An equational approach to the merging of argumentation networks

DOV. M. GABBAY, *Bar Ilan University, Department of Informatics, Ramant Gan, 52900, Israel; King's College London and University of Luxembourg, Luxembourg, L-13590, Luxembourg*
E-mail: dov.gabbay@kcl.ac.uk

ODINALDO RODRIGUES *Department of Informatics, King's College London, London, WC2R 2LS, UK*
E-mail: odinaldo.rodrigues@kcl.ac.uk

Abstract

This article concerns the merging of argumentation systems. We propose an equational approach to this problem by considering an augmented network containing the arguments and attacks of all systems to be merged and then associating a numerical weight to each of the components of the augmented network. The weights are calculated based on how the components are perceived by the agents associated with the systems being merged. The resulting weighted network is then used to define a system of equations, one for each argument, a solution of which corresponds to the overall level of acceptance of the arguments within the community.

Keywords: Argumentation, merging, equational approach.

1 Introduction

An argumentation system is a tuple $\langle S, R \rangle$, where S is a non-empty set of *arguments* and R is a binary relation on S representing *attacks* between the arguments [15]. One may argue that the main objective of an argumentation system is to identify sets of *winning* arguments in S, based on the interactions represented by R and an appropriate semantics determining which subsets of S can be taken as a coherent view. Such subsets are called extensions.

This article concerns the merging of argumentation systems. We imagine a family of k agents and a large set of possible arguments. Each agent a_i can see a subset S_i of these arguments and in her opinion, the attack relation should be $R_i \subseteq S_i^2$. Each agent a_i further adopts a set of winning arguments $E_i \subseteq S_i$. The agents form a community and a consensus is required. Thus our problem is to merge these k systems $\langle S_i, R_i, E_i \rangle$ into a single system representing the views of the community and draw appropriate conclusions from it.

At first, one may think that the merging process can be done at the *meta level*, i.e. by considering only the winning arguments E_i in each local system. However, as pointed out in [14], this not only will sometimes produce unintuitive results, but will also fail to simultaneously satisfy well-known social choice properties [27]. The reasons have to do with loss of information during the merging process. In particular, the extensions of winning arguments do not carry the information about the local attacks, and as it turns out, these may well be relevant to the collective decision making process. If we want to take both the local preferences for winning arguments and the local attack relations of the various systems into account, we need a framework that can deal with all this information.

Vol. 24 No. 6,
Published online November 20, 2013 doi:10.1093/logcom/ext060

Our starting point is an augmented argumentation system containing the arguments and attacks of all individual networks. We approach the merging problem from a voting perspective: agents put forward a vote on the components of the augmented system depending on how they perceive these components locally. However, the votes are not used as in an usual voting procedure such as majority voting, etc. For us, votes are used to support the idea of *reinforcement*: the more a component appears in individual networks, the more it is represented collectively. We aggregate the votes of the components resulting in an augmented argumentation system in which both arguments and attacks have weights with values in the interval $U=[0,1]$. Thus, we get a network of the form $\langle S,R,V\rangle$, where $\langle S,R\rangle$ is a traditional network and V is a function from $S\cup R$ into U. Such augmented systems can be seen a special case of *support and attack networks* [3]. We believe that the merging of argumentation systems is a scenario that naturally justifies the employment of weights in attacks and arguments.

We now have a situation whereby each agent has a traditional argumentation system, they all vote, and as a result we get a merged argumentation system with weights combined numerically. There is a mismatch between the 'type' of the original networks and that of the network resulting from their merging, and thus we need to explain how we understand the numerical weights and how we can extract/project a set of winning arguments from the merged system. Had we started working from the outset with numerical weighted systems, we would have more choice on how to perform the merging because we could use the original weights in the computation of the overall result, e.g. by constructing a new weighted argumentation system representing the group as a whole.

Given an augmented argumentation system with weights constructed as described above, we see the weights of the nodes as the overall initial level of support for the arguments in the community and the weights of the edges as the intensity with which the attacks between the arguments are carried out.

It is natural to expect that the overall support for an argument will decrease in proportion to the strength of its attacking arguments and the intensity with which these attacks are carried out. However, since the attacking arguments may themselves be attacked, we need to find a way to systematically propagate the values in the network and determine *equilibrium* values for the nodes based on their interactions, much in the spirit of an *interaction-based valuation* [11]. This is akin to finding the extensions in a traditional network. However, our work has two important differences: (i) we allow both arguments and attacks to have weights; and (ii) we calculate the equilibrium values using the equational approach of [18, 19]: the augmented system works as a generator of numerical equations whose solutions correspond to the equilibrium values.

Argumentation systems using weights in one form or another have been studied before. One of the first approaches using them was proposed by Besnard and Hunter who suggested a *categorizer* function assigning values to trees of arguments [5]. Subsequently, arguments were presented with weights used to express their relative strength within a particular audience [4]. Cayrol and Lagasquie-Schiex introduced the concept of *graduality* in the valuation of arguments in [11]. Since then, other frameworks using weights include the ones proposed in [2, 3, 6, 16, 24, 28].

The novelty of our approach is in the use of the weights to represent the support of the community for both arguments and attacks and in the way that equilibrium values for these components are calculated using a system of equations.

The rest of the article is structured as follows. In Section 2, we introduce some basic concepts and the equational approach. In Section 3, we show how the merging process is done. We then show how to calculate equilibrium values in Section 4 and illustrate the idea with many examples in Section 6. Some comparisons with related work are done in Section 7 and we finish with a discussion and some conclusions in Section 8.

2 Background

As mentioned in the previous section, given an argumentation system $\langle S,R\rangle$, one is generally interested in finding the *winning* arguments in S according to a particular semantics.

One way of doing this is to look at subsets $E \subseteq S$ that are as large as possible and yet whose arguments are *compatible* with each other. Two common notions of compatibility require E to be *conflict-free*, i.e. for all arguments X and Y in E, it is not the case that $(X,Y) \in R$; and that all arguments $X \in E$ are *acceptable*, i.e. for every argument Y in S, if $(Y,X) \in R$, there exists an argument Z in E such that $(Z,Y) \in R$. If E is conflict-free and only contains acceptable arguments, then we say that E is *admissible*. An admissible set $E \subseteq S$ that is also maximal with respect to set inclusion amongst all admissible sets is called a *preferred extension* of $\langle S,R\rangle$.

A preferred extension can be defined in terms of a complete labelling of the set of arguments that assigns *in* to arguments that are accepted; *out* to those that are rejected; and *undec* to those that are neither [10, Theorem 2]. Such labelling is called a *Caminada labelling* [10, Definition 5] and has advantages over the extension approach, because the latter only identifies the set of arguments that are accepted. We will return to this type of labelling later in the section.

In traditional argumentation systems, there is no notion of weight associated to an argument or attack. However, there are scenarios in which this association seems natural. In the case of arguments, the weights may come, for instance, from an underlying many-valued logic; as the normalized result of a vote put to a community of agents; or as the result of interactions between the arguments in a network (as in [11]). In the first case, the values are intrinsic to the arguments whereas in the last two, the values are conceptually *external* to the argumentation framework. Mixed approaches are also possible. We may start with each agent assigning numerical values via considerations which are conceptually connected to the arguments and their meaning and end up with merged values obtained during a voting procedure. The application area can dictate the most appropriate approach.

For similar reasons, an attack between arguments X and Y may also be given varying degrees of strength rather than just 0 or 1. Again, the strength may have conceptually related, internal, argumentation meaning or may be conceptually external to the arguments themselves. For example, it may be obtained from the statistics about the correlation between X and Y; or calculated from the proportion of members of a community supporting the attack of X on Y (as in [12]). It may even come from considerations about the geometry of the network itself.

An even more compelling scenario for the use of extended values is because they arise naturally in formalizms that are concerned with the problem of *merging* of argumentation systems, which we consider here. The concept was introduced by Coste-Marquis *et al.* in [14].

It may be wise when presenting a numerical argumentation network to provide not only the numerical values themselves but also to give their origin, internal or external, etc, because the origin provides the context that supports a proper interpretation of the weights.

Now, given the numerical network $\langle S,R,V\rangle$ we need to somehow figure out what the various values mean. We can regard the values given by V as *start-up values* that we may want to adjust depending on how the components interact in the network. The adjustment corresponds to the *valuation step* in Cayrol and Lagasquie-Schiex's terminology [11]. However, in our case we want arguments to be *weakened* in proportion to the strength of the attacks on them as well as the intensity with which these attacks are carried out. Ideally, we want to find *equilibrium* values for all arguments, i.e. stable values for the arguments resulting from their start-up values and the interactions with the equilibrium values of the arguments that attack them.

An interesting methodology for calculating these values is the so-called *equational approach* proposed in [18, 19] which sees a numerical network as a generator of equations. Each argument

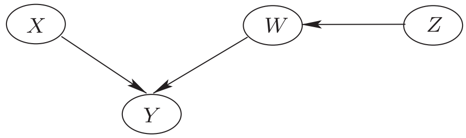


FIGURE 1. A simple argumentation system.

is associated with a variable and an equation is generated for each argument taking into account the attack relation. Provided the equations respect the meanings of the weights of the arguments and attacks an 'evaluation' of the network can be done by solving the system of equations. For an argument X, the equilibrium value 1 means that X definitely 'in'; 0 means that it is definitely 'out'; and any other value in between means how close to 'in' (or 'out') X is. If we want, we may be more flexible and settle for a threshold value for the acceptance of arguments that is lower than 1.

An example of how such equations can be generated is given by the schema $Eq_{\max}$ below, where the symbol $V_e(X)$ denotes the *equilibrium* value of a node X. Now let $Att(Y)$ denote the set of all arguments attacking Y, i.e. $Att(Y)=\{X_i \in S \mid (X_i, Y) \in R\}$. We can define the equilibrium value of Y through the $Eq_{\max}$ schema as follows.

$$V_e(Y)=1-\max_{X_i \in Att(Y)}\{V_e(X_i)\} \qquad (Eq_{\max})$$

Notice that for a node Y, $V_e(Y)=1$ if and only if $V_e(X)=0$ for all $X \in Att(Y)$ and $V_e(Y)=0$ if and only if $V_e(X)=1$ for some $X \in Att(Y)$.

Thus, the network of Figure 1. generates the following system of equations:

$$\begin{aligned} V_e(X) &= 1 \\ V_e(Z) &= 1 \\ V_e(W) &= 1-\max\{V_e(Z)\} \ (=0) \\ V_e(Y) &= 1-\max\{V_e(X), V_e(W)\} \ (=0) \end{aligned}$$

We now interpret these values as 'accept X and Z' and 'reject W and Y'. This gives the extension $\{X, Z\}$, as expected from a traditional argumentation system.

Generally speaking, Gabbay has shown that the totality of the solutions of the equations generated from a network using $Eq_{\max}$ corresponds to the totality of Caminada labellings of that network [19].

Unfortunately, $Eq_{\max}$ does not take into account the start-up value of a node or the intensity with which the attacks on it are carried out. In order to take these into account, we will consider a more sophisticated equation schema in Section 4.

3 Combining argumentation networks

In this section, we discuss some intuitions underpinning our proposed method of merging argumentation networks. Our first goal is to show how to combine a collection of networks into a single-weighted argumentation network.

As discussed in Section 1, we start by associating each network with an agent who 'votes' for components of another possibly larger network. Obviously, the more interesting scenarios involve distinct networks being merged. Consider the networks in Figure 2 and each agent's chosen extension of winning arguments in these networks.

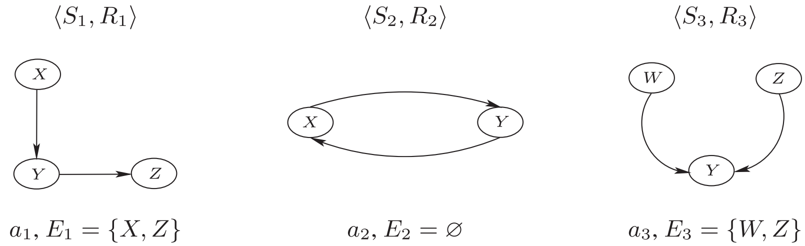


FIGURE 2. Argumentation networks of three different agents.

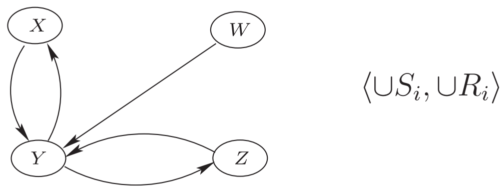


FIGURE 3. Augmented network containing all components of $\langle S_1, R_1\rangle$, $\langle S_2, R_2\rangle$ and $\langle S_3, R_3\rangle$.

One can easily see that the three agents have different sets of arguments, and even in the case where some arguments coincide, the agents may disagree with respect to the attack relationship between them. For instance, argument W is only known to agent a_3, and in her network, Z attacks Y, whereas in the network of agent a_1, Y attacks Z.

There are many reasons why agents may have different points of view regarding arguments and their attack relationship. They may use different knowledge bases, they may have different resources for inference at their disposal, they may use different inference systems, they may have different preferences for arguments, etc. All of these may result in the non-availability of some arguments to some agents as well as the generation of disagreements with respect to the direction of the attacks between them (this is not very dissimilar to the existence of cycles in a single network).

A simple way of harmonizing the differences is to consider expansions to the networks. However, unlike in [14], we do not expand each network individually and then combine the expansions. Instead, we consider the single augmented network that includes the components of all other networks and then analyse the representation of the components of this network with respect to the community of agents.

Since some components appear in more networks than others, the augmented network alone is not sufficient to represent the community. In order to do that we use weights, but let us first introduce the notion of a profile of (traditional) argumentation systems.

DEFINITION 3.1 (Profile of argumentation systems)
A profile of argumentation systems is a tuple $P = \langle AN_1, \ldots, AN_k\rangle$ where each $AN_i = \langle S_i, R_i, E_i\rangle$ is an argumentation system $\langle S_i, R_i\rangle$ for agent a_i provided with a set $E_i \subseteq S_i$ of the winning arguments in S_i selected according to some local semantics.

DEFINITION 3.2 (Augmented weighted network for a profile of argumentation systems)
Let $P = \langle AN_1, \ldots, AN_k\rangle$ be a profile of argumentation systems. The *weighted augmented network for* P is a tuple $\text{AWN}_P = \langle S, R, V_0, \xi\rangle$ where

– $S = \cup_i S_i$ and $R = \cup_i R_i$

– $V_0 : S \to [0,1]$ represents the start-up values of the arguments in S and is to be interpreted as the initial level of support for an argument $X \in S$ within P
– $\xi : R \to [0,1]$ represents the intensity of an attack $(X,Y) \in R$ within P

The values for V_0 and ξ need to be computed from the argumentation systems in the profile P. The basis for this is a policy interpreting each agent's perception of the arguments and attacks in AWN_P depending on the agent's own original network. For simplicity, we will refer generally to the arguments and attacks of a network as its 'components'.

In agreement with [14] we believe that there is an intrinsic difference between supporting a component; rejecting it and being ignorant about its existence (in which case a decision for or against it is impossible). In order to distinguish these attitudes, we let agents vote for components by assigning to them one of the three values below.

0: the agent does not know about the component
1: the agent knows about the component and supports it
−1: the agent knows about the component but does not support it

DEFINITION 3.3 (Attitude with respect to a component)
Let P be a profile of argumentation systems and AWN_P the weighted augmented network for P. The attitude of an agent a_i towards the component c of AWN_P, in symbols $v_i(c)$, is represented in the following way.[1]

$v_i(X)$	(for arguments)	$v_i\big((X,Y)\big)$	(for attacks)
0,	if $X \notin S_i$	0,	if either $X \notin S_i$ or $Y \notin S_i$ (or both)
1,	if $X \in E_i$	1,	if $(X,Y) \in R_i$
−1,	if $X \in S_i - E_i$	−1,	if $X,Y \in S_i$, but $(X,Y) \notin R_i$

That is, the agent a_i votes with 0 for *argument* X, if a_i has no knowledge about X. Otherwise, a_i will vote with 1 or −1 depending on whether X is amongst the winning arguments of S_i. Of course, X may be one of the winning arguments of S_i only because a_i is unaware about some other argument Y that defeats it. However, if that is the case, both Y and its supposed attack on X will be represented in the augmented network and this will have an effect on X's equilibrium value, as we shall see later.

The case of an attack from an argument X to an argument Y is similar, except that an attack may not exist because the agent is unaware of one or both of the arguments. Hence, the agent a_i will vote with 0 if at least one of X and Y is not known to her (in which case a judicious decision of a_i about the attack from X on Y is not possible). Otherwise, if both X and Y are known to a_i, she will vote with −1 if $(X,Y) \notin R_i$ and with 1 if $(X,Y) \in R_i$. Notice that the vote 1 for an attack (X,Y) by an agent a_i depends only on the existence of the attack in a_i's local network. Even if $Y \in E_i$ and $X \notin E_i$, a_i must still vote with 1 if $(X,Y) \in R_i$, since she knows about it. The fact that a_i chooses Y over X in spite of the attack of X on Y in this case is already taken into account in the agent's votes for the arguments X and Y.

The above voting strategy requires that there is a local semantics for deciding the winning arguments in each network but does not make any assumptions on what the semantics should be. In fact, the group as a whole may have several different local semantics. If local preference orderings are used to decide about the winning arguments of each agent, then it would make sense to aggregate all of these preference orderings and take the aggregated result into account when deciding on the extensions of the aggregated network.

[1] To simplify notation we use the same function symbol v_i for nodes and edges.

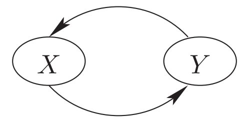


FIGURE 4. Argumentation networks for agents a_1 and a_2.

In either case, we find it reasonable to assume that an agent will be able to make some decisions about her winning arguments and these preferences should somehow be used during the merging. This is especially important when it concerns the representation of unanimity of local decisions. In order to illustrate this, suppose that agents a_1 and a_2 share the same argumentation network given in Figure 4.

Under the preferred semantics, $E(a_1)=\{X\}$ or $E(a_1)=\{Y\}$ (a_2 has the same choices). For the sake of discussion, let us assume that $E(a_1)=E(a_2)=\{X\}$, i.e., both agents have a preference for X over Y, and, as a result, they both take the extension $\{X\}$ to be their set of winning arguments. We would expect that the result of the merging of these networks would contain X only.[2] However, this is only possible if these preferences are somehow taken into account in the merging (as is the case in our approach). Thus, a formalizm that does not consider such local preferences (such as [14]) and which would still agree that merging the networks for a_1 and a_2 would result in a_1 $(=a_2)$ has no recourse when it comes to deciding between $\{X\}$ and $\{Y\}$, even though all agents involved in the merging prefer the former to the latter (we will come back to this example in Section 6).

If the local networks are themselves numerical, then a number of alternatives arise. One could compute each network individually, decide on the winning arguments and apply the same technique given above; or one could feed the equilibrium values of each network into the augmented one, normalize the values as appropriate, generate the equations and then compute the overall equilibrium values as before; or one could choose a combination of these ideas. However, in this work we want to keep our assumptions to a minimum so we only require that the agents can provide a set of winning arguments.

3.1 Calculating initial weights for nodes and attacks

We now need to generate the initial weights for the augmented network based on each agent's attitude to its components. Again, because some components are only known to some agents, the community as a whole may take two different approaches when considering the overall level of support of a component:[3]

- in the *credulous* approach, the weights are calculated based on the total number of agents *that know about a component*
- in the *sceptical* approach, the weights are calculated taking into account the total number of agents in the profile P

[2]This expectation does not hold if we consider simply the merging of *graphs*, but the objective of the merging of argumentation systems is to provide the collective view of a community of agents with respect to a set of arguments and therefore everything that involves the way the agents see the arguments and how they interact with each other must play a part in the process.

[3]Credulous and sceptical reasoning was put forward at the level of each individual agent in [8] to tailor the reasoning according to how critical the context of the argumentation is.

We will associate the credulous approach with the superscript $^+$ and the sceptical one with the superscript $^-$ in the definitions of the initial values V_0 and ξ below. Whenever the distinction is not important we will simply omit the superscripts.

Definition 3.4 (Initial values for nodes and attacks)
Let $P = \langle AN_1, \ldots, AN_k \rangle$ be a profile of argumentation systems and AWN_P the weighted augmented network for P. Let $v^+(c) = |\{i \mid v_i(c) = 1\}|$ and $v^-(c) = |\{i \mid v_i(c) = -1\}|$.[4] We define

$$V_0^+(X) = \frac{v^+(X)}{v^+(X)+v^-(X)} \qquad V_0^-(X) = \frac{v^+(X)}{k}$$

$$\xi^+\big((X,Y)\big) = \frac{v^+\big((X,Y)\big)}{v^+\big((X,Y)\big)+v^-\big((X,Y)\big)} \qquad \xi^-\big((X,Y)\big) = \frac{v^+\big((X,Y)\big)}{k}$$

where $V_0^+(X)$ (resp., $V_0^-(X)$) is the start-up value for the argument X under the credulous (resp., sceptical) approach and $\xi^+\big((X,Y)\big)$ (resp., $\xi^-\big((X,Y)\big)$) is the intensity of the attack from X to Y within AWN_P under the credulous (resp., sceptical) approach.

Notice that we have purposefully excluded the agents who do not know about a component c in the definitions of $V_0^+(c)$ and $\xi^+(c)$ above. These agents vote with 0 for c according to Definition 3.3 and hence are not counted in either $v^+(c)$ or $v^-(c)$. $V_0^-(c)$ and $\xi^-(c)$ on the other hand look at the components more sceptically and consider their representation across all voters.

For the example in Figure 2 we get the weights shown in Figure 5 for the components of the augmented networks under both approaches. Given these weights, we then need to calculate equilibrium values for the nodes, which will be done in Section 4.

By referring to Figure 2, where the winning arguments in each network are given, one can easily see why the initial weight of the argument Y is 0 under both approaches (see Figure 5). This is because Y is not a winning argument in any of the initial networks (see Proposition 3.5).

Analogously, the initial value of Z in Figure 5 is 1 only under the credulous approach. This is because even though Z is a winning argument in every network in which it is known, it is not known in every network in the profile. Similarly, W's initial weight is 1 under the credulous approach, but only $1/3$ under the sceptical one. This is reasonable, since it is only known by one out of the three agents, but for that agent (a_3), it is one of the winning arguments. The weights for the attacks are assigned following the same pattern.

Generally speaking, we have the following.

Proposition 3.5
Let $P = \langle AN_1, \ldots, AN_k \rangle$ be a profile of argumentation systems where each $AN_i = \langle S_i, R_i, E_i \rangle$ and let $\text{AWN}_P = \langle S, R, V, \xi \rangle$ be the weighted augmented network for P. The following hold for all arguments $X \in S$.

1. if $X \in \cap_i E_i$, then $V_0^+(X) = V_0^-(X) = 1$
2. if $V_0^-(X) = 1$, then $X \in \cap_i E_i$
3. if $X \in E_i$ for all i such that $X \in S_i$, then $V_0^+(X) = 1$
4. if $X \notin \cup_i E_i$, then $V_0^+(X) = V_0^-(X) = 0$

[4]It is not difficult to see that $v^+(c) + v^-(c) > 0$, for all components c.

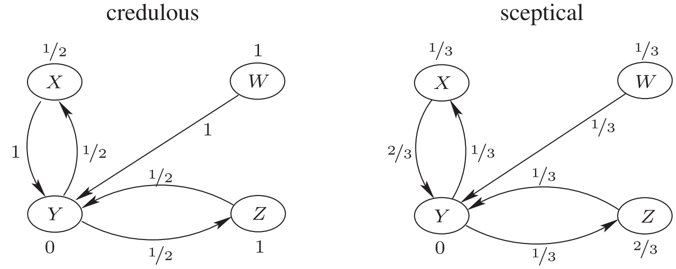


FIGURE 5. Merged networks of Figure 2 under the credulous and sceptical approaches.

PROOF. 3.5. and 3.5. follow directly from Definitions 3.3 and 3.4. For 3.5., note that if $V_0^-(X)=1$, then $v^+(X)=k$, and hence $X\in E_i$ for every agent a_i. For 3.5., note that if $X\in E_i$ for all i such that $X\in S_i$, then $v^-(X)=0$, and hence $V_0^+(X)=1$. ■

The situation with attacks is similar, but simpler.

PROPOSITION 3.6
For all attacks $(X,Y)\in R$.

1. if $(X,Y)\in\cap_i R_i$, then $\xi^+\big((X,Y)\big)=1$ and $\xi^-\big((X,Y)\big)=1$.
2. if $\xi^-\big((X,Y)\big)=1$, then $(X,Y)\in\cap_i R_i$.
3. $(X,Y)\in\cup_i R_i$ if and only if $\xi^+\big((X,Y)\big)>0$ (resp. $\xi^-\big((X,Y)\big)>0$)

PROOF. These follow directly from Definitions 3.3 and 3.4. ■

Proposition 3.6 states the following. Attacks appearing in all networks are transmitted with full intensity (i.e. 1) in both approaches. If the intensity of an attack is 1 under the sceptical approach, then the attack appears in all networks. Notice that under the credulous approach, an attack may be transmitted with full intensity even if it does not appear in all networks, as long as the networks in which it does not appear do not have at least one of the nodes involved in the attack. The corresponding agents will vote with 0, which under the credulous approach indirectly means that they vote with the networks that support the attack. Under the credulous approach, the value of the weight of a component can only be reduced by the agents that vote against it (i.e. vote with -1). Finally, any attack appearing in at least one network is carried out with some non-null intensity (and vice versa).

We now turn to the problem of calculating equilibrium values for the arguments of a weighted augmented network.

4 Equilibrium values in a weighted augmented network

One important aspect in the calculation of the equilibrium values of the arguments in a weighted augmented network is the decision of how the attacks to an argument should affect its initial support value.

As in an usual argumentation system, arguments may be attacked by any number of arguments. Since we work with numerical values, we want to reduce the strength of the attacked node by a measure that is the result of the aggregation of the strength of the attacking nodes. The strength of a single attack itself depends on the strength of the attacking node and the intensity with which the attack is carried out. However, the attacking nodes may be themselves attacked, so we

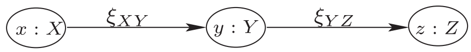


FIGURE 6. A typical weighted argument network.

need to perform the aggregation systematically.[5] Let us start by analysing the effect of attacks in general.

Consider the network in Figure 6, in which x, y and z are the initial values of the arguments X, Y and Z, respectively. Let us for a moment ignore these values.

If we want to mimic the behaviour of attacks in a standard argumentation system [15], we need to accept arguments X and Z and reject argument Y. The reasoning is as follows. Since no arguments attack X, it *persists*. X then attacks Y, which is *defeated*, and hence no persisting arguments attack Z, which then consequently also persists. In our numerical semantics, persistence is associated with the values $[t,1]$ (for some $t>0$) and (strong) defeat with the value 0. Of course there is a grey area for the values $0<v<t$, which show *some* support for arguments whose values did not meet the threshold t for acceptance. In other words, acceptance in Dung's sense for us means to have an equilibrium value e equal or higher than a minimum acceptance level $t>0$. If we want to be strict, we can set $t=1$. Otherwise, we may settle for any value greater than 0 (up to 1).

Ideally, we would like to remain close to the basic semantics, taking care of the arguments' start-up values (which are all in the unit interval U) and the intensity with which the attacks between them are carried out. Hence, our objective in Figure 6 is to calculate the values $V_e(X)$, $V_e(Y)$ and $V_e(Z)$, based on x, y, z, ξ_{XY} and ξ_{YZ}. Arguably, since X is not attacked by any node, its equilibrium value $V_e(X)$ can be calculated directly by some manipulation on the value x alone. The simplest procedure is to make $V_e(X)=x$, its initial value. On the other hand, the value of $V_e(Y)$ depends both on $V_e(X)$ and the *intensity* ξ_{XY} with which the attack from X to Y is carried out. Once $V_e(Y)$ is calculated, the equilibrium value for $V_e(Z)$ can be calculated using ξ_{YZ} in the same way. If there are cycles, the equations get more complex, but they are solvable, as long as the functions involved are all continuous, because of Brouwer's fixed-point theorem.[6]

Now suppose we give initial value 1 to all arguments and consider all attacks being transmitted with full intensity. Since X has initial value 1 and it is not attacked by any arguments, its equilibrium value is the same as its initial value, i.e. 1. It then attacks Y with full intensity (i.e. $\xi_{XY}=1$), which means that Y's initial value, $y=1$, is weakened by 1 and its equilibrium value is set to 0. Effectively, this annihilates Y's attack on Z, which then gets as its equilibrium value the same value as its initial one, i.e. 1. As a result, we end up with the acceptance of X (because of its equilibrium value 1); the rejection of Y (because of its equilibrium value 0); and the acceptance of Z (also because of its equilibrium value 1), as expected.

We stress that, in general, we are free to decide on the minimum value we require for considering an argument as being accepted. As we mentioned, we may decide this to be the value 1 itself, leaving all values $0<x<1$ to represent *undecided* arguments; or we may even do away with the notion of undecidedness altogether and divide the interval in two halves only: acceptance or rejection of an argument depends on which interval its equilibrium value falls into.

In terms of the effect of the attacks on some argument X, our problem is to determine a factor $0\leq\pi(X)\leq 1$ representing the combined strength of the attacks on it. The equilibrium value for X

[5]If there are no loops, the problem is simpler because we can always start with the nodes that are not attacked by any nodes and propagate the values up the graph.

[6]This and some other related issues will be explored in more detail in a forthcoming paper dealing with the more mathematical issues of this approach.

can then be calculated by multiplying X's initial value by this factor, i.e. $V_e(X) = V_0(X) \cdot \pi(X)$. The function π must *aggregate* the value of the attacks on an argument. In order to remain close to the standard argumentation semantics, we want π to satisfy at least the three conditions below.

(SSC1) $\pi(X) = 1$, if $max_{Y \in Att(X)}\{\xi\big((Y,X)\big)V_e(Y)\} = 0$
(SSC2) $\pi(X) = 0$, if $max_{Y \in Att(X)}\{\xi\big((Y,X)\big)V_e(Y)\} = 1$
(SSC3) π is continuous

(SSC1) says that if all arguments attacking X are fully defeated or transmitted with null intensity, then X should retain its initial value fully. (SSC2) says that if any argument that attacks X has full strength *and* the attack is carried out with full intensity, then X should be fully defeated. (SSC3) ensures that the considerations about the interactions between the nodes are robust, i.e. that small changes in the initial values do not cause sudden variations in the equilibrium ones.

Thus, the basic idea is that the stronger an attack is, the closer its value gets to 1 and hence the closer we want π to get to 0 so that the equilibrium value of the attacked argument can decrease proportionally (since its initial value is multiplied by π). In the case of a single attack of strength u to node X with transmission factor κ, one possibility is to make $\pi(X) = 1 - \kappa u$. In the network of Figure 6, this would make $\pi(Y) = 1 - \xi\big((X,Y)\big)V_e(X)$ and hence Y's equilibrium value would be $V_e(Y) = V(Y) \cdot (1 - V_e(X)) = 1 \cdot 0 = 0$, as expected.

Besnard and Hunter's *categoriser* function [5] is an example of a function satisfying (SSC1)–(SSC3) (more on this in Section 7).

We still have to tackle the problem of a node being attacked by multiple arguments, i.e. what can we say about $\pi(X)$ when X is attacked by multiple arguments?

As usual, attacking arguments combine via *multiplication*, which is compatible with the behaviour of conjunction in Boolean logic and in probability. The equations for the equilibrium values of the nodes of a weighted augmented network are defined below.

DEFINITION 4.1 (Equilibrium value of an argument)
Let $P = \langle AN_1, \ldots, AN_k \rangle$ be a profile of argumentation systems and $\text{AWN}_P = \langle S, R, V, \xi \rangle$ the weighted augmented network for P. The equilibrium value of an argument $X \in S$ is defined be the equation:

$$V_e(X) = V_0(X) \cdot \prod_{Y_i \in Att(X)} (1 - \xi\big((Y_i, X)\big)V_e(Y_i)) \qquad (Eq_{\text{inv}})$$

One can choose V_0 and ξ to be V_0^+ and ξ^+ or V_0^- and ξ^- depending on whether a credulous or sceptical approach is desired (this will be explored further in Section 6). Notice that the highest possible intensity of the attack by an argument Y is $V_0(Y)$ itself. This happens when the attack is carried out with full intensity and Y is not itself attacked by any node—in this case it retains its initial value fully, i.e. $V_e(Y) = V_0(Y)$. Because we take the complement of this attack to 1, in such circumstances this is sufficient to make the equilibrium value of the attacked argument 0.

The equation Eq_{max} (defined on Section 2) decreases the initial support value of an argument according to the value of the strongest attack on it. The effect of attacks according to equation Eq_{inv} (see Definition 4.1) on the other hand is *cumulative*: it aggregates the strength of the attacking nodes. The intuition is that each challenge to an argument contributes to decreasing its overall credibility.

Henceforth, we formally set the value $\pi(X)$ to $\prod_{Y_i \in Att(X)} (1 - \xi\big((Y_i, X)\big)V_e(Y_i))$.

PROPOSITION 4.2
π satisfies (SSC1)–(SSC3).

PROOF. If $\max_{Y \in Att(X)}\{\xi\big((Y,X)\big)V_e(Y)\}=0$, then by Definition 4.1, $\prod_{Y_i \in Att(X)}(1-\xi\big((Y_i, X)\big)V_e(Y_i))=1$. Therefore, (SSC1) is satisfied. If $\max_{Y \in Att(X)}\{\xi\big((Y,X)\big)V_e(Y)\}=1$, then by Definition 4.1, for some $Y' \in Att(X)$, $1-\xi\big((Y',X)\big)V_e(Y')=0$, and then $\prod_{Y_i \in Att(X)}(1-\xi\big((Y_i,X)\big)V_e(Y_i))=0$. Hence (SSC2) is also satisfied. (SSC3) is trivially satisfied. ■

Combining attacks in this way was initially proposed in [3].

It is easy to see that when all attacks are carried out with full intensity, $\pi(X)$ can be defined in a simpler way as

$$\pi(X)=\prod_{Y \in Att(X)}\big(1-V_e(Y)\big)$$

which is equivalent to

$$\pi(X)=1-\curlyvee_{Y \in Att(X)}V_e(Y) \tag{1}$$

where $a \curlyvee b=a+b-a.b$ and for $\Delta=\{a_1,\dots,a_k\}$, $\curlyvee\Delta=((a_1 \curlyvee a_2)\curlyvee\dots\curlyvee a_k)$. The expression in (1) corresponds to the complement of the probabilistic sum t-conorm used by Leite and Martins in [24]. In probability theory, the probabilistic sum expresses the probability of the occurrence of independent events. Since we want to weaken the value of the attacked node, we take the complement of this sum to 1.

It is worth emphasizing that the equilibrium value of a node can never be higher than its initial value.

PROPOSITION 4.3
For arguments X, $V_e(X) \leq V_0(X)$.

PROOF. Straightforward. Note that $V_e(X)=V_0(X)\cdot\pi(X)$. By Definition 3.4, for all arguments Y, $0 \leq V_0(Y) \leq 1$. By Definition 4.1, $0 \leq \pi(X) \leq 1$ and hence $V_e(X) \leq V_0(X)$. ■

PROPOSITION 4.4 (Preservation of acceptance under consensus)
Let $P=\langle AN_1,\dots,AN_k\rangle$ be a profile of argumentation systems where each $AN_i=\langle S_i,R_i,E_i\rangle$ and let $\text{AWN}_P=\langle S,R,V,\xi\rangle$ be the weighted augmented network for P. If each E_i is conflict-free and $X \in \cap_i E_i$, then $V_e(X)=1$.

PROOF. By Proposition 3.5, if $X \in \cap_i E_i$, then $V_0^+(X)=V_0^-(X)=1$. Suppose $(Y,X) \in R_i$, for some argumentation framework AN_i. Since each E_i is conflict-free, then $Y \notin E_i$ and hence $Y \notin \cup_i E_i$. By Proposition 3.4, $V_0^+(Y)=V_0^-(Y)=0$ and by Proposition 4.3, $V_e(Y)=0$. It follows that $\pi(X)=1$ and hence $V_e(X)=1$. ■

If each E_i is conflict-free and $V_e(X)=1$ under the sceptical approach, then it follows that $X \in \cap_i E_i$. This is not necessarily the case under the credulous approach, because under the latter the initial support value of an argument is set to 1 as long as it wins in every argumentation system *in which it is known*. Consequently, as long as there are no attacks on X its equilibrium value will also be 1 (see Example 6. in Section 6). It is worth emphasizing that the flip-side of this credulity is that attacks (which also influence the equilibrium value of an argument) are treated in the same way. This is illustrated in Example 6. of Section 6, where the equilibrium value of the argument Y is lower in the credulous approach than in the sceptical one as the result of credulously accepting an argument that attacks it.

PROPOSITION 4.5 (Preservation of rejection under consensus)
Let $P=\langle AN_1,\dots,AN_k\rangle$ be a profile of argumentation systems where each $AN_i=\langle S_i,R_i,E_i\rangle$ and let $\text{AWN}_P=\langle S,R,V,\xi\rangle$ be the weighted augmented network for P. If $X \notin \cup_i E_i$, then $V_e(X)=0$.

PROOF. By Proposition 3.5, if $X \notin \cup_i E_i$, then $V_0^+(X) = V_0^-(X) = 0$. By Definition 4.1, $V_e(X) = 0$. ■

The equilibrium values are largely (but not solely) dependent on the initial values, which in turn depend on the level of acceptance of each argument in the profile. If an argument is not accepted in any of the networks in the profile, then its initial support value will be null and, consequently, so will its equilibrium value. The converse is not necessarily true though, since $V_e(X)$ is also null when $\pi(X) = 0$.

5 Thresholds for acceptance

The equilibrium values simply represent how the initial overall level of support for a component is affected by the interactions with the other components in the network. Propositions 4.4 and 4.5 go some way into explaining what to expect from the equilibrium values in special cases. If one wants to make a decision on what arguments to accept overall, an appropriate threshold for acceptance for the network at hand must be decided.

The value 1 represents the strongest possible level of acceptance. We have seen that under the sceptical approach an equilibrium value of 1 means that the argument belongs to the sets of winning arguments of all agents, but setting 1 as the minimum acceptance level could prove too strict. The concept of *majority* is sometimes used in some voting systems. Acceptance of an argument by a clear majority of the networks in a profile produces a start-up value for the argument strictly greater than $1/2$, as can be seen below.

PROPOSITION 5.1 (Majority of acceptance)
Let $P = \langle AN_1, \ldots, AN_k \rangle$ be a profile of argumentation systems where each $AN_i = \langle S_i, R_i, E_i \rangle$ and let $\text{AWN}_P = \langle S, R, V, \xi \rangle$ be the weighted augmented network for P. If $|\{i \mid X \in E_i\}| > k/2$, then $V_0^+(X) > 1/2$ and $V_0^-(X) > 1/2$.

PROOF. This comes straight from Definition 3.4. $v^+(X) = |\{i \mid v_i(X) = 1\}| = |\{i \mid X \in E_i\}|$. Notice that $v^+(X) + v^-(X) \leq k$. If $v^+(X) > k/2$, then $V_0^+(X) > 1/2$. The same applies to $V_0^-(X)$, since $v^+(X) = k/2 + \epsilon$, for some ϵ and $\frac{k/2+\epsilon}{k} > 1/2$. ■

However, having a start-up value greater than $1/2$ is not sufficient to guarantee an *equilibrium* value greater than $1/2$, because there could be arguments that attack X in those networks in which it was not accepted and if the equilibrium values of those arguments is greater than 0, they will contribute to bring X's equilibrium value down. We argue that this is a good feature and addresses the problem of 'voting relying only on the selected extensions' (this is presented as 'Problem 2.' in [14]). Majority of acceptance alone does not guarantee overall group acceptance, precisely because of the attack relations and this is reflected in the calculation of the equilibrium values.

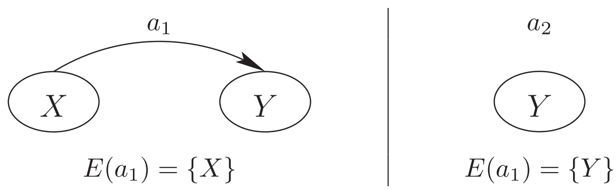
Alternatively, one could define the acceptance value as the maximum of the equilibrium values (if that is greater than 0). A simpler approach (adopted here) is to set the threshold for acceptance as the average value of the equilibrium values and accept the arguments whose equilibrium values are greater or equal to it. This is akin to deciding the grades in an exam by looking at the distribution of the marks, dividing it into clusters according to the average, and associating the top cluster with the best grade.

In specific scenarios a different threshold value can be defined through a more sophisticated analysis of the networks in the profile in a similar way to how it is done in [2] (which is itself based on the notion of the 'inconsistency degree' of a knowledge base). This investigation itself is quite complex and left for future work.

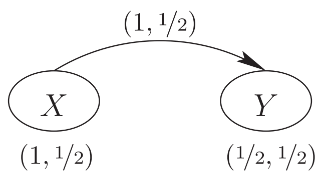
6 Worked examples

We now illustrate our technique with a few examples. In each example, we show the input networks and the resulting augmented (merged) network with its components annotated with the initial weights for each approach in the form (credulous,sceptical). The equations for all nodes are given as well as the corresponding equilibrium values calculated under both the credulous and sceptical approaches. The accepted arguments are indicated within a shadowed box (they have equilibrium values greater or equal than the average). Each example is followed by a discussion of the results.

1. **Input networks**



Merged network



Credulous

$V_e(X) = V_0(X)$
$V_e(Y) = V_0(Y)(1 - V_0(X))$

$V_e(X) = 1$
$V_e(Y) = 0$
$avg = 1/2$

Sceptical

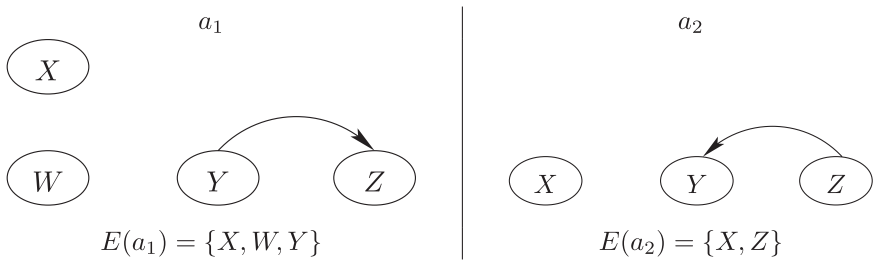
$V_e(X) = V_0(X)$
$V_e(Y) = V_0(Y)(1 - 1/2 \cdot V_0(X))$

$V_e(X) = 1/2$
$V_e(Y) = 3/8 = 0.375$
$avg = 0.437$

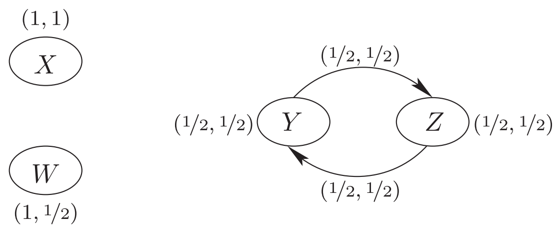
In this example, under the credulous approach $V_0(X) = 1$ and since X has no attacks, $V_e(X) = 1$. Its attack on Y is transmitted with full intensity. $V_0(Y) = 1/2$. Therefore, $V_e(Y) = 1/2 \cdot (1 - 1) = 0$. Under the sceptical approach $V_0(X) = 1/2$ and hence $V_e(X) = 1/2$. Its attack on Y is transmitted with intensity $1/2$. Therefore, $V_e(Y) = 1/2 \cdot (1 - 1/2 \cdot 1/2) = 3/8$. Note that the sceptical approach produces a higher equilibrium value for Y because under the credulous approach X is fully accepted and its attack on Y fully defeats it. The only argument with equilibrium value above the average of the values is X in both approaches and therefore it is the only one accepted. As for the credulous approach, it is easy to see why X should be accepted and Y rejected. There is no dispute about the acceptance of X, but under the credulous approach so is the case about its attack on Y, which is then defeated. Under the sceptical approach, the results can be explained as follows. Both agents know about Y, so its equilibrium value should not be null. However, even though

X is not unanimously known (and hence accepted), there is no knowledge of any attacks on it. X's equilibrium value of $1/2$ reflects its support across the community. Y's initial support value has to be adjusted down by the fact that one of the agents supports an attack on it. Taking the average of the two values means that Y ends up being rejected.

2. **Input networks**



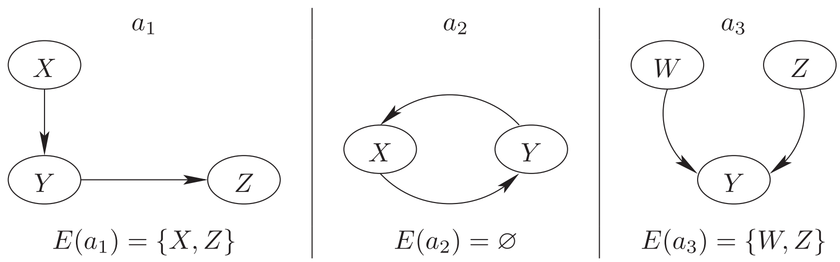
Merged network



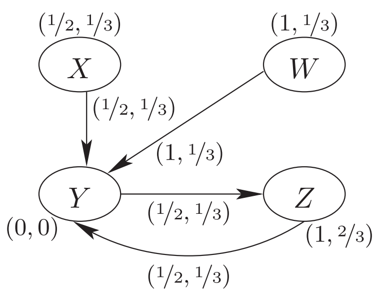
Credulous	**Sceptical**
$V_e(X)=V_0(X)$	$V_e(X)=V_0(X)$
$V_e(W)=V_0(W)$	$V_e(W)=V_0(W)$
$V_e(Y)=V_0(Y)(1-1/2\cdot V_e(Z))$	$V_e(Y)=V_0(Y)(1-1/2\cdot V_e(Z))$
$V_e(Z)=V_0(Z)(1-1/2\cdot V_e(Y))$	$V_e(Z)=V_0(Z)(1-1/2\cdot V_e(Y))$
$V_e(X)=1$	$V_e(X)=1$
$V_e(W)=1$	$V_e(W)=1/2$
$V_e(Y)=2/5$	$V_e(Y)=2/5$
$V_e(Z)=2/5$	$V_e(Z)=2/5$
$avg=0.7$	$avg=0.575$

In this example, both agents accept argument X and there are no attacks on it in any network. Thus, regardless of the approach, the equilibrium value of X is 1. This value under the sceptical approach indicates unanimity of acceptance of X. Compare this with the previous example, where X's equilibrium value smaller than 1 under the sceptical approach signals some disagreement about its universal acceptance. In spite of there not being any attacks on W, this argument is only known by agent a_2. Under the credulous approach $V_e(W)=1$, but under the sceptical approach $V_e(W)=1/2$, since it is accepted by only half of the community. Arguments Y and Z are also accepted by half of the community, but in each case, the other half supports a complementary attack of one on the other. As a result, their equilibrium values are both reduced from $1/2$ to $2/5$. Arguments X and W have equilibrium values above the average under the credulous approach and hence are accepted there, but under the sceptical approach only X is accepted, since W's equilibrium value takes into account the fact that it is not known by a_2. Notice that our approach provides some indication about the relative acceptability levels of all arguments, including those that are not accepted. In this case, Y and Z are at the same level but below W. This reflects the fact that no attacks on W are known.

3. **Input networks**



Merged network



Credulous

$$
\begin{aligned}
V_e(X) &= V_0(X)\\
V_e(W) &= V_0(W)\\
V_e(Y) &= V_0(Y)(1-1/2\cdot V_e(X))\\
&\quad (1-1/2\cdot V_e(W))\\
&\quad (1-1/2\cdot V_e(Z))\\
V_e(Z) &= V_0(Z)(1-1/2\cdot V_e(Y))
\end{aligned}
$$

$V_e(X)=1/2$

$V_e(W)=1$

$V_e(Y)=0$

$V_e(Z)=1$

$avg=5/8$

Sceptical

$$
\begin{aligned}
V_e(X) &= V_0(X)\\
V_e(W) &= V_0(W)\\
V_e(Y) &= V_0(Y)(1-1/3\cdot V_e(X))\\
&\quad (1-1/2\cdot V_e(W))\\
&\quad (1-1/2\cdot V_e(Z))\\
V_e(Z) &= V_0(Z)(1-1/3\cdot V_e(Y))
\end{aligned}
$$

$V_e(X)=1/3$

$V_e(W)=1/3$

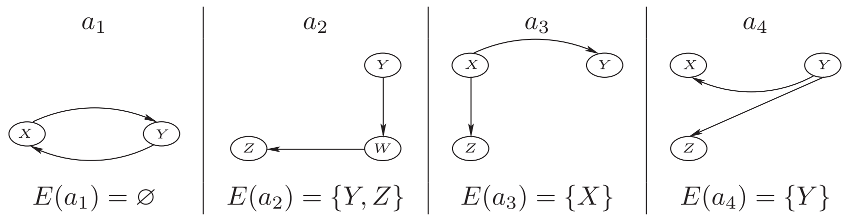
$V_e(Y)=0$

$V_e(Z)=2/3$

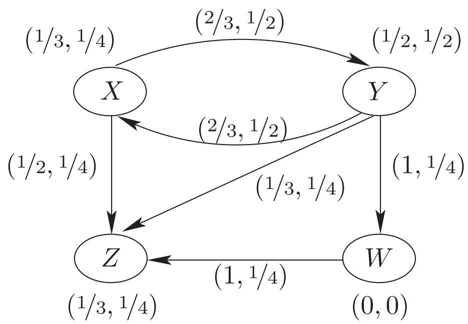
$avg=1/3$

This is the example appearing in Figure 2. We start with argument Y, which does not feature in any of the sets of the agents' winning arguments. Its initial support value is null and hence its equilibrium value is also null. This leaves X's initial support values unchanged in both approaches. Under the credulous approach both W and Z get equilibrium value 1 (there are no attacks on W and Z is accepted by the majority). Under the sceptical approach Z's equilibrium value is the highest, because it is accepted by 2/3 of the agents (as opposed to X and W which are accepted by only 1/3 of them). Both W and Z have equilibrium values equal to 1 under the credulous approach and hence are accepted, but under the sceptical approach only Z is accepted (note that it is the only argument accepted by the majority of the agents).

4. **Input networks**



Merged network



Credulous

$$
\begin{aligned}
V_e(X) &= V_0(X)(1-{}^{2}\!/_{3}\cdot V_e(Y))\\
V_e(W) &= V_0(W)(1-V_e(Y))\\
V_e(Y) &= V_0(Y)(1-{}^{2}\!/_{3}\cdot V_e(X))\\
V_e(Z) &= V_0(Z)(1-{}^{1}\!/_{2}\cdot V_e(X))\\
&\quad (1-{}^{1}\!/_{3}\cdot V_e(Y))\\
&\quad (1-V_e(W))
\end{aligned}
$$

$V_e(X)={}^{12}\!/_{50}=0.24$
$V_e(W)=0$
$V_e(Y)={}^{21}\!/_{50}=0.42$
$V_e(Z)={}^{473}\!/_{1875}\approx 0.252$
$avg\approx 0.228$

Sceptical

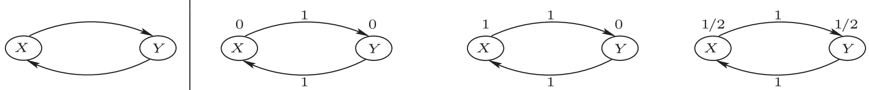
$$
\begin{aligned}
V_e(X) &= V_0(X)(1-{}^{1}\!/_{2}\cdot V_e(Y))\\
V_e(W) &= V_0(W)(1-{}^{1}\!/_{4}\cdot V_e(Y))\\
V_e(Y) &= V_0(Y)(1-{}^{1}\!/_{2}\cdot V_e(X))\\
V_e(Z) &= V_0(Z)(1-{}^{1}\!/_{4}\cdot V_e(X))\\
&\quad (1-{}^{1}\!/_{4}\cdot V_e(Y))\\
&\quad (1-{}^{1}\!/_{4}\cdot V_e(W))
\end{aligned}
$$

$V_e(X)={}^{6}\!/_{31}\approx 0.194$
$V_e(W)=0$
$V_e(Y)={}^{14}\!/_{31}\approx 0.452$
$V_e(Z)\approx 0.211$
$avg\approx 0.214$

This example was introduced in [14] and is further discussed in Section 7. It shows a high level of disagreement between all agents. In particular, no single argument is accepted by all of these agents. The argument which is accepted by the highest number of agents is Y (accepted by half of the agents), followed by X and Z (accepted by one agent only). Y obtains the highest equilibrium values in both the credulous and sceptical approaches. This value is well above the average values of the arguments in each approach and as a result Y is accepted under both. X and Z are accepted under the credulous approach with equilibrium values just above the average of the values there. However, they fail to be accepted under the sceptical approach because their values are below the average of all values. W is not accepted by any of the agents and as a result it is not accepted in the result of the merging under either approach.

5. This example was introduced in Section 3. We consider the case of two agents sharing the same argumentation network in scenarios in which they sometimes agree and sometimes disagree with respect to the winning arguments in the network.

Input networks | **Merged networks**



Input networks	Merged networks		
a_1, a_2	$E(a_1) = E(a_2) = \varnothing$	$E(a_1) = E(a_2) = \{X\}$	$E(a_1) = \{X\}$ $E(a_2) = \{Y\}$

Credulous and Sceptical

$V_e(X) = V_0(X)(1 - V_e(Y))$
$V_e(Y) = V_0(Y)(1 - V_e(Y))$

(a)	**(b)**	**(c)**
$E(a_1) = E(a_2) = \varnothing$	$E(a_1) = E(a_2) = \{X\}$	$E(a_1) = \{X\}\ E(a_2) = \{Y\}$
$V_e(X) = 0$	$V_e(X) = 1$	$V_e(X) = 1/3$
$V_e(Y) = 0$	$V_e(Y) = 0$	$V_e(Y) = 1/3$
$avg = 1/2$	$avg = 1/2$	$avg = 1/3$

This example includes three subcases. The equations are the same because the attacks are the same, only the start-up values need to change and they correspond to the subcases. (a) If both agents do not accept any of the arguments, then no arguments will be accepted in the result of the merging. This is because the equilibrium value of a node is always less than or equal its initial value and the initial value of the arguments in this case are all null since they are not supported by any of the agents (cf. Proposition 4.5). (b) On the other hand, if they all accept the same arguments (for instance, $\{X\}$[7]) and provided their extensions are conflict-free (which is the case here), these arguments will get equilibrium value 1 and consequently will be accepted (cf. Proposition 4.4), agreeing with their own local preferences. These local preferences are not taken into account in [14] and, consequently, some other decision mechanism would have to be employed to deal with the results of the merged network there. (c) Finally, if the agents are divided with respect to the winning arguments, the resulting equilibrium values will reflect that. In this case, the agents are equally split. Even though each argument starts with value $1/2$, they end up with the lower value $1/3$ due to the attacks these arguments make on each other. Notice that the use of *avg* causes all of the arguments to be accepted (since they have equal equilibrium values). Setting the threshold at $1/2$ would cause them to be rejected.

7 Comparisons with other work

As mentioned in Section 1, many frameworks consider extensions to Dung's argumentation systems that are capable of representing in one way or another the notion of strength of arguments or

[7] Choosing $\{Y\}$ would give symmetrical results.

attacks. In this section, we compare some of these approaches with ours. There are two main aspects to consider: the numerical nature of the networks itself (independently of any merging) and the methodology for merging.

In terms of numerical merging, the formalizm that most resembles ours is the one proposed in [12], which uses a weighted argumentation system. The idea is also based on the combination of all networks into a single augmented one in which attacks are assigned weights that correspond to ours under the credulous approach. However, the similarities stop there. In particular, there is no notion of sceptical support; no mechanism to consider the weights of arguments; and the concept of acceptance is based on the notion of 'various-strength defence': an argument X defends an argument Y against an argument Z, if the weight of the attack of X on Z is greater than the weight of the attack of Z on Y. This notion of defence is then used in the definition of *admissibility*. We believe that once we are prepared to associate strengths to the attacks based on the opinions of the agents, we should also be prepared to take into account the opinions of these agents about the arguments themselves during the merging process.

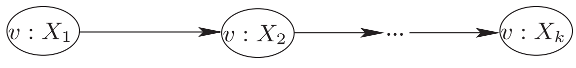
Bistarelli and Santini also propose a numerical approach to merging in [6], but as in the formalizm above, weights are given only to the attacks.

Amgoud and Kaci take a different approach to merging by considering the merging of knowledge bases whose underlying formalizm is a possibilistic logic [2]. This allows for the calculation of the *inconsistency degree* of a base, which in turn can be used to determine its 'plausible' consequences. The inconsistency degree of a knowledge base is an interesting concept and something we would like to investigate in the future to help us to provide a more robust definition of the threshold of acceptance of arguments.

Given an adequate meaning for the initial weights, the equational approach can be used for a single-weighted network independently of the merging process. Leaving considerations about the merging aside, it is possible to compare our formalizm with other weighted argumentation systems. One of the first systems of this type was proposed in [5]. In that formalizm, the weight of an argument is calculated by a so-called *categorizer* function, an example of which is the ***h*-categorizer**. For an argument X, this is defined as follows

$$h(X) = \begin{cases} 1, & \text{if } Att(Y) = \varnothing;\ \text{or} \\ 1/(1 + \sum\limits_{Y \in Att(X)} h(Y)), & \text{otherwise} \end{cases}$$

We can think of h as the function that calculates the equilibrium values of the arguments. Let us now analyse what happens with these values in a sequence of attacks like the one below. For comparison, we assume that all nodes have the same initial value v and that the intensity with which all attacks are carried out is also 1.



Given that X_1's initial value is 1 in the example above, we would have that $h(X_1) = 1$; $h(X_2) = 0.5$; $h(X_3) = 0.66$; and so forth. This obviously does not agree with Dung's semantics. Using the equational approach, we get that $V_e(X_1) = v$, $V_e(X_2) = v(1-v)$, $V_e(X_3) = v(1-(v(1-v)))$, ... If $v = 1$, then $V_e(X_1) = 1$, $V_e(X_2) = 0$, $V_e(X_3) = 1$, and so forth, agreeing with Dung's semantics as expected. If $v = 0$, then $V_e(A_i) = 0$ for all i. This is as expected, since in this case no arguments have any initial support and we have seen that $V_e(X) \leq V_0(X)$ (cf. Proposition 4.3). If $v = 0.5$, we get $V_e(X_1) = 0.5$, $V_e(X_2) = 0.25$, $V_e(X_3) = 0.375$, ...[8]

In [24], Leite and Martins proposed the so-called *social abstract argumentation frameworks* (SAAFs), which can be seen as an extension of Dung's abstract argumentation frameworks to allow the representation of information about votes to arguments. The motivation for a SAAF is to provide a means to calculate the result of the interaction between arguments using approval and disapproval ratings from users of news forums. The idea is that when a user sees an entry in a forum, she may approve it, disapprove it, or she may simply abstain from expressing her opinion about it. This leads to the notion of the initial support level for an argument, which can be seen as the initial value of an argument in the resulting weighted augmented network for a profile of argumentation systems (which we originally proposed in [21]). Their initial support levels for arguments is calculated differently from ours and there was then no notion of the strength of attack like in ours.[9] However, the aggregation of attacks in both cases is done as was originally formalized and extensively discussed by Gabbay *et al.* in [3]. In a follow-up paper [17], social abstract argumentation frameworks were then extended so that they could also deal with the notion of strength of attack which was formalized in the same way we originally defined it in [3, 21].

7.1 Comparisons with the merging methodology

In terms of the merging methodology itself Coste-Marquis *et al.*'s work [14] has a number of similarities with ours that are worth discussing. As in our case, their starting point is an augmented network containing all nodes in the profile to be merged. However, they consider the expansions to *each* of the networks being merged as an intermediate step. The expansions are called *partial argumentation frameworks* (PAFs). The process can be described as follows. Nodes unknown to an agent are simply added to her intermediate network. For every pair of nodes in an agent's own (non-expanded) network, the underlying assumption about the absence of an edge is that of a *non-attack* between the nodes. Therefore, an agent is never 'ignorant' about attacks within her own arguments. However, as the network is expanded to include all arguments in the profile, a decision has to be taken with respect to the attacks between the arguments the agent does not know about (as is the case in our own formalizm when we consider the attitude of an agent with respect to components of the augmented network) and she is left with the choice of either supporting the attack; or rejecting it; or neither (i.e. to assert her ignorance about it).

In a single agent's (smaller) network, the attack, non-attack and ignorance relations above can be reduced to two only if one assumes omniscience of the agent about attacks between her own arguments. This renders a traditional argumentation system as a special case of a PAF in which the ignorance relation is empty and the non-attack relation is the complement of the attack relation with respect to the Cartesian product of the set of arguments.

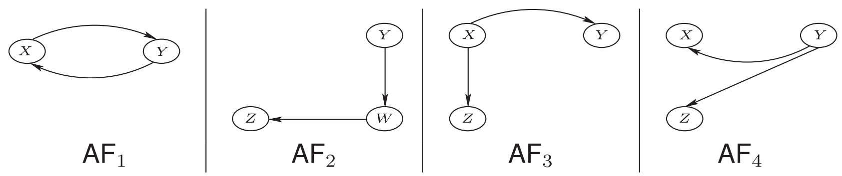
[8] We can think of an infinite sequence of this kind as a node with an attack on itself. In the limit $k \to \infty$, for $V_0(X_1) = 0.5$, $V_e(X_k) = 1/3$. Further considerations of this kind are left for a forthcoming paper on the numerical aspects of these networks.

[9] Our point of view about initial support is different. Support must be calculated in terms of the relative representation of the component in the community (even though one may act credulously with respect to a component, as we have shown). Simply using the ratio between the votes for an argument and the total number of votes it receives (i.e. the total of votes for and against it), may have undesired consequences. For instance, if an argument has a single vote supporting it and none against it, it gets 100% initial support, even within a possibly large community of voters. In our case, this can only happen if we choose to act credulously and the argument is unknown to all but one of the agents in the community.

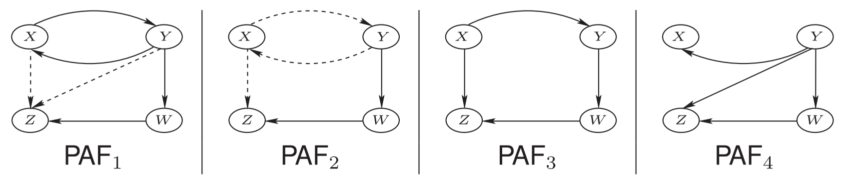
Any PAF can be ‘completed’ by moving all edges from the ignorance relation into one of the other two.[10]

Now given a network and a profile of networks, we want to consider the expansion containing all arguments in the profile, and hence we need to decide what to do with the attacks between arguments not present in the original network. Coste-Marquis *et al.* advocate for a ‘consensual’ expansion: if an attack that is unknown to an agent appears in the profile the agent will add the edge to her PAF in the attack relation only if all agents that know about the two arguments have that attack; otherwise the agent will add the edge to her ignorance relation.

For comparison, the idea is illustrated by the example below taken from [14]. In the example, the four networks AF_1–AF_4 are to be merged.



The first step is the expansion of each individual network, which results in the four partial argumentation frameworks PAF_1–PAF_4 below.



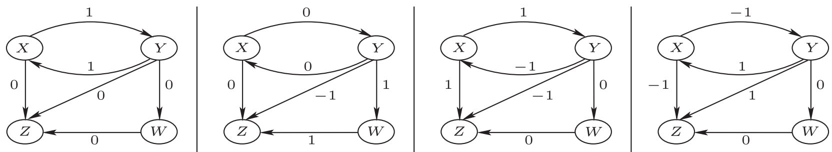
Notice that edges not appearing in any network are therefore (implicitly) added as non-attacks. Each network maintains its original edges in the expansion. Take for instance the network PAF_1. It contains the edges originally present in AF_1, but also the attacks from Y to W and W to Z, because they appear in AF_2, the only network that contains W (as a result, all expanded networks will contain all attacks to and from W—see our discussion later). An ignorance edge is added from X to Z, because even though X attacks Z in AF_3, this is not supported by AF_4, and hence there is no ‘consensus’ on the attack (the absence of the edge between X and Z in AF_4 is treated explicitly as a non-attack).

Once the networks of all agents are expanded, the merging is computed by selecting the ‘super-networks’ with minimal aggregated distance to all of the expanded networks. These super-networks are networks with all arguments, where the absence/presence of an edge between arguments is taken into account by a distance function. They consider the distances *sum*, max and *leximax*.

Our aggregation is done implicitly through the equations, so a direct comparison is not possible. However, we can interpret each individual expansion as the attitude of the agent with respect to the augmented network containing all components, since the expanded networks are what will be used in the aggregation. With this in mind, we can compare the two approaches. In our case, each agent would ‘vote’ for each component of the augmented network as depicted below (remember that the

[10]It is sufficient to have only the attack relation, since assuming the ignorance relation is empty, the non-attack relation can be defined as the complement of the attack relation as explained above.

absence of an edge in a PAF means a non-attack—effectively a vote against the attack, which is what a '-1' vote means in our methodology).



The attitudes are similar, except that in their formalizm an agent adopts the views of the others with respect to an attack between unknown arguments as long as no agent explicitly disagrees about the attack. Our approach is more cautious about this acceptance. For instance, in our case, a_1, a_3 and a_4 remain neutral with respect to the attacks from Y to W and from W to Z suggested by AF_2, since none of a_1, a_3 and a_4 know about W. We argue that this difference is crucial. In their methodology, if there are only two agents a_1 and a_2 who disagree about everything they know in common, they would still forcibly agree with each other on everything they do not know in a consensual expansion.[11] In this sense, consensual expansion is always biased towards reaching an agreement, but this may not always be appropriate. In contrast, the votes of the agents themselves are completely independent of how the other agents vote in our framework, which seems more natural. However, our credulous approach exhibits the same bias towards agreement at the time of merging even though other considerations can also potentially be taken at that stage based, e.g. on the overall level of disagreement in the community. We feel that if an agent is to lean towards the attitude of some other agents with respect to a component, at the very least the agent should lean more towards the agents whose opinions are more in agreement with her own.

Furthermore, notice that the merging of the PAFs can still produce more than one network as the result, since there could be more than one candidate aggregation with the minimal distance. And of course there is also the matter of choosing the winning arguments from the result.

We conclude this section by mentioning a few other frameworks that also deal with the problem of the merging of argumentation systems. The first of these was proposed in [27] and it was based on the idea of *labelling aggregation*.[12] The approach only considers different perceptions of a fixed argumentation system by different agents and uses a *plurality operator* defined in terms of how each argument is labelled by each agent. The approach is also non-numerical in the sense that no augmented network is generated as a result of the merging and nodes and attacks are not assigned any weights, as in our case.

Finally, two entirely different approaches were proposed by [7] and [8]. Both approaches make use of some extra information, such as agents' preferences and the argumentation 'context'. This makes the process of merging a more 'instantiated' approach where one can use a wealth of extra information available to decide how to merge. Our approach is completely abstract and relies only on the argumentation graphs and the agents' extensions.

[11] It is sufficient for the set of non unanimous arguments to be partitioned between disagreeing agents for them all to take the same view on those arguments.

[12] It is well known that a standard way to give an interpretation to an argumentation system is through a labelling semantics that identifies arguments as *in*, *out* or *undecided* [9, 10].

8 Conclusions and future work

In this article, we showed how to merge a profile of argumentation systems through the use of an augmented argumentation network provided with weights for the arguments and the attacks between them.

The weights of the augmented network are calculated based on how representative each component features in the profile and are independent of the local semantics of each network. We proposed two approaches for calculating these weights. In the *credulous approach* agents vote for components and the votes generate weights that take into account only the number of agents that know about a component. The *sceptical approach* considers the votes with respect to the total number of networks in the profile.

Weighted argumentation networks have been proposed before. Sometimes weights have been assigned to the arguments (e.g. as in [2, 3, 5, 11, 24]) and sometimes they have been assigned to the attacks (e.g. as in [3, 6, 16, 28]). In our approach, both arguments and attacks have weights and the network is seen as a generator for equations. The idea is to calculate equilibrium values for the arguments based on their initial support value within the profile and on how they interact with other arguments through the attack relation. The equilibrium values can be calculated by solving a system of equations generated by the augmented network, following [18]. Once calculated, the notion of acceptance can be defined in terms of reaching a minimum threshold value for acceptance. A strict definition of this threshold is the value 1. However, our framework is flexible in the sense that a particular application is free to partition the unit interval in different ways and give an interpretation for values occurring within each of these segments. For instance, one could associate 0 with rejection; 1 with acceptance and consider anything else in between as being undecided.

We can interpret the initial values in our augmented networks as coming from an extended form of approval voting in which voters can also express ignorance and rejection for some components. There are variations of this idea that are worth investigating, including giving different degrees of preference to the components depending on the expertise level of the agents supporting them as well as allowing the agents to change their positions in an iterative way with the aim of achieving a collective consensus (in the spirit of [7]).

Furthermore, there are interesting connections with several other areas of research. From the aggregation perspective, it is worth exploring similarities with other mechanisms for voting and formalizms for merging of knowledge bases as in [13, 20, 22, 23]. Some similarities also exist in the way the interactions are calculated with the approaches taken in the areas of network flows [1], belief propagation and Bayesian networks [25]. We hope to explore these issues in more detail in future work.

The merging of argumentation systems is an application that leads naturally to the employment of weights in a network. However, one need not restrict its use to such scenarios only. All that is required is a suitable interpretation for the weights; an adequate schema for generating the equations; and an interpretation for the equilibrium values. This article paves the way for a new type of research in argumentation networks not only because its approach is numerical, but also because it is an initial study of *vector evaluations*. We can see this work as a preliminary investigation on how to aggregate many-dimensional values of the components of a network and propagate the aggregated values through the network taking the network's interactions into account.

To realise the potential, consider the very well developed area of many-dimensional temporal logics. In these logics, a formula is evaluated at several indices. As a complex formula is evaluated in the model we move from one set of indices to another. The analogous movement in the case of argumentation is that of an attack. One can move from one node to the next evaluating and propagating the values on the way.

The equational approach can also be used in a more general context. For instance, if the underlying representation is itself based on a fuzzy or possibilistic logic, the initial weights can be obtained from the computations in the logic themselves, in the spirit of Prakken [26] or Amgoud-Kaci's 'force of an argument' [2]. The weights can then be subsequently combined taking the topology of the network into account as done here.

Acknowledgements

The authors would like to thank Sanjay Modgil and the reviewers for useful comments and suggestions to this article. D. M. G. was partially supported by the Israel Science Foundation Project 1321/10: Integrating logic and networks.

References

[1] R. K. Ahuja, T. L. Magnanti, and J. B. Orlin. *Network flows – Theory, Algorithms and Applications*. Prentice Hall, 1993.

[2] L. Amgoud and S. Kaci. An argumentation framework for merging conflicting knowledge bases. *Journal of Approximate Reasoning*, **45**, 321–340, 2007.

[3] H. Barringer, D. M. Gabbay, and J. Woods. Temporal dynamics of support and attack networks: from argumentation to zoology. In *Mechanizing Mathematical Reasoning*, D. Hutter and W. Stephan, eds, pp. 59–98. Vol. 2605 of *LNCS*, Springer, 2005.

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

[5] P. Besnard and A. Hunter. A logic-based theory of deductive arguments. *Artificial Intelligence*, **128**, 203–235, 2001.

[6] S. Bistarelli and F. Santini. A common computational framework for semiring-based argumentation systems. In *Proceedings of the 2010 conference on ECAI 2010: 19th European Conference on Artificial Intelligence*, H. Coelho, R. Studer, and M. Wooldridge, eds, pp. 131–136. Amsterdam, The Netherlands. IOS Press, 2010.

[7] E. Bonzon and N. Maudet. On the outcomes of multiparty persuasion. In *Proceedings of the 8th international conference on Argumentation in Multi-Agent Systems (ArgMAS'11)*, P. McBurney, S. Parsons, and I. Rahwan, eds, pp. 86–101. Berlin, Heidelberg. Springer-Verlag, 2012.

[8] G. Brewka and T. Eiter. Argumentation context systems: a framework for abstract group argumentation. In *IN: LPNMR 2009. LNCS*. Springer, 2009.

[9] M. Caminada. A labelling approach for ideal and stage semantics. *Argument and Computation*, **2**, 1–21, 2011.

[10] M. Caminada and D. M. Gabbay. A logical account of formal argumentation. *Studia Logica*, **93**, 109–145, 2009.

[11] C. Cayrol and M.-C. Lagasquie-Schiex. Graduality in argumentation. *Journal of Artificial Intelligence Research*, **23**, 245–297, 2005.

[12] C. Cayrol and M.-C. Lagasquie-Schiex. Merging argumentation systems with weighted argumentation systems: a preliminary study. *Technical Report RR2011-18FR*, IRIT, 2011.

[13] S. Chopra, A. Ghose, and T. Meyer. Social choice theory, belief merging, and strategy-proofness. *Information Fusion*, **7**, 61–79, 2006.

[14] S. Coste-Marquis, C. Devred, S. Konieczny, M.-C. Lagasquie-Schiex, and P. Marquis. On the merging of Dung's argumentation systems. *Artificial Intelligence*, **171**, 730–753, 2007.

[15] P. M. Dung. On the acceptability of arguments and its fundamental role in nonmonotonic reasoning, logic programming and n-person games. *Artificial Intelligence*, **77**, 321–357, 1995.

[16] P. E. Dunne, A. Hunter, P. McBurney, S. Parsons, and M. Wooldridge. Weighted argument systems: basic definitions, algorithms, and complexity results. *Artificial Intelligence*, **175**, 457–486, 2011.

[17] S. Eğilmez, J. Leite, and J. Martins. Extending social abstract argumentation with votes on attacks. In *Proceedings of the 2nd International Workshop on Theory and Applications of Formal Argumentation (TAFA'13)*, E. Black, S. Modgil, and N. Oren, eds. (to appear 2014).

[18] D. M. Gabbay. Introducing equational semantics for argumentation networks. In *Symbolic and Quantitative Approaches to Reasoning with Uncertainty*, W. Liu, ed., Vol. 6717 of *Lecture Notes in Computer Science*, pp. 19–35. Springer, 2011.

[19] D. M. Gabbay. An equational approach to argumentation networks. *Argumentation and Computation*, **3**, 87–142, 2012.

[20] D. M. Gabbay, G. Pigozzi, and O. Rodrigues. Belief revision, belief merging and voting. In *Proceedings of the Seventh Conference on Logic and the Foundations of Games and Decision Theory (LOFT06)*, G. Bonanno, W. van der Hoek, and M. Woolridge, eds, pp. 71–78. University of Liverpool, 2006.

[21] D. M. Gabbay and O. Rodrigues. A numerical approach to the merging of argumentation networks. In *Proceedings of CLIMA XIII*, M. Fisher, L. van der Torre, M. Dastani, and G. Governatori, eds, pp. 195–212. Springer-Verlag, 2012.

[22] S. Konieczny and R. Pino-Pérez. On the logic of merging. In *Proceedings of KR'98*, T. Cohn and L. Schubert, eds, pp. 488–498. Morgan Kaufmann, 1998.

[23] S. Konieczny and R. Pino-Pérez. Logic based merging. *Journal of Philosophical Logic*, **40**, 239–270, 2011.

[24] J. Leite and J. Martins. Social abstract argumentation. In *Proceedings of the 22nd International Joint Conference on Artificial Intelligence*, T. Walsh, ed., pp. 2287–2292. AAAI Press/International Joint Conferences on Artificial Intelligence, 2011.

[25] J. Pearl. Fusion, propagation, and structuring in belief networks. *Artificial Intelligence*, **29**, 241–288, 1986.

[26] H. Prakken. An abstract framework for argumentation with structured arguments. *Argument and Computation*, **1**, 93–124, 2010.

[27] I. Rahwan and F. Tohmé. Collective argument evaluation as judgement aggregation. In *Proceedings of the 9th International Conference on Autonomous Agents and Multiagent Systems: volume 1 - Volume 1 (AAMAS '10)*, M. Luck and S. Sen, eds, pp. 417–424. Richland, SC. International Foundation for Autonomous Agents and Multiagent Systems, 2010.

[28] J. Wang, G. Luo, and B. Wang. Argumentation framework with weighted argument structure. In *10th IEEE International Conference on Cognitive Informatics Cognitive Computing (ICCI*CC)*, pp. 385–391. IEEE, 2011.

Received 19 January 2013