

# A Numerical Approach to the Merging of Argumentation Networks

Dov Gabbay<sup>1</sup> and Odinaldo Rodrigues<sup>2</sup>

<sup>1</sup> Bar Ilan University, Israel; Department of Informatics,  
King's College London; and University of Luxembourg  
dov.gabbay@kcl.ac.uk

<sup>2</sup> Department of Informatics, King's College London  
odinaldo.rodrigues@kcl.ac.uk

**Abstract.** In this paper, we propose a numerical approach to the problem of merging of argumentation networks. The idea is to consider an augmented network containing the arguments and attacks of all networks to be merged and then associate a weight to each of its components based on how they are perceived by the agents associated with the local networks. The combined weighted network is then used to define a system of equations from which the overall strength of the arguments is calculated.

## 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 [12]. 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 paper 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  $R_i \subseteq S_i^2$ . 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.

At first, one may think that the merging process can be done at the *meta level*, i.e., by considering only the winning arguments in each local system. However, this not only will sometimes produce unintuitive results [11], but will also fail to simultaneously satisfy well-known social choice properties [22]. The reasons have to do with the fact that attacks known only locally are not represented by the local extensions of winning arguments, but they may well be relevant during the collective decision as a whole. If we want to take both the local decisions and the local topologies of the various systems into account, we need a framework that can take all this information into account.

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 locally, 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 get a merged combined numerically weighted argumentation system. This is a mismatch. So we need to explain how we understand the numerical weights and then extract/project from the merged system a set of winning arguments. 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* [8]. This is akin to finding the extensions in a traditional network. However, our work has two important differences: 1) we allow both arguments and attacks to have weights; and 2) we calculate the equilibrium values using the equational approach of [14,15]: we see the augmented system as a generator of numerical equations whose solutions correspond to the equilibrium values.

Argumentation systems in which weights are associated to arguments have been studied before. One of the first approaches was proposed in [4] where the weight of an argument is used to express its relative strength for a particular audience. Besnard and Hunter proposed a *categoriser* function that assigned a value to a tree of arguments [5]; Cayrol and Lagasquie-Schiex introduced the concept of *graduality* in the valuation of arguments in [8]; and other examples of systems using weights in one form or another include [3,13,2,6,23,19]. 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 paper 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 some examples in Section 5. Some comparisons with related work are done in Section 6 and we finish with some discussions and conclusions in Section 7.

## 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.,  $\forall X, Y \in E$ , it is not the case that  $(X, Y) \in R$ ; and that all arguments  $X \in E$  are *acceptable*, i.e.,  $\forall Y \in S$ , if  $(Y, X) \in R$ , then  $\exists Z \in E$  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 [7, Theorem 2]. Such labelling is called a *Caminada labelling* [7, 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 normalised result of a vote put to a community of agents; or as the result of interactions between the arguments in a network (as in [8]). 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 [9]). 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 formalisms 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 [11].

Because of these considerations, 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.

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 Lagasque-Schieux’s terminology [8]. However, in our case we want arguments to be *weakened* in proportion to the strength of the attacks and the intensity with which these attacks are carried out. Ideally, we want to find *equilibrium* values for all arguments.

One good option to calculate these values is by using the *equational approach* proposed in [14,15] which sees a numerical network as a generator of equations. Provided the equations respect the meaning of the weights of the arguments and attacks an “evaluation” of the network can be done according to the solutions found for the system of equations. For an argument  $X$ , the equilibrium value 1 means definitely “in”; 0 means definitely “out”; and any other value inbetween means how close to in (or out)  $X$  is. We may even decide on an appropriate threshold value for the acceptance of arguments.

An example of how such equations can be generated is given by the schema  $Eq_{\max}$  below. The symbol  $V_e(X)$  will be used to denote 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 equation

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

Note 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 Fig. 1. generates the following system of equations:

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

If we set the threshold for acceptance of arguments at the value 1, we get that only the arguments  $X$  and  $Z$  are accepted as traditionally expected.

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 [15]. However, note that  $Eq_{\max}$  does not take into account a node’s initial value or the intensity with which the attacks to it have been carried out. We will consider a more sophisticated equation schema to take these into account in Section 4.

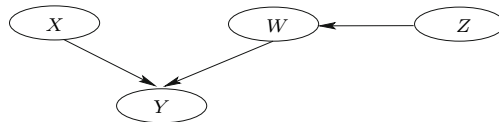
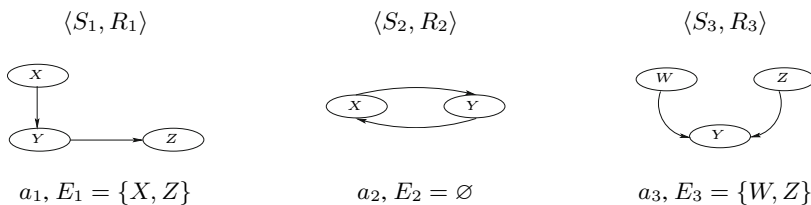


Fig. 1. A simple argumentation system

### 3 Merging Argumentation Networks

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

As discussed in Section 1, we start by associating each network with an agent who “votes” for its components. Obviously, in any interesting scenario, the networks being merged are distinct. Consider the networks in Fig. 2 and the chosen extensions of each network containing its winning arguments.



**Fig. 2.** Argumentation networks of three different agents

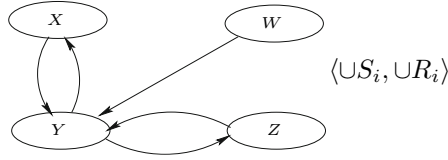
We immediately notice that the three agents have different sets of arguments, and even in the case where the arguments coincide, the agents may disagree with respect to the attack relationships between them. For instance, argument  $W$  is only known to agent  $a_3$ , and in her network,  $Z$  attacks  $Y$ , whereas in the network of agent  $a_1$ ,  $Y$  attacks  $Z$ .

There are many reasons why agents may have different argument systems. They may use different knowledge bases; they may have different deductive capabilities; they may use different inference systems; etc. These may also generate disagreements with respect to the direction of the attacks between arguments which are arguably akin to the existence of cycles in a single network. In fact, individual acyclic networks when combined into a single network may well end up containing cycles.

A simple way of harmonising the differences is to consider expansions to the networks. Unlike in [11], we do not expand each network individually, but rather we consider the single augmented network that includes the components of all other networks.

However, the augmented network alone is not sufficient to represent the community, because some components appear in more networks than others and we would like to reflect that by using weights. We first introduce the notion of a profile of (traditional) networks and then we define the notion of an augmented network with weights for a profile.

**Definition 1.** A profile of argumentation systems is a tuple  $P = \langle AN_1, \dots, AN_k \rangle$  where each  $AN_i = \langle S_i, R_i \rangle$  is an argumentation system. We assume each agent  $a_i$  has a procedure  $w_i$  for selecting a subset of  $S_i$  representing the winning arguments in  $S_i$  according to  $a_i$ 's local semantics and we use  $E_i$  to denote this set, i.e.,  $w_i(S_i) = E_i$ .



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

**Definition 2.** Let  $P = \langle AN_1, \dots, AN_k \rangle$  be a profile of argumentation systems and let  $AN_i = \langle S_i, R_i \rangle$ . The weighted augmented network for  $P$  is a tuple  $AWN_P = \langle S, R, V_0, \xi \rangle$  where

- $S = \cup_i S_i$  and  $R = \cup_i R_i$
- $V_0 : S \rightarrow [0, 1]$  represents the initial level of support for an argument  $X \in S$  within  $P$ , as to be calculated/voted from all the  $AN_i$
- $\xi : R \rightarrow [0, 1]$  represents the intensity of an attack  $(X, Y) \in R$  within  $P$ , as to be calculated/voted from all the  $AN_i$

Now we need to decide on a policy for representing each agent’s perception of the arguments and attacks in  $AWN_P$  depending on the agent’s original network. We shall see that these will later be used to define  $V_0$  and  $\xi$ . For simplicity, we will refer generally to the arguments of these functions, i.e., arguments and attacks respectively, as the “components” of the network.

In agreement with [11] 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.** Let  $P$  be a profile. 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.<sup>1</sup>

$v_i(X)$	$v_i((X, Y))$
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 it; otherwise  $a_i$  will vote with 1 or -1 depending on whether  $X$  is amongst the winning arguments of  $S_i$ . The case of an attack from  $X$  to  $Y$  is similar but an attack may not exist because one or both arguments are not known. Hence, the agent  $a_i$  will vote with 0 if at least one of  $X$  and  $Y$  is not known (in which case a judicious decision about the

<sup>1</sup> To simplify notation we use the same function symbol  $v_i$  for nodes and edges.

attack is not possible). Otherwise, if both  $X$  and  $Y$  are known, the agent will vote with  $-1$  if  $(X, Y) \notin R_i$  and with  $1$  if  $(X, Y) \in R_i$ . Note that the vote  $1$  depends only on the existence of the attack in the agent's local network. Even if  $Y \in E_i$  and  $X \notin E_i$ , the agent  $a_i$  must still vote with  $1$  if  $(X, Y) \in R_i$ , since she knows about it. The agent's choice for  $Y$  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  $X$  and  $Y$ .

The above voting strategy only requires that there is a local semantics for deciding the winning arguments in each network and does not make any assumptions on what it should be. In fact, the group as a whole may have several different local semantics.

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, normalise 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.

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 may take two different approaches when considering the overall level of support for a component:

- 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  $P$

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  $\xi$  below. Whenever the distinction is not important we will simply omit the superscripts.

**Definition 4.** Let  $P = \langle AN_1, \dots, AN_k \rangle$  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  $v^-(c) = |\{i \mid v_i(c) = -1\}|$ . We define

$$\left. \begin{aligned} V_0^+(X) &= \frac{v^+(X)}{v^+(X)+v^-(X)} \\ \xi^+((X, Y)) &= \frac{v^+((X, Y))}{v^+((X, Y))+v^-((X, Y))} \end{aligned} \right| \begin{aligned} V_0^-(X) &= \frac{v^+(X)}{k} \\ \xi^-((X, Y)) &= \frac{v^+((X, Y))}{k} \end{aligned}$$

Note 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 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 Fig. 2 we get the initial weights shown in Fig. 4 for the components of the augmented network under each approach. Given these weights, we then need to calculate equilibrium values for the nodes (this will be done in Section 4).

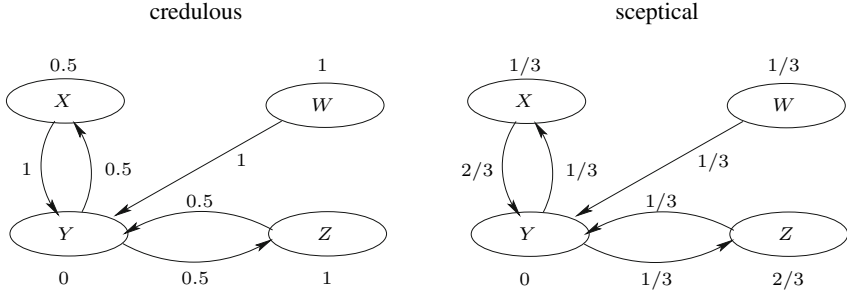


Fig. 4. Merged networks of Fig. 2 under the credulous and sceptical approaches

Note in Fig. 4 that under either approach, the initial weight of argument  $Y$  is 0. This is because  $Y$  is not a winning argument in any of the initial networks. Analogously, the initial value of  $Z$  is 1 only in the credulous approach. This is because  $Z$  is a winning argument in every network in which it is known, but it is not known in every network. Similarly,  $W$ 's initial weight is 1 under the credulous approach, but  $1/3$  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 follow the same pattern.

Generally speaking, we have the following.

**Proposition 1.** Let  $P = \langle AN_1, \dots, AN_k \rangle$  be a profile of argumentation systems where each  $AN_i = \langle S_i, R_i \rangle$  and  $E_i$  identifies the winning arguments in  $S_i$ . Let  $AWN_P = \langle S, R, V, \xi \rangle$  be the weighted augmented network for  $P$  according to Definition 2. 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  $X \notin \cup_i E_i$ , then  $V_0^+(X) = V_0^-(X) = 0$
3. if  $X \in E_i$  for all  $i$  such that  $X \in S_i$ , then  $V_0^+(X) = 1$

*Proof.* 1. and 2. follow directly from Definitions 3 and 4. For 3., 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 2.** For all attacks  $(X, Y) \in R$ .

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

*Proof.* These follow directly from Definitions 3 and 4.

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

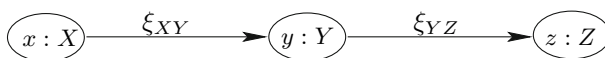


## 4 Equilibrium Values in a Weighted Augmented Network

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 any usual argumentation system, arguments may be attacked by any number of arguments. Since we work with numerical values, we want to aggregate the strength of these attacks somehow in order to *weaken* the weight of the attacked node. The strength of an attack itself depends on the strength of the attacking node and the intensity with which the attack is carried out. The attacking nodes may be themselves attacked, so we need to perform the aggregation systematically. We start by analysing the effect of attacks in general.

Consider the network in Fig. 5, in which  $x$ ,  $y$  and  $z$  are the initial weights of the arguments  $X$ ,  $Y$  and  $Z$ , respectively. Let us for a moment ignore these initial weights.



**Fig. 5.** A typical weighted argument network

If we want to mimic the standard behaviour of the attacks in an argumentation system [12], 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 defeat with the value 0. For us, “to be defeated” means to end up with equilibrium value 0 and “to persist” means to end up with a value 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’ initial weights (which are all in the unit interval  $U$ ) and the intensity with which the attacks between them are carried out. Hence, our objective is to calculate the values  $V_e(X)$ ,  $V_e(Y)$  and  $V_e(Z)$ , based on  $x$ ,  $y$ ,  $z$ ,  $\xi_{XY}$  and  $\xi_{YZ}$ . 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*  $\xi_{XY}$  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  $\xi_{YZ}$  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.<sup>2</sup>

If we give initial value 1 to all arguments and consider all attacks being transmitted with full intensity, then since  $X$  has initial value 1 and it is not attacked by any other argument, its equilibrium value becomes 1. It then attacks  $Y$  with full intensity (i.e.,

<sup>2</sup> This and some other related issues will be explored in more detail in a forthcoming paper.

$\xi_{XY} = 1$ ), which means that the initial value of  $Y$ ,  $y = 1$ , is weakened by 1 and its equilibrium value becomes 0. Effectively, this annihilates the 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).

We stress that, in general, we are free to decide on the minimum value we require for considering an argument as 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: one with the values which we consider accepted and the other with the ones we consider rejected.

If we want to think in terms of the effect of the attacks on an argument  $X$ , our problem is to determine a factor  $0 \leq \pi(X) \leq 1$  representing the combined strength of these attacks. The equilibrium value for  $X$  can be calculated by multiplying  $X$ 's initial value by this factor, i.e.,  $V_e(X) = V_0(X) \cdot \pi(X)$ .

When there are multiple attacks to an argument  $X$ ,  $\pi$  must *aggregate* the value of these attacks. In order to remain close to the standard argumentation semantics, we want  $\pi$  to satisfy at least the three conditions below.

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

(SSC1) says that if all arguments attacking  $X$  are fully defeated or transmitted with null intensity, then  $X$  retains 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  $X$  is 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.

So the idea is that the stronger an attack is, the closer the attack gets to the value 1 and hence the closer we want  $\pi$  to get to 0 so that the equilibrium value of the attacked argument decreases proportionally (since its initial value is multiplied by  $\pi$ ). In the case of a single attack of strength  $u$  to node  $X$  with transmission factor  $\kappa$ , one possibility is to make  $\pi(X) = 1 - \kappa u$ . In the network of Fig. 5 above, this would make  $\pi(Y) = 1 - \xi((Y, X)) V_e(X)$  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* [5] is an example of a function satisfying (SSC1)–(SSC3) (more on this in Section 6).

But 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 5.** Let  $P = \langle AN_1, \dots, AN_k \rangle$  be a profile of argumentation systems and  $AWN_P = \langle S, R, V, \xi \rangle$  the weighted augmented network for  $P$  as defined before. The equation for the equilibrium value of an argument  $X \in S$  is defined as

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

One can choose  $V_0^+$  and  $\xi^+$  or  $V_0^-$  and  $\xi^-$  depending on whether a credulous or sceptical approach is desired (this will be explored further in Section 5). Note 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 the equilibrium value of the attacked argument would be 0.

$Eq_{max}$  decreases the initial support value of an argument according to the value of the strongest attack.  $Eq_{inv}$  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 decrease the argument’s overall credibility.

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

**Proposition 3.**  $\pi$  satisfies (SSC1)–(SSC3).

*Proof.* If  $\max_{Y \in Att(X)} \{\xi((Y, X))V_e(Y)\} = 0$ , then by Definition 5,  $\prod_{Y_i \in Att(X)} (1 - \xi((Y_i, X))V_e(Y_i)) = 1$ . Therefore, (SSC1) is satisfied. If  $\max_{Y \in Att(X)} \{\xi((Y, X))V_e(Y)\} = 1$ , then by Definition 5, for some  $Y' \in Att(X)$ ,  $1 - \xi((Y', X))V_e(Y') = 0$ , and then  $\prod_{Y_i \in Att(X)} (1 - \xi((Y_i, X))V_e(Y_i)) = 0$ . 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 written simply as

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

which is equivalent to

$$1 - \gamma_{Y \in Att(X)} V_e(Y) \tag{1}$$

where  $a \gamma b = a + b - a.b$  and for  $\Delta = \{a_1, \dots, a_k\}$ ,  $\gamma \Delta = ((a_1 \gamma a_2) \gamma \dots \gamma a_k)$ . The expression in (1) is the complement of the probabilistic sum t-conorm used by Leite and Martins in [19]. 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 support value.

**Proposition 4.** 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 4, for all arguments  $Y$ ,  $0 \leq V_0(Y) \leq 1$ . By Definition 5,  $0 \leq \pi(X) \leq 1$  and hence  $V_e(X) \leq V_0(X)$ .

**Proposition 5 (Unanimity of acceptance).** Let  $P = \langle AN_1, \dots, AN_k \rangle$  be a profile of argumentation systems where each  $AN_i = \langle S_i, R_i \rangle$  and let  $AWN_P = \langle S, R, V, \xi \rangle$  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 1, 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 4,  $V_0^+(Y) = V_0^-(Y) = 0$  and by Proposition 4,  $V_e(Y) = 0$ . It follows that  $\pi(X) = 1$  and hence  $V_e(X) = 1$ .

If each  $E_i$  is conflict-free and  $V_e(X) = 1$ , then  $X \in \cap_i E_i$  only if  $V_e$  is calculated under the sceptical approach. The credulous approach is more lenient, because it gives initial support value 1 to an argument as long as it wins in every argumentation system in which it is known and this may be sufficient to make the argument's equilibrium value 1 too as long as the sets of winning arguments are conflict-free (see Example 2. in Section 5). It is worth emphasizing that the flip side of this is that attacks are treated in the same way and as a consequence the value of an argument may also decrease as a result. This is illustrated in Example 1. of Section 5, 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  $X$  which attacks  $Y$ .

### 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. 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, but setting this as the minimum acceptance level could prove too strict even under the credulous approach. One could base the minimum acceptance level on the maximum or minimum of the equilibrium values or simply take the average and accept the arguments whose equilibrium values are above it. Another possibility is to accept arguments with equilibrium value above 0.5. Although this may seem arbitrary, in fact values above 0.5 can be associated with the concept of *majority*, because acceptance of an argument by clear majority produces a initial support value strictly greater than 0.5.

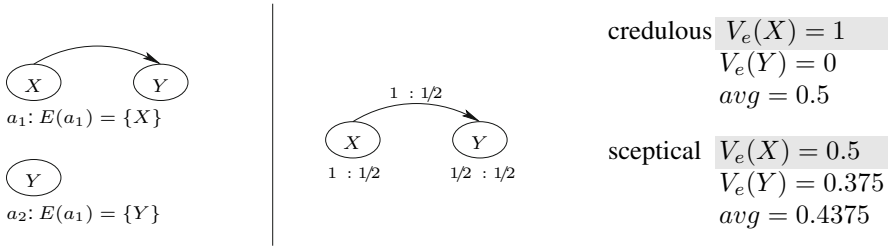
However, in the examples that follow, we will simply use the average of the equilibrium values calculated. In more realistic scenarios an appropriate threshold value can be determined through a more sophisticated analysis of the networks in the profile in a similar way to how it is done in [2] (which 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.

## 5 Worked Examples

We now illustrate our technique with a few examples. In each example, we show the networks to be merged on the left and the augmented (merged) network in the middle with its components annotated with the initial weights in the form “credulous : sceptical”. The equilibrium values obtained are given on the right and the accepted arguments

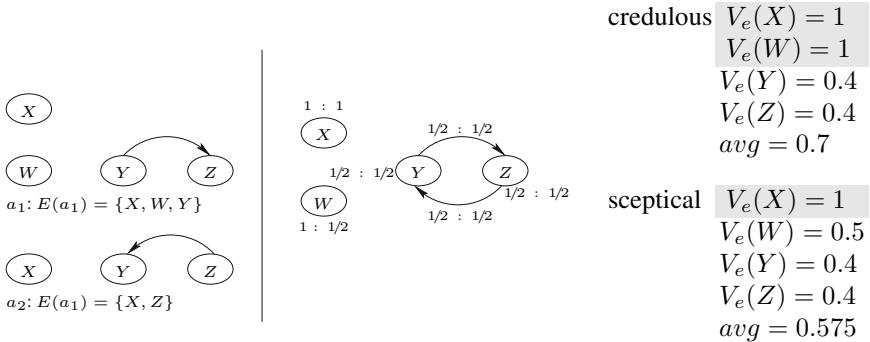
indicated with a shadowed box. Due to space limitations we cannot include the equations here, but they can be easily obtained using Definition 5.

1.



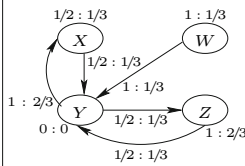
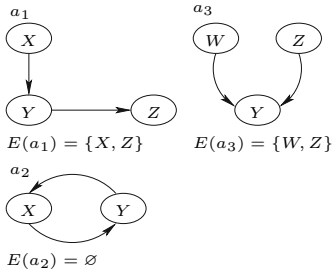
Under the credulous approach  $V_0(X) = 1$  and hence  $V_e(X) = 1$ . Its attack on  $Y$  is transmitted with full intensity.  $V_0(Y) = 0.5$ . Therefore,  $V_e(Y) = 0.5 \times (1 - 1) = 0$ . Under the sceptical approach  $V_0(X) = 0.5$  and hence  $V_e(X) = 0.5$ . Its attack on  $Y$  is transmitted with intensity 0.5. Therefore,  $V_e(Y) = 0.5 \times (1 - 0.5 \times 0.5) = 0.375$ . 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.

2.



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. In spite of there not being any attacks on  $W$ , it is only known by agent  $a_2$ . Under the credulous approach  $V_e(W) = 1$ , but under the sceptical approach  $V_e(W) = 0.5$ , since it is accepted by only half of the community.  $Y$  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 0.5 to 0.4.  $X$  and  $W$  have equilibrium values above the average under the credulous approach and are hence accepted, but under the sceptical approach only  $X$  is accepted.

3.



credulous	$V_e(X) = 0.5$
	$V_e(W) = 1$
	$V_e(Y) = 0$
	$V_e(Z) = 1$
	$avg = 0.625$
sceptical	$V_e(X) = 1/3$
	$V_e(W) = 1/3$
	$V_e(Y) = 0$
	$V_e(Z) = 2/3$
	$avg = 1/3$

This is the example appearing in Fig. 2.  $Y$  does not feature in any of the agents’ winning arguments. Its initial support value is 0 and hence its equilibrium value is also 0. This leaves  $X$ ’s initial support values unchanged. Under the credulous approach both  $W$  and  $Z$  get value 1. 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 above the average under the credulous approach and are hence accepted, but only  $Z$  is accepted under the sceptical one (note that it is the only argument accepted by the majority of the agents).

## 6 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 the strength of arguments or attacks. In this section, we discuss the relationship between some of these approaches and ours.

In terms of numerical *merging*, the formalism that most resembles ours is the one proposed in [9], 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 associate weights to arguments; and the concept of acceptance is based on the notion of “various-strength” defence: an argument  $X$  defends an argument  $Y$  against 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 is then used in the definition of admissibility. We believe that once we are prepared to evaluate the strength of the attacks based on the opinions of the agents, we should also be prepared to take these opinions into account in the evaluation of the support for the arguments as well.

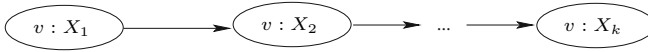
Bistarelli and Santini also consider a numerical approach but, as in the above, their formalism only assigns weights to attacks [6]. Amgoud and Kaci take a different approach to merging by considering the merging of knowledge bases whose underlying formalism is a possibilistic logic [2]. This allows for the calculation of the *inconsistency*

degree of a base, which in turn can be used to determine its “plausible” consequences. This notion of a degree of inconsistency is something we would like to investigate further to provide a more robust definition for the threshold of acceptance of arguments.

Other formalisms for merging not based on numerical approaches include [11,22].

If we have an adequate meaning for the initial weights, we can use the equational approach for a single weighted network independently of the merging process. Leaving considerations about merging aside, it is possible to compare our formalism with other weighted argumentation systems. In [5], the equilibrium value of an argument is calculated by a so-called *categoriser* function, an example of which is the ***h*-categoriser**  $h$ , which defined for an argument  $X$  as  $h(X) = 1$ , if  $Att(Y) = \emptyset$ ; and  $h(X) = 1/(1 + \sum_{Y \in Att(X)} h(Y))$ , otherwise.

Let us now analyse what happens in a sequence of attacks like the one below. For comparison, we assume that all nodes have the same initial values  $v$  and consider the intensity with which all attacks are carried out to be 1.



Assuming initial value  $v = 1$  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. In fact, for all nodes  $X$ ,  $V_e(X) \leq V(X)$ , since  $\pi(X) = \prod_{Y \in Att(X)} (1 - \xi((Y, X))V_e(Y)) \leq 1$ . If  $v = 0.5$ , we get  $V_e(X_1) = 0.5$ ,  $V_e(X_2) = 0.25$ ,  $V_e(X_3) = 0.375$ , ...<sup>3</sup>

Leite and Martins proposed *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 of these networks 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 argument, she may approve it, disapprove it, or simply abstain from expressing an opinion. The ‘weights’ associated with the arguments in this case can also be seen as being generated by how the agents perceive the arguments. However, the initial support level for an argument is calculated differently in their formalism and there is no notion of strength of attack, even though, as in our case attacks are aggregated using the probabilistic sum t-conorm.

## 7 Conclusions and Future Work

In this paper, we showed how a profile of argumentation systems can be merged through the use of an augmented argumentation network provided with weights for the arguments and the attacks between them. The initial weights are calculated based on how

<sup>3</sup> We can think of an infinite sequence of this kind as a node with an attack on itself. In the limit  $k \rightarrow \infty$ , for  $V_0(X_1) = 0.5$ ,  $V_e(X_k) = \frac{1}{3}$ .

representative each component features in the profile and are independent of the local semantics of each network. We proposed credulous and sceptical approaches for calculating the weights. The credulous approach calculates the overall representation based on how many agents “know” about a component whereas the sceptical approach looks at 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 [2,3,5,8,19]) and sometimes they have been assigned to the attacks (e.g., as in [3,6,13,23]). 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 the interactions with other arguments through the attack relation. These values can be calculated by solving a system of equations generated by the augmented network, following [14]. Once calculated, the notion of acceptance can be defined in terms of a threshold value for the equilibrium values, for which a strict interpretation is the value 1. However, the framework is flexible in the sense that a particular application is free to associate segments of the unit interval in different ways. For instance, one could associate 0 with rejection; 1 with acceptance and consider anything else inbetween as undecided.

We can see the initial values in our augmented network as coming from an extended form of approval voting in which voters can also express ignorance and rejection for some components. There are variations on this idea that are worth investigating, including giving varying degrees of preference to the components depending on the expertise level of the agents supporting them. Furthermore, there are interesting connections with several other areas of research. From the aggregation perspective, it is worth exploring similarities with other procedures for voting and formalisms for merging of knowledge bases as in [10,16,17,18]. Some similarities also appear in the spirit of the calculation of the interactions with the areas of network flows [1], belief propagation and Bayesian networks [20]. We will 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 paper 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 them through the network taking its attack relation 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 another evaluating and propagating the values.

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 [21] or Amgoud-Kaci's "force of an argument" [2]. The weights can then subsequently be combined taking the topology of the network into account as done here.

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

## References

1. Ahuja, R.K., Magnanti, T.L., Orlin, J.B.: Network flows – Theory, algorithms and applications. Prentice-Hall (1993)
2. Amgoud, L., Kaci, S.: An argumentation framework for merging conflicting knowledge bases. *Journal of Approximate Reasoning* 45, 321–340 (2007)
3. Barringer, H., Gabbay, D.M., Woods, J.: Temporal Dynamics of Support and Attack Networks: From Argumentation to Zoology. In: Hutter, D., Stephan, W. (eds.) *Mechanizing Mathematical Reasoning*. LNCS (LNAI), vol. 2605, pp. 59–98. Springer, Heidelberg (2005)
4. Bench-Capon, T.J.M.: Persuasion in practical argument using value-based argumentation frameworks. *Journal of Logic and Computation* 13(3), 429–448 (2003)
5. Besnard, P., Hunter, A.: A logic-based theory of deductive arguments. *Artificial Intelligence* 128(1-2), 203–235 (2001)
6. Bistarelli, S., Santini, F.: A common computational framework for semiring-based argumentation systems. In: *Proceedings of the 2010 conference on ECAI 2010: 19th European Conference on Artificial Intelligence*, pp. 131–136. IOS Press, Amsterdam (2010)
7. Caminada, M., Gabbay, D.M.: A logical account of formal argumentation. *Studia Logica* 93(2-3), 109–145 (2009)
8. Cayrol, C., Lagasque-Schiex, M.-C.: Graduality in argumentation. *Journal of Artificial Intelligence Research* 23, 245–297 (2005)
9. Cayrol, C., Lagasque-Schiex, M.-C.: Merging argumentation systems with weighted argumentation systems: a preliminary study. Technical Report RR 2011-18 FR, IRIT (2011)
10. Chopra, S., Ghose, A., Meyer, T.: Social choice theory, belief merging, and strategy-proofness. *Information Fusion* 7(1), 61–79 (2006)
11. Coste-Marquis, S., Devred, C., Konieczny, S., Lagasque-Schiex, M.-C., Marquis, P.: On the merging of Dung's argumentation systems. *Artificial Intelligence* 171, 730–753 (2007)
12. Dung, P.M.: On the acceptability of arguments and its fundamental role in nonmonotonic reasoning, logic programming and n-person games. *Artificial Intelligence* 77, 321–357 (1995)
13. Dunne, P.E., Hunter, A., McBurney, P., Parsons, S., Wooldridge, M.: Weighted argument systems: Basic definitions, algorithms, and complexity results. *Artificial Intelligence* 175(2), 457–486 (2011)
14. Gabbay, D.M.: Introducing Equational Semantics for Argumentation Networks. In: Liu, W. (ed.) *ECSQARU 2011*. LNCS, vol. 6717, pp. 19–35. Springer, Heidelberg (2011)
15. Gabbay, D.M.: An equational approach to argumentation networks. *Argumentation and Computation* 3 (2012)
16. Gabbay, D.M., Pigozzi, G., Rodrigues, O.: Belief revision, belief merging and voting. In: *Proceedings of the Seventh Conference on Logic and the Foundations of Games and Decision Theory (LOFT 2006)*, pp. 71–78. University of Liverpool (2006)
17. Konieczny, S., Pino-Pérez, R.: On the logic of merging. In: *Proceedings of KR 1998*, pp. 488–498. Morgan Kaufmann (1998)

18. Konieczny, S., Pino-Pérez, R.: Logic based merging. *Journal of Philosophical Logic* 40(2), 239–270 (2011)
19. Leite, J., Martins, J.: Social abstract argumentation. In: *Proceedings of the 22nd International Joint Conference on Artificial Intelligence*, pp. 2287–2292 (2011)
20. Pearl, J.: Fusion, propagation, and structuring in belief networks. *Artificial Intelligence* 29(3), 241–288 (1986)
21. Prakken, H.: An abstract framework for argumentation with structured arguments. *Argument and Computation* 1, 93–124 (2010)
22. Rahwan, I., Tohmé, F.: Collective argument evaluation as judgement aggregation. In: *Proceedings of the 9th International Conference on Autonomous Agents and Multiagent Systems, AAMAS 2010*, vol. 1, pp. 417–424. International Foundation for Autonomous Agents and Multiagent Systems, Richland (2010)
23. Wang, J., Luo, G., Wang, B.: Argumentation framework with weighted argument structure. In: *10th IEEE International Conference on Cognitive Informatics Cognitive Computing (ICCI\*CC)*, pp. 385–391 (2011)