)}(s) = \alpha x + (1-\alpha)(-1) = \alpha x + \alpha - 1 < \alpha\frac{1-\alpha}{\alpha} + \alpha - 1 = 0$. Since $v_{B_\alpha(P)} \not> 0$, the proof goes analogously for the negative case of Sided unanimity.

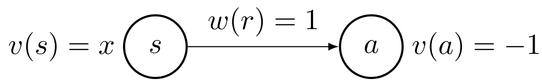

Figure 20

□

Proposition A.5. The family of α-recursive aggregation functions $\{R_\alpha\}_{\alpha\in(0,1)}$ satisfies the following properties:

(i) Collective Coherence for $\alpha < \frac{\epsilon}{2}$;

(ii) Exhaustive Domain and Coherent Domain;

(iii) Anonymity and Non-Dictatorship;

(iv) Weak Unanimity for $\alpha > \frac{1}{2}$;

and does not satisfy:

(v) Sided unanimity;

(vi) Endorsed Unanimity;

(vii) Familiar Monotonicity.

Proof. (i) Collective Coherence. Given $\epsilon > 0$ and a DRF, we must prove that $|v_{R_\alpha(P)}(s) - e_{R_\alpha(P)}(s)| < \epsilon$ for any sentence $s \in \mathcal{S}$. First, we develop the difference between valuation and estimation for the collective opinion:

$$
\begin{aligned}
v_{R_\alpha(P)}(s) - e_{R_\alpha(P)}(s) &= v_{R_\alpha(P)}(s) - \frac{\sum_{r \in R^+(s)} w_{R_\alpha(P)}(r) v_{R_\alpha(P)}(s_r)}{\sum_{r \in R^+(s)} w_{R_\alpha(P)}(r)} \\
&= [\alpha v_{D(P)}(s) + (1-\alpha) v_{R(P)}(s)] - \frac{\sum_{r \in R^+(s)} w_{D(P)}(r) \big[\alpha v_{D(P)}(s_r) + (1-\alpha) v_{R(P)}(s_r)\big]}{\sum_{r \in R^+(s)} w_{R_\alpha(P)}(r)} \\
&= \alpha \big[v_{D(P)}(s) - \frac{\sum_{r \in R^+(s)} w_{R_\alpha(P)}(r) v_{D(P)}(s_r)}{\sum_{r \in R^+(s)} w_{D(P)}(r)}\big] + \\
&+ (1-\alpha) \big[v_{R(P)}(s) - \frac{\sum_{r \in R^+(s)} w_{R(P)}(r) v_{R(P)}(s_r)}{\sum_{r \in R^+(s)} w_{R(P)}(r)}\big] \\
&= \alpha \big[v_{D(P)}(s) - e_{D(P)}(s)\big] + (1-\alpha) \big[v_{R(P)}(s) - e_{R(P)}(s)\big] \\
&= \alpha (v_{D(P)}(s) - e_{D(P)}(s)).
\end{aligned}
$$

Notice that we get rid of $v_{R(P)}(s) - e_{R(P)}(s)$ because is zero. Now the coherence of R_α directly depends on the coherence of the direct aggregation function D and α. Figure 21 depicts a DRF with a single-opinion profile representing a worst-case scenario for D because $v_{D(P)}(s) - e_{D(P)}(s) = 2$. Considering the coherence of R_α, we have that $|v_{R_\alpha(P)}(s) - e_{R_\alpha(P)}(s)| = \alpha |v_{D(P)}(s) - e_{D(P)}(s)| \leq 2\alpha$ for any profile P. Therefore, we must ensure that $\alpha < \frac{\epsilon}{2}$ so that $|v_{R_\alpha(P)}(s) - e_{R_\alpha(P)}(s)| < \epsilon$ holds for any profile of the domain, and hence Collective coherence holds for R_α.

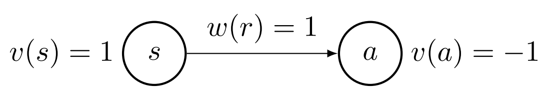

Figure 21

(ii) Exhaustive Domain and Coherent Domain directly follow from propositions A.1 and A.3 and Lemma A.1.

(iii) Anonymity and Non-Dictatorship follow directly from propositions A.1 and A.3 and Lemma A.2.

(iv) Weak unanimity. To prove Weak unanimity we can resort to the proof built to prove Weak unanimity for B_α in Proposition A.4. We simply have to substitute B_α for R_α.

(v) Sided Unanimity. Consider the DRF and the single-opinion profile depicted in figure 22, where $x \in (0,1)$ is such that $0 < x < \frac{1-\alpha}{\alpha}$. Since $x > 0$, the assumption for Sided Unanimity holds at s. However, $v_{R_\alpha(P)}(s)$ is not positive, since $v_{R_\alpha(P)}(s) = x\alpha - 1 + \alpha < \frac{1-\alpha}{\alpha}\alpha - 1 + \alpha = 0$, and hence Sided unanimity does not hold.

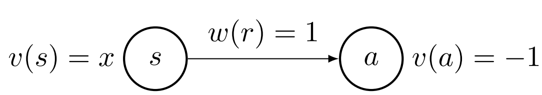

Figure 22

(vi) Endorsed Unanimity. Next we build a DRF and an opinion profile for which Endorsed Unanimity does not hold. Consider the DRF and the opinion profile P depicted in figure 23. The assumptions for Endorsed unanimity hold at s because s has full negative support. However, $v_{R_\alpha(P)}(s)$ is not negative: since $v_{D(P)}(s) = 1$ and $v_{R(P)}(s) = 1$, we obtain that $v_{R_\alpha(P)}(s) = 1$ for any $\alpha \in (0, 1)$. Therefore, Endorsed Unanimity does not hold.

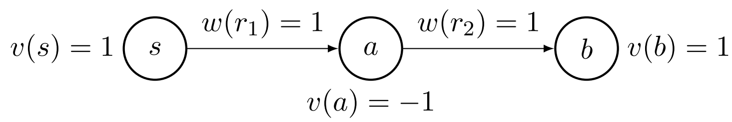

Figure 23

(vii) Familiar Monotonicity. We build a DRF and an opinion profile for which Familiar Monotonicity does not hold despite satisfying the assumptions. Consider the DRF and single-opinion profile P in figure 23 together with another single-opinion profile P' in figure 24. Clearly, the assumptions of Familiar Monotonicity are fulfilled at s because $v_i(s) \leq v'_i(s)$ and the descendant of s has the same value. However, we will show that $v_{R_\alpha(P)}(s) \leq v_{R_\alpha(P')}(s)$ is not true. For both profiles we have that $v_{D(P)}(s) = v_{D(P')}(s) = 1$, and, $v_{R(P)}(s) = 1$ and $v_{R(P')}(s) = 1 - x$. Thus, for any $\alpha \in (0, 1)$: $v_{R_\alpha(P)}(s) = \alpha + (1 - \alpha) = 1$ and $v_{R_\alpha(P')}(s) = \alpha + (1 - \alpha)(1 - x) = 1 - x(1 - \alpha) < 1$ for any $x \in (0, 1)$. Therefore, $v_{R_\alpha(P)}(s) > v_{R_\alpha(P')}(s)$ and Familiar Monotonicity is not satisfied.

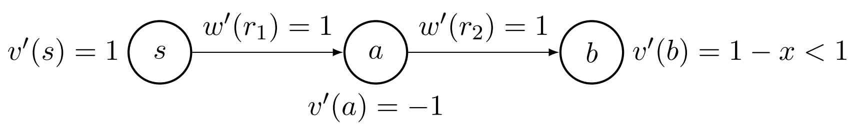

Figure 24

$\square$

A.2 Constrained opinion profiles: assuming consensus on acceptance degrees

This section relates to Section 7.2, where we assume that opinion profiles share consensus on their acceptance degrees on relationships, i.e. for each relationship $r \in \mathcal{R}$ of a DRF all the agents agree on their acceptance degrees: $w_i(r) = w_j(r)\ \forall i, j \in Ag$.

In previous Section A.1, each proof and counterexample used to demonstrate that an aggregation function does or does not satisfy a property uses opinion profiles composed by one single agent. Thus, those proofs serve as well in this section when assuming consensus on acceptance degrees. For this reason, adding this assumption does not change any of the properties fulfilled by the aggregation functions in the general case (Table 3), and therefore there are no further desirable properties gained in this scenario with respect to the more general scenario thoroughly analysed in Section A.1.

A.3 Constrained opinion profiles: assuming coherent profiles

This section corresponds to the results displayed in Table 4 in Section 7.3. We prove the results regarding the social choice properties satisfied by the aggregation functions introduced in Section 6 when assuming the domain of the aggregation functions to be ϵ-coherent for some $\epsilon \in (0,1)$. This means that we consider that our aggregation functions take in coherent opinion profiles.

Since in the previous section many properties have been proven for the general case, we will not need to prove them again for this more restrictive scenario. For each opinion aggregation function, we will prove only those results regarding social choice properties that change by the addition of the coherence assumption and disprove again, this time for coherent domains, those properties which are yet not satisfied.

Proposition A.6. D over a coherent domain satisfies:

(i) Endorsed Unanimity;

and does not satisfy:

(ii) Collective coherence.

Proof. (i) Endorsed Unanimity. Let s a statement in a DRF and let $R^+(s)$ be the set of relationships r from s to its descendants s_r. Let P be an ϵ-coherent profile for $\epsilon \in (0,1)$ with full positive support on s, i.e. $v_i(s_r) = 1$ for any i and descendant $s_r \in D(s)$. Then:

$$e_i(s) = \frac{1}{\sum_{r\in R^+(s)} w_i(r)} \sum_{r\in R^+(s)} v_i(s_r) w_i(r) = \frac{1}{\sum_{r\in R^+(s)} w_i(r)} \sum_{r\in R^+(s)} w_i(r) = 1$$

By the ϵ-coherence of P we have that:

$$|v_i(s) - e_i(s)| < \epsilon \Longrightarrow v_i(s) > e_i(s) - \epsilon = 1 - \epsilon.$$

Therefore, for any $\epsilon \in (0,1)$ we can ensure that $v_i(s) > 0$ for any i and the conditions for Sided Unanimity hold. Now, since D satisfies Sided Unanimity (by Proposition A.1), we obtain that $v_D(s) > 0$, and hence D fulfils Endorsed Unanimity.

(ii) Collective coherence. Consider the DRF and the δ-coherent opinion profile P depicted in figure 25 and any $\delta \in (0,1)$. We will show that the collective opinion yield by the direct function for this example is never ϵ-coherent for any $\epsilon \in (0,1)$.

Clearly, this profile is δ-coherent for any $\delta > 0$. Computing the direct function at s we obtain that: $v_{D(P)}(s) = -1$, $v_{D(P)}(a) = 0$ and $w_{D(P)}(r) = \frac{1}{2}$. Now, if we check collective coherence at s, we see that: $|v_{D(P)}(s) - e_{D(P)}(s)| = |-1-0| = 1 > \epsilon$. Thus, since 1 is larger than any ϵ value that we take in $(0,1)$, D does not satisfy ϵ-Collective coherence.

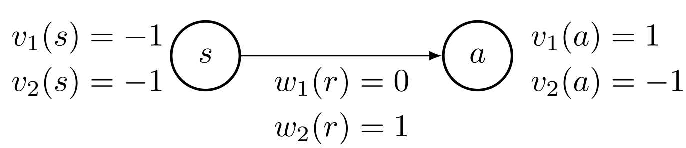

Figure 25

□

Proposition A.7. *I* over a coherent domain satisfies:

(i) Weak Unanimity;

and does not satisfy:

(ii) Sided Unanimity.

(iii) Collective Coherence.

Proof. (i) Weak Unanimity. Consider a DRF with a statement $s \in \mathcal{S}$ and $P = (O_1 = (v_1, w_1), \ldots, O_n = (v_n, w_n))$ an opinion profile such that $v_i(s) = 1$ for every i. Hence, the conditions for Weak unanimity hold. If the profile P is ϵ-coherent, where $\epsilon \in (0,1)$, then we can conclude that for any i: $1 - \epsilon < e_i(s) < 1 + \epsilon$, being $1 - \epsilon > 0$ for any $\epsilon \in (0,1)$. Now, computing v_I at s we get:

$$v_{I(P)}(s) = \frac{1}{n} \sum_i e_i(s) > \frac{1}{n} \sum_i 1 - \epsilon > 0$$

Since $v_{I(P)}(s) > 0$, Weak unanimity holds. The proof for the negative case of Weak unanimity is analogous.

(ii) Sided unanimity. Consider the DRF and one-opinion profile depicted in figure 26 such that $\epsilon \in (0,1)$ and x, y such that $0 < x < y < \epsilon$. The assumptions of Sided unanimity are fulfilled at s. We check that the opinion in the profile is ϵ-coherent because $|v(s) - e(s)| = |x - y + \epsilon| < \epsilon$. However, $v_{I(P)}(s) = y - \epsilon < 0$, instead of positive, and hence Sided Unanimity is not satisfied.

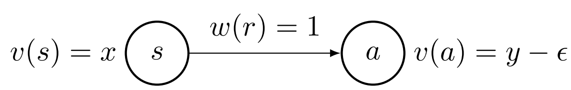

Figure 26

(iii) Collective Coherence. To prove that this property does not hold, it suffices to find a *DRF* and an opinion profile for which there is no ϵ-Collective coherence. Consider the DRF and opinion profile P in figure 27. Clearly, opinions O_1 and O_2 of P are δ-coherent for any $\delta > 0$. Now, we compute the indirect function for all statements: $v_{I(P)}(s) = \frac{-1}{2}$, $v_{I(P)}(a) = \frac{1}{2}, v_{I(P)}(b) = 0$, and, $w_{I(P)}(r_1) = \frac{1}{2} = w_{I(P)}(r_2)$. If we check coherence at s we see that:

$$|v_{I(P)}(s) - e_{I(P)}(s)| = |v_{I(P)}(s) - v_{I(P)}(a)| = |\frac{-1}{2} - \frac{1}{2}| = 1 > \epsilon.$$

Thus, since 1 is larger that any ϵ value that we take in $(0,1)$, I does not satisfy ϵ-Collective coherence.

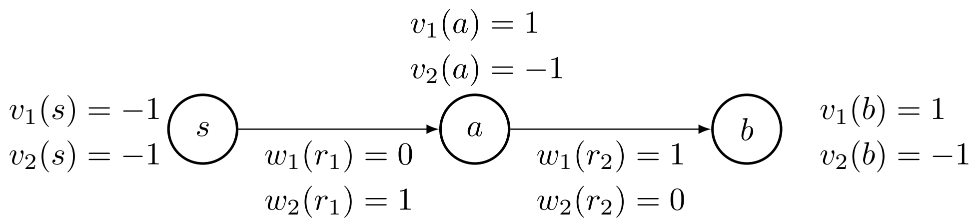

Figure 27

□

Proposition A.8. R over a coherent domain fulfils the following properties:

(i) Weak Unanimity and SidedUnanimity;

(ii) Endorsed Unanimity;

(iii) Familiar Monotonicity.

Proof. (i) Weak Unanimity, and Sided unanimity. To prove that neither of these properties hold, it suffices to build a DRF and opinion profile to show that Weak Unanimity does not hold. This is sufficient because Weak unanimity is a weaker variant of Sided unanimity. Indeed, Proposition 5.1 tells us that Sided unanimity will not hold if Weak Unanimity does not. Consider the DRF and opinion profile $P = ((v, w))$ in figure 28 such that $w(r) = 1$ for any relationship $r \in \mathcal{R}$, $\epsilon \in (0, 1)$ and $\delta \in (0, \epsilon)$ and $m \in \mathbb{N}$ so that $m\delta \geq 1 > (m-1)\delta$.

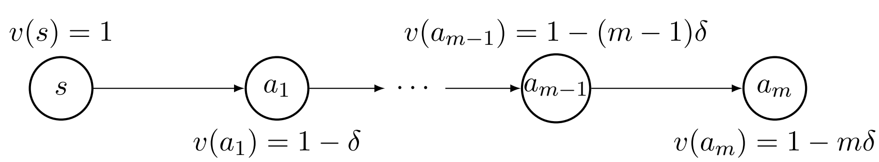

Figure 28

Clearly, the outcome of the recursive function at each sentence is obtained from the value of the recursive function at the previous sentence, i.e.:

$$v_{R(P)}(a_m) = v_{R(P)}(a_{m-1}) = ... = v_{R(P)}(a_1) = v_{R(P)}(s),$$

which actually is the value $v(a_m) = 1 - m\delta \leq 0$. So, this is an ϵ-coherent opinion profile fulfilling the assumptions of Weak unanimity at sentence s because $v(s) = 1$. However, the value of the recursive function at s is negative. Therefore, R does not fulfil Weak unanimity.

(ii) Endorsed Unanimity. We build a DRF and opinion profile to show that Endorsed unanimity does not hold from the example in the previous proof. Figure 29 shows our example, which extends the one in figure 28 with an additional sentence a. Since

$v(a_i) - v(a_{i-1}) = \delta$, likewise in the proof above, we have an ϵ-coherent opinion profile. Since $v(s) = 1$ the assumption for Endorsed unanimity at a is satisfied, but since $v(a) = 1 - m\delta \leq 0$, Endorsed unanimity does not hold.

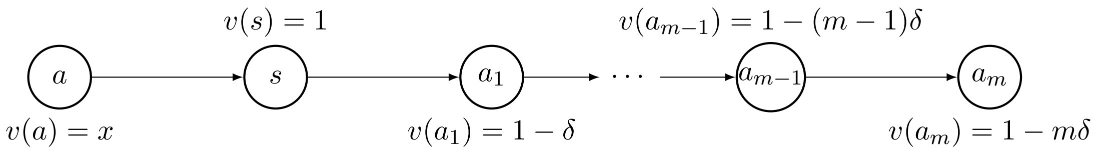

Figure 29

(iii) Familiar Monotonicity. Consider the opinion profiles P and P' over the same DRF depicted in figures 30 and 31 respectively. Since $v(s) = v(a) = v(b) = 1$, P is ϵ-coherent. By setting $0 < x < \epsilon$, we also obtain that P' is ϵ-coherent. Therefore, both P and P' are ϵ-coherent and the assumptions for familiar monotonicity hold at s. However, since $1 = v_{R(P)}(s) > v_{R(P')}(s) = 1 - x$, Familiar monotonicity cannot hold.

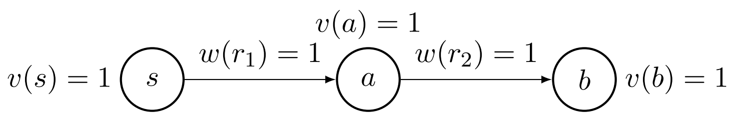

Figure 30

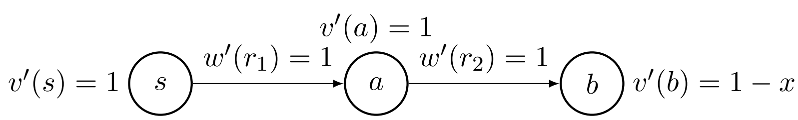

Figure 31

□

Proposition A.9. The family $\{B_\alpha\}_{\alpha \in (0,1)}$ over a coherent domain satisfies:

(i) Weak Unanimity; and

(ii) Endorsed Unanimity

and does not satisfy:

(iii) Sided Unanimity;

(iv) Collective Coherence.

Proof. (i) Weak Unanimity follows from propositions A.1, A.7 and Lemma A.5.

(ii) Endorsed Unanimity follows from propositions A.2, A.6 and Lemma A.6.

(iii) Sided Unanimity. It suffices to build a DRF and an opinion profile for which Sided unanimity does not hold for any values of α and ϵ, where $\alpha, \epsilon \in (0,1)$.

Consider the set $A = \{(x,y) \in (0,1) \quad | \quad 0 < y < \epsilon \text{ and } 0 < x < y - \alpha y\}$. We check first, that this set is actually not empty. For $\alpha \in (0,1)$, $y - \alpha y > 0$, thus $y > \alpha y > 0$. So for $y \in (0,\epsilon)$, there are $x \in (0,1)$ satisfying $x < y - \alpha y$.

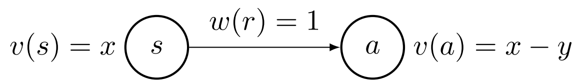

Figure 32

Now, we consider the DRF and opinion profile depicted in figure 32 where x and y are values from A, namely $(x,y) \in A$. Since $|v(s) - e(s)| = |x - (x-y)| = |y| < \epsilon$, the opinion profile in the figure is ϵ-coherent, and satisfies the assumptions for Sided Unanimity at s because $v(s) = x > 0$. However,

$$\begin{aligned} v_{B_\alpha(P)}(s) &= \alpha v_{D(P)}(s) + (1-\alpha) v_{I(P)}(s) \\ &= \alpha x + (1-\alpha)(x-y) \\ &= x - y + \alpha y < 0 \end{aligned}$$

since $(x,y) \in A$. So, clearly this example shows that Sided Unanimity does not hold for the family B_α in an ϵ-coherent profile.

(iv) First, we show that the ϵ-coherence condition for B_α depends on the functions employed in its definition, namely on D and I:

$$\begin{aligned} |v_{B_\alpha(P)}(s) - e_{B_\alpha(P)}(s)| &= \left| v_{B_\alpha(P)}(s) - \frac{\sum_{r \in R^+(s)} \big(\alpha v_{D(P)}({}_r s) + (1-\alpha) v_{I(P)}(s_r)\big) w_{D(P)}(r)}{\sum_{r \in (R^+(s)} w_{D(P)}(r)} \right| \\ &= \Big| \big(\alpha(v_{D(P)}(s) + (1-\alpha)(v_{I(P)}(s)\big) - \big(\alpha e_{D(P)}(s)) + (1-\alpha) e_{I(P)}(s))\big) \Big| \\ &= \Big| \alpha\big(v_{D(P)}(s) - e_{D(P)}(s)\big) + (1-\alpha)\big(v_{I(P)}(s) - e_{I(P)}(s)\big) \Big| \end{aligned}$$

Thus, since D and I do not satisfy ϵ-collective coherence for any δ-coherent profile (by propositions A.6 and A.7 respectively), neither will B_α satisfy the property for any $\alpha \in (0,1)$. Indeed, consider for instance the DRF and δ-coherent opinion profile P, with any $\delta \in (0,1)$, in figure 27 as employed in Proposition A.7. If we compute ϵ-collective coherence for B_α at sentence s we obtain that:

$$|v_{B_\alpha(P)}(s) - e_{B_\alpha(P)}(s)| = \left| \alpha(-1-0) + (1-\alpha)(-\frac{1}{2} - \frac{1}{2}) \right| = |-1| = 1 > \epsilon$$

for any $\epsilon \in (0,1)$. So, B_α does not fulfill ϵ-coherence for any $\alpha \in (0,1)$.

$\square$

Proposition A.10. Let $\epsilon \in (0, 1)$ such that the domain of R_α is an ϵ-coherent domain, then the family $\{R_\alpha\}_{\alpha\in(0,1)}$ satisfies:

(i) Weak Unanimity for $\alpha \in (\frac{1}{2}, 1)$, and hence independently of ϵ;

(ii) Endorsed unanimity for $\alpha \in (\frac{1}{2-\epsilon}, 1)$; and

(iii) ϵ-Collective coherence for $\alpha \in (0, \frac{\epsilon}{2})$.

and does not satisfy:

(iv) Sided Unanimity; and

(v) Familiar Monotonicity.

Proof. (i) Weak Unanimity. Consider a DRF with sentences $\mathcal{S}$, P an opinion profile over the DRF and $s \in \mathcal{S}$ a sentence such that $v_i(s) = 1$ for any agent i. We know that

$$v_{D(P)}(s) = \frac{1}{n} \sum_{i\in Ag} v_i(s) = 1.$$

and hence $v_{D(P)}(s) = \frac{1}{n} \sum_{i\in Ag} v_i(s) = 1$. Note that in that case $R_\alpha(s)$ function is given by $\alpha + v_{R(P)}(s) - \alpha v_{R(P)}(s)$. Note $v_{R(P)}(s) \in [-1, 1]$ and $v_{R_\alpha}(s)$ takes its minimum value when $v_{R(P)}(s) = -1$. The DRF and profile depicted in figure 15 above shows that, in fact, this scenario exists with $v_{D(P)}(s) = 1$ and $v_{R(P)}(s) = -1$, and hence $v_{R_\alpha(P)}(s) = \alpha + (1 - \alpha)(-1) = 2\alpha - 1$. To fulfil Weak unanimity, we need that $v_{R_\alpha(P)}(s) > 0$ holds, but we also know that $v_{R_\alpha(P)}(s) \geq 2\alpha - 1$. Therefore, we can guarantee Weak Unanimity for those R_α functions for which $\alpha \in (\frac{1}{2}, 1)$. Hence, Weak Unanimity holds for the family of functions $\{R_\alpha\}_{\alpha\in(\frac{1}{2},1)}$. The proof for the negative case of Weak Unanimity goes analogously.

(ii) Endorsed Unanimity. To prove this property we will build a customised DRF and opinion profile to demonstrate the worst case that we can find when fulfilling the assumptions of Endorsed unanimity.

Consider a DRF and let $P = (O_1 = (v_1, w_1), \dots, O_n = (v_n, w_n))$ be an ϵ-coherent profile with full positive support on statement $s \in \mathcal{S}$.

First, we consider the worst case where $v_{R(P)}(s) = -1$ can be achieved when s has full positive support. Figure 33 depicts a DRF and an opinion profile illustrating this situation.

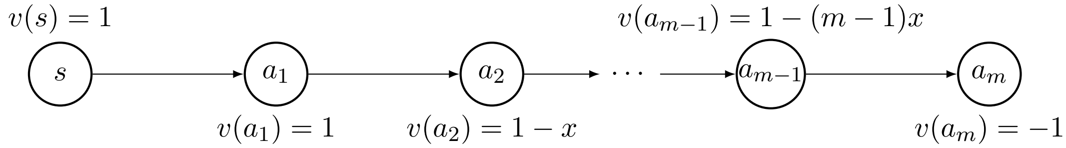

Figure 33

By choosing $0 < x < \epsilon$ and $m \in \mathbb{N}$ such that $mx > 2 \geq (m-1)x$, this example shows an ϵ-coherent profile where $v(a_1) = 1$ (full positive support) and $v_R(s) = -1$. Next, we move to the general setting considered by the proof, an opinion profile with n agents, knowing that the worst case for this property is possible. Since $v_i(s_r) = 1$ for any descendant $s_r \in D(s)$ and any agent i, the estimation function on s will be $e_i(s) = 1$ for any agent. Therefore, from the coherence condition at s we conclude that

$$1 - \epsilon < v_i(s) < 1 + \epsilon.$$

Consider that for every i, $v_i(s) = 1 - \delta_i$ such that $0 \leq \delta_i < \epsilon$. This clearly satisfies the previous inequality. Now we take $\delta = \max_i\{\delta_1, \ldots, \delta_n\}$ to create a new opinion profile $P' = (O'_1 = (v'_1, w_1), \ldots, O'_n = (v'_n, w_n))$ such that $v_i(a) = v'_i(a)$ for $a \in \mathcal{S} \setminus \{s\}$ and $v'_i(s) = 1 - \delta$ for any i. Then, since D fulfils Monotonicity, we know that $v_{D(P')}(s) \leq v_{D(P)}(s)$. Furthermore, from the definition of D, it follows that $v_{D(P')}(s) = \frac{1}{n}\sum_{i=1}^{n} v'_i(s) = \frac{1}{n}\sum_{i=1}^{n} 1 - \delta = 1 - \delta$. And, from the example in figure 33 we know that for any ϵ-coherent opinion profile $v_{R(P)}(s) \geq -1$. Therefore,

$$\begin{aligned} v_{R_\alpha(P)}(s) &= \alpha v_{D(P)}(s) + (1-\alpha) v_{R(P)}(s) \\ &\geq \alpha v_{D(P')}(s) + (1-\alpha)(-1) \\ &= (1-\delta)\alpha - (1-\alpha) = (2-\delta)\alpha - 1 \end{aligned}$$

So, if we choose $\alpha \in (0,1)$ so that $(2-\delta)\alpha - 1 > 0$, the R_α aggregation function will satisfy Endorsed Unanimity. Since $\delta < \epsilon$, by choosing a value for α such that $\alpha \geq \frac{1}{2-\epsilon} > \frac{1}{2-\delta}$, then we can ensure that the aggregation function R_α satisfies Endorsed Unanimity. Therefore, given an ϵ value $\epsilon \in (0,1)$, the family of aggregation functions $\{R_\alpha\}_{\alpha \in (\frac{1}{2-\epsilon}, 1)}$ satisfies Endorsed Unanimity.

(iii) As seen before in Proposition A.3, the collective coherence of R_α entirely depends on the collective coherence of D, i.e.:

$$|v_{R_\alpha(P)}(s) - e_{R_\alpha(P)}(s)| = \alpha |v_{D(P)}(s) - e_{D(P)}(s)|$$

Thus, finding the worst case scenario for D will give us the condition on α that ensures that R_α satisfies ϵ-collective coherence for any ϵ. Next, we consider an example showing that $|v_{D(P)}(s) - e_{D(P)}(s)|$ can be as close to 2 as wanted depending on the number of agents.

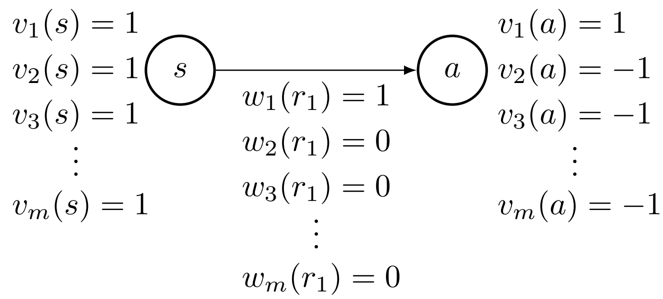

Figure 34

Let P be the δ-coherent opinion profile over a DRF depicted in figure 34, for any $\delta \in (0,1)$. For any $i > 1$: $v_i(s) = 1$, $v_i(a) = -1$ and $w_i(r_1) = 0$; whereas $v_1(s) = 1$, $v_1(a) = 1$ and $w_1(r_1) = 1$. We check the condition for collective coherence at s to find that:

$$|v_{R_\alpha(P)}(s) - e_{R_\alpha(P)}(s)| = \alpha(1 + \frac{m-2}{m}) < \epsilon$$

if $\alpha < \frac{\epsilon}{1+\frac{m-2}{m}}$. Thus, by choosing a value for α such that $\alpha < \frac{\epsilon}{2} < \frac{\epsilon}{1+\frac{m-2}{m}}$ we ensure that R_α satisfies ϵ-coherence for the worst case. Therefore, for any δ-coherent opinion profile, $\delta \in (0,1)$, choosing $\alpha \in (0, \frac{\epsilon}{2})$ will ensure that R_α satisfies ϵ-collective coherence for any $\epsilon \in (0,1)$.

(iv) Sided Unanimity. Consider the DRF and opinion profile P depicted in figure 35, where: $x \in (0,1)$ is such that $0 < x < \frac{1-\alpha}{\alpha}$, $0 < \delta < \epsilon$; $m \in \mathbb{N}$ satisfies $(m-1)\delta \leq 1+ < m\delta$; and for any $r \in \mathcal{R}$, $w(r) = 1$.

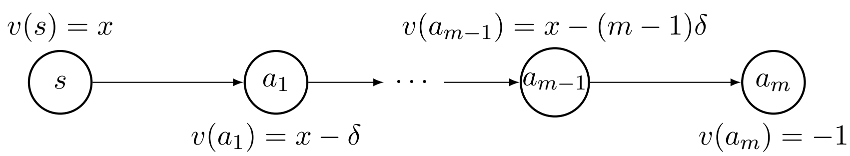

Figure 35

Clearly, P is an ϵ-coherent because $|v(a_m) - e(a_m)| = 0$, and for any $i < m$, $|v(a_i) - e(a_i)| = v(a_i) - v(a_{i+1}) = \delta < \epsilon$, and $|v(s) - e(s)| = x - x + \delta < \epsilon$. Furthermore, P satisfies the assumptions of Sided unanimity at s since $v(s) = x > 0$. It is straightforward to see that $v_{D(P)}(s) = x$ and $v_{R(P)}(s) = v_{R(P)}(a_m) = -1$. Hence, $v_{R\alpha(P)}(s) = x\alpha + (1-\alpha)(-1) = x\alpha + \alpha - 1$. But since $x < \frac{1-\alpha}{\alpha}$, we conclude that $v_{R\alpha(P)}(s) < \frac{1-\alpha}{\alpha}\alpha + \alpha - 1 = 0$. We can proceed analogously for the negative case of Sided Unanimity.

(v) Familiar monotonicity. The counterexample employed in Proposition A.8 to show that Familiar monotonicity does not hold for R serves here as well to prove that R_α does not satisfy Familiar monotonicity for any $\alpha \in (0,1)$. From opinion profiles P and P' depicted in figures 30 and 31 respectively, we extract that $v_{D(P)}(s) = v_{D(P')}(s) = 1$ and $1 = v_{R(P)}(s) > v_{R(P')}(s) = 1-x$. Therefore, it follows that $v_{R_\alpha(P)}(s) > v_{R_\alpha(P')}(s)$, hence proving that Familiar monotonicity does not hold.

□

A.4 Constrained opinion profiles: assuming Consensus on acceptance degrees and coherent profiles

Next, we show the results regarding our fourth, and last, scenario. We now assume that opinion profiles are both ϵ-coherent, for some $\epsilon \in (0,1)$, and agree on their acceptance degrees over relationships. The results that follow are summarised in Table 5 in Section 7.4.

Likewise in previous sections, next we only prove per aggregation function those properties that either were partially satisfied or not satisfied at all in previous scenarios, but do hold in this new scenario. We do not prove those properties for which the proofs in the previous sections serves as well for this scenario.

Proposition A.11. Let be a *DRF* and an opinion profile $P = (O_1, \ldots, O_n)$. For any $s \in \mathcal{S}$, assume that for each $r \in R^+(s)$ $w_i(r) = \lambda_r \in (0,1]$ for any i, then:

(i) For any $\epsilon \in (0,1)$, if $0 < \delta \leq \epsilon$ and the domain $\mathcal{D}$ is δ-coherent then $D(P)$ is ϵ-coherent, so satisfies ϵ-Collective coherence.

(ii) For any $\epsilon \in (0,1)$, if $0 < \delta \leq \epsilon$ and the domain $\mathcal{D}$ is δ-coherent then $I(P)$ is ϵ-coherent, so satisfies ϵ-Collective coherence.

(iii) For any $\epsilon \in (0,1)$, if $0 < \delta \leq \epsilon$ and the domain $\mathcal{D}$ is δ-coherent then $B_\alpha(P)$ is ϵ-coherent for any $\alpha \in (0,1)$, so satisfies ϵ-Collective coherence.

(iv) For any $\epsilon \in (0,1)$, if $0 < \delta \leq \epsilon$ and the domain $\mathcal{D}$ is δ-coherent then $R_\alpha(P)$ is ϵ-coherent for any $\alpha \in (0,1)$, so satisfies ϵ-Collective coherence.

Proof. (i) Collective Coherence of D. Let $s \in \mathcal{S}$. We assume that for any i, $|v_i(s) - e_i(s)| < \delta \leq \epsilon$. Next we calculate the coherence condition for D at sentence s:

$$
\begin{aligned}
|v_{D(P)}(s) - e_{D(P)}(s)| &= \left| \frac{1}{n} \sum_i v_i(s) - \frac{1}{\sum_{r \in R^+(s)} w_{D(P)}(r)} \sum_{r \in R^+(s)} w_{D(P)}(r) v_{D(P)}(s_r) \right| \\
&= \left| \frac{1}{n} \sum_i v_i(s) - \frac{1}{\sum_{r \in R^+(s)} \lambda_r} \sum_{r \in R^+(s)} \lambda_r \big(\frac{1}{n} \sum_i v_i(s_r) \big) \right| \\
&= \left| \frac{1}{n} \sum_i \Big(v_i(s) - \frac{1}{\sum_{r \in R^+(s)} w_i(r)} \sum_{r \in R^+(s)} w_i(r) v_i(s_r) \Big) \right| \\
&= \left| \frac{1}{n} \sum_i \Big(v_i(s) - e_i(s) \Big) \right| \leq \frac{1}{n} \sum_i \Big| v_i(s) - e_i(s) \Big|
\end{aligned}
$$

Thus, by δ-coherence of the domain we obtain that:

$$|v_{D(P)}(s) - e_{D(P)}(s)| \leq \frac{1}{n}\sum_i \Big|v_i(s) - e_i(s)\Big| < \frac{1}{n}\sum_i \delta \leq \epsilon$$

This proves that the collective opinion by D is ϵ-coherent.

(ii) Collective Coherence of I. We prove collective coherence for I similarly to the proof above for D. Let $s \in \mathcal{S}$. We assume that for any i, $|v_i(s) - e_i(s)| < \delta \leq \epsilon$. We compute the condition for the collective coherence of I at sentence s as follows:

$$\begin{aligned}
|v_{I(P)}(s) - e_{I(P)}(s)| &= \left|\frac{1}{n}\sum_i e_i(s) - \frac{1}{\sum_{r\in R^+(s)} w_{I(P)}(r)} \sum_{r\in R^+(s)} w_{I(P)}(r) v_{I(P)}(s_r)\right| \\
&= \left|\frac{1}{n}\sum_i e_i(s) - \frac{1}{\sum_{r\in R^+(s)} \lambda_r} \sum_{r\in R^+(s)} \lambda_r \big(\frac{1}{n}\sum_i e_i(s_r)\big)\right| \\
&= \left|\frac{1}{n}\sum_i \Big(e_i(s) - \frac{1}{\sum_{r\in R^+(s)} w_i(r)} \sum_{r\in R^+(s)} w_i(r) e_i(s_r)\Big)\right| \\
&= \left|\frac{1}{n}\sum_i \frac{\sum_{r\in R^+(s)} w_i(r) v_i(s_r) - \sum_{r\in R^+(s)} w_i(r) e_i(s_r)}{\sum_{r\in R^+(s)} w_i(r)}\right| \\
&= \left|\frac{1}{n}\sum_i \sum_{r\in R^+(s)} \frac{w_i(r)\Big(v_i(s_r) - e_i(s_r)\Big)}{\sum_{r\in R^+(s)} w_i(r)}\right|
\end{aligned}$$

So, by δ-coherence of the domain, we obtain that:

$$\begin{aligned}
|v_{D(P)}(s) - e_{D(P)}(s)| &\leq \frac{1}{n}\sum_i \sum_{r\in R^+(s)} \frac{w_i(r)\Big|v_i(s_r) - e_i(s_r)\Big|}{\sum_{r\in R^+(s)} w_i(r)} \\
&< \frac{1}{n}\sum_i \frac{\sum_{r\in R^+(s)} w_i(r)\ \delta}{\sum_{r\in R^+(s)} w_i(r)} \\
&= \frac{1}{n}\sum_i \delta = \delta \leq \epsilon
\end{aligned}$$

This proves that the collective opinion by I is ϵ-coherent.

(iii) Collective Coherence of B_α. We have just proven that D and I satisfy ϵ-collective coherence assuming consensus on acceptance degrees and a δ-coherent domain with $\delta < \epsilon$. It directly follows that for any $\alpha \in (0, 1)$, then B_α on a δ-coherent domain also satisfies ϵ-collective coherence.

(iv) Collective Coherence of R_α. We have just proven that D satisfies ϵ-collective coherence assuming consensus on acceptance degrees and a δ-coherent domain. ϵ-collective coherence also holds for R under the same assumptions following Proposition A.3 (see collective coherence for R). Hence, it follows that for any $\alpha \in (0, 1)$, R_α on a δ-coherent domain also satisfies ϵ-collective coherence.

□

References

Amgoud, L., Cayrol, C., Lagasquie-Schiex, M.-C., & Livet, P. (2008). On bipolarity in argumentation frameworks. *International Journal of Intelligent Systems*, *23*(10), 1062–1093.

Appgree (2021) http://www.appgree.com/. Last visited:04/2022.

Aragón, P. (2019). *Characterizing online participation in civic technologies*. Ph.D. thesis, Universitat Pompeu Fabra. Departament de Tecnologies de la Informació i les Comunicacions.

Arrow, K. J., & Maskin, E. S. (2012). *Social choice and individual values: Third edition*. Yale University Press.

Awad, E., Bonnefon, J.-F., Caminada, M., Malone, T. W., & Rahwan, I. (2017a). Experimental assessment of aggregation principles in argumentation-enabled collective intelligence. *ACM Transactions on Internet Technology*, *17*(3), 1–21.

Awad, E., Booth, R., Tohmé, F., & Rahwan, I. (2017b). Judgement aggregation in multi-agent argumentation. *Journal of Logic and Computation*, *27*(1), 227–259.

Aziz, H., Brandt, F., Elkind, E., & Skowron, P. (2017). Computational social choice: The first ten years and beyond. *Computer Science Today*, *10000*.

Baoqu (2020) https://twitter.com/_baoqu_. Last visited: 05/2023.

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

Baroni, P., & Giacomin, M. (2009). Semantics of abstract argument systems. In *Argumentation in Artificial Intelligence*, pp. 25–44. Springer.

BBC news (2019). “Cancel Brexit” petition passes 2m signatures on Parliament site. https://www.bbc.com/news/uk-politics-47652071. last visited 05/2023.

Besnard, P., & Hunter, A. (2001). A logic-based theory of deductive arguments. *Artificial Intelligence*, *128*(1-2), 203–235.

BetterReykjavik (2021) https://betrireykjavik.is. Last visited: 05/2023.

Bikakis, A., Cohen, A., Dvorák, W., Flouris, G., & Parsons, S. (2021). Joint attacks and accrual in argumentation frameworks. *Journal of Applied Logics*, *8*(6), 1437–1501.

Bochman, A. (2003). Collective argumentation and disjunctive logic programming. *Journal of logic and computation*, *13*(3), 405–428.

Bodanza, G., Tohmé, F., & Auday, M. (2017). Collective argumentation: A survey of aggregation issues around argumentation frameworks. *Argument & Computation*, *8*(1), 1–34.

Brewka, G., Pührer, J., & Woltran, S. (2019). Multi-valued grappa. In Calimeri, F., Leone, N., & Manna, M. (Eds.), *Logics in Artificial Intelligence*, pp. 85–101, Cham. Springer International Publishing.

Brewka, G., & Woltran, S. (2010). Abstract dialectical frameworks. In *Proceedings of the Twelfth International Conference on Principles of Knowledge Representation and Reasoning*, pp. 102–111.

Brewka, G., & Woltran, S. (2014). Grappa: A semantical framework for graph-based argument processing. In *Proceedings of the Twenty-First European Conference on Artificial Intelligence*, ECAI'14, p. 153–158, NLD. IOS Press.

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

Carr, C. S. (2003). Using computer supported argument visualization to teach legal argumentation. In *Visualizing argumentation: Software tools for collaborative and educational sense-making*, pp. 75–96. Springer.

Cayrol, C., & Lagasquie-Schiex, M. (2005a). Gradual valuation for bipolar argumentation frameworks. In *Symbolic and quantitative approaches to reasoning with uncertainty*, pp. 366–377. Springer.

Cayrol, C., & Lagasquie-Schiex, M.-C. (2005b). On the acceptability of arguments in bipolar argumentation frameworks. In *European Conference on Symbolic and Quantitative Approaches to Reasoning and Uncertainty*, pp. 378–389. Springer.

Chen, W., & Endriss, U. (2019). Preservation of semantic properties in collective argumentation: The case of aggregating abstract argumentation frameworks. *Artificial Intelligence*, *269*, 27–48.

City of Helsinki (2019) `https://www.hel.fi/helsinki/en/administration/participate/channels/participation-model/`. last visited 05/2023.

Consul (2021) `http://consulproject.org/en/`. Last visited: 05/2023.

Correia, M., Cruz, J., & Leite, J. (2014). On the efficient implementation of social abstract argumentation.. In *Proceedings of the European Conference on Artificial Intelligence*, pp. 225–230.

Coste-Marquis, S., Devred, C., Konieczny, S., Lagasquie-Schiex, M.-C., & Marquis, P. (2007). On the merging of dung's argumentation systems. *Artificial Intelligence*, *171*(10), 730–753.

Decidim Barcelona (2016). City of Barcelona participation portal. `https://www.decidim.barcelona/`. Last visited: 05/2023.

Decidim.Org (2016) `https://decidim.org/`. Last visited: 05/2023.

Dietrich, F. (2007). A generalised model of judgment aggregation. *Social Choice and Welfare*, *28*(4), 529–565.

Dietrich, F., & List, C. (2007). Strategy-proof judgment aggregation. *Economics & Philosophy*, *23*(3), 269–300.

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

Dung, P. M., Kowalski, R. A., & Toni, F. (2006). Dialectic proof procedures for assumption-based, admissible argumentation. *Artificial Intelligence*, *170*(2), 114–159.

Dunne, P. E., Hunter, A., McBurney, P., Parsons, S., & Wooldridge, M. (2011). Weighted argument systems: Basic definitions, algorithms, and complexity results. *Artificial Intelligence*, *175*(2), 457–486.

Eğilmez, S., Martins, J., & Leite, J. (2013). Extending social abstract argumentation with votes on attacks. In *International Workshop on Theory and Applications of Formal Argumentation*, pp. 16–31. Springer.

Endriss, U., & Moulin, H. (2016). Judgment aggregation. In Brandt, F., Conitzer, V., Endriss, U., Lang, J., & Procaccia, A. D. (Eds.), *Handbook of Computational Social Choice*, pp. 399–426. Cambridge University Press.

Escañuela Gonzalez, M. G., Budán, M. C., Simari, G. I., & Simari, G. R. (2021). Labeled bipolar argumentation frameworks. *Journal of Artificial Intelligence Research*, *70*, 1557–1636.

Fox, J., Barber, D., & Bardhan, K. (1980). Alternatives to Bayes? a quantitative comparison with rule-based diagnosis. *Methods of Information in Medicine*, *19*(04), 210–215.

Friess, D., & Eilders, C. (2015). A systematic review of online deliberation research.. Vol. 7, pp. 319–339.

Gaertner, W. (2009). *A primer in social choice theory: Revised edition*. Oxford University Press.

Gallo, G., Longo, G., Pallottino, S., & Nguyen, S. (1993). Directed hypergraphs and applications. *Discrete Applied Mathematics*, *42*(2-3), 177–201.

Ganzer, J., Criado, N., Lopez-Sanchez, M., Parsons, S., & Rodriguez-Aguilar, J. A. (2020). A model to support collective reasoning: Formalization, analysis and computational assessment. arXiv:2007.06850.

Ganzer-Ripoll, J., Criado, N., Lopez-Sanchez, M., Parsons, S., & Rodriguez-Aguilar, J. A. (2019). Combining social choice theory and argumentation: Enabling collective decision making. *Group Decision and Negotiation*, *28*(1), 127–173.

Ganzer-Ripoll, J., Criado, N., Lopez-Sanchez, M., Parsons, S., & Rodriguez-Aguilar, J. A. (2020). The directed relational framework. `https://bitbucket.org/jariiia/workspace/projects/DRF`.

Ganzer-Ripoll, J., López-Sánchez, M., & Rodriguez-Aguilar, J. A. (2016). A multi-agent argumentation framework to support collective reasoning. In *International Workshop on Conflict Resolution in Decision Making*, pp. 100–117. Springer.

García, A. J., & Simari, G. (2004). Defeasible logic programming: an argumentative approach. *Theory and Practice of Logic Programming*, *4*(1), 95–138.

Gómez, V., Kaltenbrunner, A., & López, V. (2008). Statistical analysis of the social network and discussion threads in Slashdot. In *Proceeding of the 17th International World Wide Web Conference*.

Gonzalez-Bailon, S., Kaltenbrunner, A., & Banchs, R. E. (2010). The structure of political discussion networks: A model for the analysis of online deliberation. *Journal of Information Technology*, *25*(2), 230–243.

Jackson, S. K., & Kuehn, K. M. (2016). Open source, social activism and "necessary tradeoffs" in the digital enclosure: A case study of platform co-operative, loomio. org. *tripleC: Communication, Capitalism & Critique. Open Access Journal for a Global Sustainable Information Society*, *14*(2), 413–427.

Kialo (2021) `https://www.kialo.com/`. Last visited: 05/2023.

Klein, M. (2012). Enabling large-scale deliberation using attention-mediation metrics. *Computer Supported Cooperative Work (CSCW)*, *21*(4-5), 449–473.

Klein, M., & Convertino, G. (2015). A roadmap for open innovation systems. *Journal of Social Media for Organizations*, *2*(1), 1.

Krause, P., Ambler, S., Elvang-Goransson, M., & Fox, J. (1995). A logic of argumentation for reasoning under uncertainty. *Computational Intelligence*, *11*(1), 113–131.

Lagasquie-Schiex, M.-C. (2023). Handling support cycles and collective interactions in the logical encoding of higher-order bipolar argumentation frameworks. *Journal of Logic and Computation*, *33*(2), 289–318.

Lang, J., Slavkovik, M., & Vesic, S. (2016). Agenda separability in judgment aggregation.. In *Proceedings of the 30th AAAI Conference*, pp. 1016–1022.

Leite, J., & Martins, J. (2011). Social abstract argumentation. In *Proceedings of the 22nd International Joint Conference on Artificial Intelligence*, pp. 2287–2292. AAAI Press.

List, C. (2018). Democratic deliberation and social choice. *The Oxford Handbook of Deliberative Democracy*, 463.

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

Loomio (2021) `https://www.loomio.org/`. Last visited: 05/2023.

Lopez-Sanchez, M., Serramia, M., & RODRIGUEZ-AGUILAR, J. A. (2021). Improving online debates by aggregating citizen support. In *Frontiers in Artificial Intelligence and Applications: Artificial Intelligence Research and Development*, Vol. 339, pp. 425—434. IOS Press.

Maly, J., & Wallner, J. P. (2021). Ranking sets of defeasible elements in preferential approaches to structured argumentation: Postulates, relations, and characterizations. In *Proceedings of the AAAI Conference on Artificial Intelligence*, Vol. 35, pp. 6435–6443.

McGuire, R., Birnbaum, L., & Flowers, M. (1981). Opportunistic processing in arguments.. In *Proceedings of the 7th International Joint Conference on Artificial Intelligence*, pp. 58–60.

Modgil, S., & Caminada, M. (2009). Proof theories and algorithms for abstract argumentation frameworks. In *Argumentation in Artificial Intelligence*, pp. 105–129. Springer.

Modgil, S., & Prakken, H. (2013). A general account of argumentation with preferences. *Artificial Intelligence*, *195*, 361–397.

Mongin, P. (2008). Factoring out the impossibility of logical aggregation. *Journal of Economic Theory*, *141*(1), 100–113.

Noveck, B. S. (2009). *Wiki government: How technology can make government better, democracy stronger, and citizens more powerful*. Brookings Institution Press.

Parlement & Citoyens (2015). Draft law for the recovery of biodiversity, nature and landscapes. `https://parlement-et-citoyens.fr/consultation/projet-de-loi-pour-la-reconquete-de-la-biodiversite-de-la-nature-et-des-paysages/`

presentation/presentation-et-suivi-6. last visited 10/2020. On 05/2023 redirected to https://purpoz.com/consultation/projet-de-loi-pour-la-reconquete-de-la-biodiversite-de-la-nature-et-des-paysages/presentation/presentation-et-suivi-6.

PBNYC (2021). New York City Participatory Budgeting. https://council.nyc.gov/pb/. Last visited: 05/2023.

Petitions (2019). "Revoke Article 50 and remain in the EU" at Petitions UK Government and Parliament. https://petition.parliament.uk/archived/petitions/241584. last visited 05/2023.

Polberg, S., & Hunter, A. (2018). Empirical evaluation of abstract argumentation: Supporting the need for bipolar and probabilistic approaches. *International Journal of Approximate Reasoning*, *93*, 487–543.

Prakken, H. (2005). A study of accrual of arguments, with applications to evidential reasoning. In *Proceedings of the 10th International Conference on Artificial Intelligence and Law*, pp. 85–94, Bologna, Italy.

Prakken, H. (2020). On validating theories of abstract argumentation frameworks: the case of bipolar argumentation frameworks.. In *Proceedings of the 20th Workshop on Computational Models of Natural Argument*, pp. 21–30. CEUR Workshop Proceedings.

Purpoz (2015) https://parlement-et-citoyens.fr/. last visited 10/2020. On 05/2023 redirected to https://purpoz.com/projects.

Rahwan, I., & Simari, G. R. (Eds.). (2009). *Argumentation in artificial intelligence*, Vol. 47. Springer.

Rahwan, I., & Tohmé, F. (2010). Collective argument evaluation as judgement aggregation. In *Proceedings of the 9th International Conference on Autonomous Agents and Multiagent Systems*, Toronto.

Reed, C., & Rowe, G. (2004). Araucaria: Software for argument analysis, diagramming and representation. *International Journal on Artificial Intelligence Tools*, *13*(04), 961–979.

Serramia, M., Ganzer, J., Lopez-Sanchez, M., Rodriguez-Aguilar, J. A., Criado, N., Parsons, S., Escobar, P., & Fernandez, M. (2019). Citizen support aggregation methods for participatory platforms. In *Frontiers in Artificial Intelligence and Applications: Artificial Intelligence Research and Development.*, Vol. 319, pp. 9–18. IOS Press.

Shum, S. B. (2003). The roots of computer supported argument visualization. In *Visualizing argumentation*, pp. 3–24. Springer.

Sinnott-Armstrong, W. (1999). Begging the question. *Australasian Journal of Philosophy*, *77*(2), 174–191.

Suthers, D., Weiner, A., Connelly, J., & Paolucci, M. (1995). Belvedere: Engaging students in critical discussion of science and public policy issues. In *Proceedings of the 7th World Conference on Artificial Intelligence in Education*, pp. 266–273. Washington, DC.

Sycara, K. P. (1990). Persuasive argumentation in negotiation. *Theory and decision*, *28*(3), 203–242.

Thagard, P. (2002). *Coherence in thought and action.* MIT Press.

Van Gelder, T. (2003). Enhancing deliberation through computer supported argument visualization. In *Visualizing argumentation*, pp. 97–115. Springer.

Verheij, B. (1995). Accrual of arguments in defeasible argumentation. In *Proceedings of the Second Dutch/German Workshop on Nonmonotonic Reasoning*, pp. 217–224, Utrecht.

Villata, S., Boella, G., Gabbay, D. M., & Van Der Torre, L. (2012). Modelling defeasible and prioritized support in bipolar argumentation. *Annals of Mathematics and Artificial Intelligence*, *66*, 163–197.

Vreeswijk, G. A. (1997). Abstract argumentation systems. *Artificial intelligence*, *90*(1-2), 225–279.

Walton, D., & Krabbe, E. C. (1995). *Commitment in dialogue: Basic concepts of interpersonal reasoning*. SUNY press.

Consider.it (2022) `https://consider.it/`. last visited: 05/2023.