

# Degrees of “in”, “out” and “undecided” in Argumentation Networks

Dov M GABBAY<sup>a</sup> and Odinaldo RODRIGUES<sup>b,1</sup>

<sup>a</sup>King’s College London, Department of Informatics, Ashkelon Academic College and Bar Ilan University, Israel and University of Luxembourg, Luxembourg

<sup>b</sup>King’s College London, Department of Informatics

**Abstract.** The traditional 3-valued semantics of an argumentation framework  $(\mathcal{A}, \mathcal{R})$  identifies arguments that are “in”, “out” and “undecided”. Yet, it has long been recognised by the community that some elements can be at different degrees in each of these categories [1,2,3]. For example, Dung’s semantics can only classify some elements as “out”, but cannot reflect how much “out” they really are or if elements are “in” are they as much “in” as elements which are not attacked at all?

In this paper we shall use a numerical approach to give a measure of “in”, “out” and “undecided” to the nodes of a network. We shall devise equations which allow for solutions that reflect these distinctions.

**Keywords.** Numerical argumentation, degrees of acceptance, numerical methods

## 1. Introduction and Preliminary Discussion

Consider the situation depicted in the argumentation network of Figure 1 (L) where we have the set of arguments  $\mathcal{A} = \{A, B, C, X, Y, W, Z\}$  with a relative complex geometrical configuration of attacks. In spite of that, all of the traditional argumentation semantics give the network one single extension, namely  $\mathcal{E} = \{X, W, Z\}$ . This single extension fails to capture a lot of the information in the network. For instance, it does not distinguish between the accepted argument  $Z$ , which has one attack, and the accepted arguments  $X$  and  $W$  which have no attackers and therefore are uncontroversially accepted. From  $\mathcal{E}$  and  $\mathcal{A}$ , we can also deduce that the nodes  $Y$  and  $A$  are rejected, but this also fails to capture the fact that  $A$  has three attackers (including itself) and therefore is arguably more rejected than  $Y$ . The statuses of  $B$  and  $C$  are undecided, but although  $B$  is more attacked than  $C$ , the semantics also fails to reflect that. All of these facts can clearly be seen from the geometry of the network, but the traditional three-valued semantics is too coarse to capture them.

Various papers have tried to consider the geometry by looking at a node and the nodes attacking it, and the attackers of these attackers, and so on, until it went back to the top of the network to somehow measure how strongly each node is “in”, “out” or “undecided”. Our own approach is *numerical* using equations describing the node interactions to be able to naturally reflect numerically these geometrical considerations. What this means in principle is that the object-level instrument of traditional extensions cannot be

<sup>1</sup>Corresponding Author: Odinaldo Rodrigues, King’s College London, Department of Informatics; email: [odinaldo.rodrigues@kcl.ac.uk](mailto:odinaldo.rodrigues@kcl.ac.uk)

solely used to make the kind of distinctions we want to make about the nodes. We need to resort to external meta-level considerations. A recent brilliant numerical approach to tackle this problem was suggested in [3]. The authors however, do not connect with the traditional extensions. We notice that the *equational approach* for obtaining extensions does give numerical values between 0 and 1 for undecided nodes and these values reflect the degree of undecidedness of such nodes [4]. The approach does not distinguish however, between the nodes that are “in” and between the nodes that are “out”. The solution proposed here, which we call the *U-approach using Eq<sub>inv</sub>*, is conceptually very simple. We make all nodes undecided by having an external additional self-attacking node *U* attack every node. Solving equations for this augmented network now gives us the degree of “in”, “out” as well as “undecided” whilst still connecting with the traditional extensions as we shall see later in the paper.

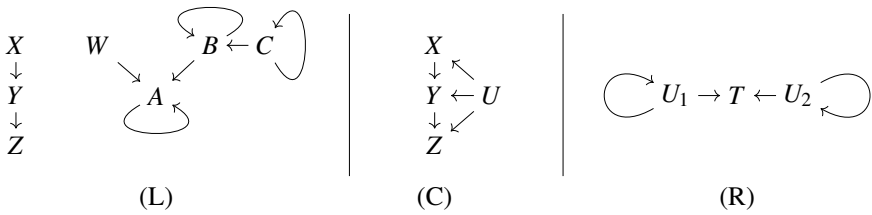


Figure 1. Sample argumentation networks used in the text.

To explain the main idea, consider the sub-network on the left of Figure 1 (L), which has the single extension  $\mathcal{E}' = \{X, Z\}$ . We would like to say that *Z* is more controversially accepted than *X* in  $\mathcal{E}'$ , because it is attacked by *Y* whereas *X* has no attackers. Adding a new self-attacking node *U*, which also attacks every other node in the sub-network gives the network in Figure 1 (C). We then write equations for each node, such as the ones that follow:<sup>2</sup>  $U = 1 - U$ ;  $X = 1 - U$ ;  $Y = (1 - U)(1 - X)$ ; and  $Z = (1 - U)(1 - Y)$ .

From the figure and solution to the equations, we see that the value of  $U = \frac{1}{2}$  is propagated to every node and moreover that the width and the depth of attacks is naturally reflected in the results, since the factor *U* is applied within each chain of attack: once to *X*, twice to *Y* (through *X* and through *Y* itself) and three times to *Z* (twice through *Y* and once from *Z* itself). These equations have the solution:  $U = X = \frac{1}{2}$ ,  $Y = \frac{1}{4}$ , and  $Z = \frac{3}{8}$ . From this solution, we have  $X > Z > Y$ . In the context of the extension  $\mathcal{E}'$ , *X* has a higher value than *Z* and is therefore more “in” than *Z*. *Y* is attacked by a node in  $\mathcal{E}'$  and is “out”. Applying the same reasoning to the sub-network on the right of Figure 1 (L), we get  $A = \frac{3}{19}$ ,  $B = \frac{1}{4}$ ,  $C = \frac{1}{3}$  and  $W = \frac{1}{2}$ . The reader will note that *B* and *Y* will get the same values. This is because the two weaker attackers of *B* are counterbalanced by the stronger attacker of *Y*. Geometrically they are indistinguishable, but a second meta-level criteria, such as the status with respect to the extension  $\mathcal{E}$  can be used. Looking back at  $\mathcal{E}$  for the whole network, we can now distinguish between the arguments in the categories “in”, “out” and “undecided” as follows:

$$\begin{array}{c} \text{More} \\ \downarrow \\ \text{Less} \end{array} \quad \text{In} \left\{ \begin{array}{l} X, W = \frac{1}{2} \\ | \\ Z = \frac{3}{8} \end{array} \right. \quad \text{Undecided} \left\{ \begin{array}{l} C = \frac{1}{3} \\ | \\ B = \frac{1}{4} \end{array} \right. \quad \text{Out} \left\{ \begin{array}{l} Y = \frac{1}{4} \\ | \\ A = \frac{3}{19} \end{array} \right.$$

<sup>2</sup>How to arrive at these equations will be explained in detail later.

Note in the solution above that it makes sense for  $A$  to be more “out” than  $Y$ , because it has the three attackers  $W$ ,  $B$  and itself, and at least one of these is as strong as  $X$ . So the calculations take into account the *number* of attackers as well as their *strength*.

The rest of the paper is structured as follows: Section 2 provides the background of the equational approach employed in this paper. Section 3 deals with  $U$ -approach using  $Eq_{inv}$  in detail. Section 4 compares our solution with the literature, and we conclude in Section 5.

## 2. Background

The *equational approach* views an argumentation network  $\langle \mathcal{A}, \mathcal{R} \rangle$  as a mathematical graph generating equations for functions in the unit interval  $[0, 1]$ . Any solution  $\mathbf{f}$  to these equations conceptually corresponds to an extension. Of course, the end result depends on how the equations are generated and we can get different solutions for different equations. Once the equations are fixed, the totality of the solutions to the system of equations is viewed as the totality of extensions via an appropriate mapping. Two equation schema we can possibly use for generating equations are  $Eq_{max}$  and  $Eq_{inv}$  below, where  $f(X)$  is the value of a node  $X \in \mathcal{A}$ :

$$(Eq_{max}) \quad f(X) = 1 - \max_{Y \in Att(X)} \{f(Y)\} \qquad (Eq_{inv}) \quad f(X) = \prod_{Y \in Att(X)} (1 - f(Y))$$

It is easy to see that according to  $Eq_{max}$  the value of any source argument will be 1 (since they have no attackers) and the value of any argument with an attacker with value 1 will be 0. Gabbay has shown that in the case of  $Eq_{max}$  the totality of solutions to the system of equations corresponds to the totality of extensions in Dung’s sense. Let  $\mathcal{N} = \langle \mathcal{A}, \mathcal{R} \rangle$  be an argumentation framework, the following two theorems, whose proofs can be found in [5], show the relationship between the solutions of  $Eq_{inv}$  and extensions of  $\mathcal{N}$ .

**Theorem 2.1 (Theorem 2.2 in [5])** *Every solution  $\mathbf{f}$  of  $Eq_{inv}$  equations written for an argumentation framework  $\mathcal{N}$  yields a complete extension for  $\mathcal{N}$ .*

**Theorem 2.2 (Theorem 2.3 in [5])** *Every preferred extension of an argumentation framework  $\mathcal{N}$  can be obtained from a solution  $\mathbf{f}$  of  $Eq_{inv}$  equations written for  $\mathcal{N}$ .*

In general terms, the following correspondence will relate a solution  $\mathbf{f}$  with the traditional semantics:  $f(X) = 1 \ :: \ X$  is “in”;  $f(X) = 0 \ :: \ X$  is “out”; and  $0 < f(X) < 1 \ :: \ X$  is “undecided”.

## 3. The $U$ -approach using $Eq_{inv}$

We now ask what equation schema is more appropriate to capture the geometry of a network?  $Eq_{max}$  will disregard all but the attacks with maximum value, so it is not ideal. If we think in terms of *probability*, we want the values obtained as solutions to the equations to reflect the probability of being “in”. Thus, 1 is definitely “in” and 0 is definitely **not** “in”, i.e., definitely “out”.  $\frac{1}{2}$  is right in the middle of “in” and “out” and hence means definitely “undecided”. Following this reasoning, now consider Figure 1 (R). We ask what is the probability that  $T$  is “in”?  $T$  is “in”, if both  $U_1$  and  $U_2$  are “out” (i.e., if both have value 0). So it is the product of the probability of each  $U_i$  being “out”, which is

the product  $(1 - U_1) \times (1 - U_2)$ . This motivates the use of the product operation in our computations (i.e., via  $Eq_{inv}$ ).

An admirable discussion of these issues by Cayrol and Lagasque-Schiex can be found in [6]. We would like to use  $Eq_{inv}$  in a way that responds to these intuitions, without resorting to all kinds of meta-level geometrical analyses and distinctions. However, if we simply use  $Eq_{inv}$  we would not be able to distinguish the nodes in the categories “in” and “out”,<sup>3</sup> so our idea is to make every node “undecided” in a uniform way thus allowing for a larger spread of values in all categories.

So given  $\langle \mathcal{A}, \mathcal{R} \rangle$ , we move to  $\langle \mathcal{A}_U, \mathcal{R}_U \rangle$  with a new node  $U$  making all nodes in  $\mathcal{A}$  “undecided”. The relative values in solving  $Eq_{inv}$  for  $\langle \mathcal{A}_U, \mathcal{R}_U \rangle$  give the relative strength of all nodes in  $\langle \mathcal{A}, \mathcal{R} \rangle$ .<sup>4</sup> So for any extension  $\mathcal{E}$  of  $\mathcal{A}$  and any  $X, Y \in \mathcal{E}$  both are “in” in  $\langle \mathcal{A}, \mathcal{R} \rangle$  but they are undecided in  $\langle \mathcal{A}_U, \mathcal{R}_U \rangle$  and may have different values in a solution  $\mathbf{f}$  to the  $Eq_{inv}$  equations of  $\langle \mathcal{A}_U, \mathcal{R}_U \rangle$ . These different values will give us an indication of how much “in”  $X, Y$  are in  $\mathcal{E}$  (and similarly for nodes that are “out” or “undecided”).

**Definition 3.1** Let  $f : \mathcal{A} \mapsto [0, 1]$  be an assignment of values to elements of  $\mathcal{A}$ . We define the sets  $in(f) = \{X \in \mathcal{A} \mid f(X) = 1\}$  and  $out(f) = \{X \in \mathcal{A} \mid f(X) = 0\}$ .

**Definition 3.2 (U-Augmentation of an Argumentation Framework)** The  $U$ -augmentation of the argumentation network  $\langle \mathcal{A}, \mathcal{R} \rangle$  is the network  $\langle \mathcal{A}_U, \mathcal{R}_U \rangle$ , where  $U \notin \mathcal{A}$ ,  $\mathcal{A}_U = \mathcal{A} \cup \{U\}$  and  $\mathcal{R}_U = \mathcal{R} \cup \{(U, U)\} \cup \{U\} \times \mathcal{A}$ .

The relative degree of membership of each node in the categories “in”, “out” and “undecided” is defined as follows.

**Definition 3.3 (Numerical evaluation of the degree of “in”, “out”, and “undecided” in abstract argumentation frameworks)** Let  $\mathcal{N} = \langle \mathcal{A}, \mathcal{R} \rangle$  be a network and  $\mathcal{N}_U$  its  $U$ -augmentation. Let  $\mathcal{E}$  be an extension for  $\mathcal{N}$ . Let  $\mathbf{f}$  be a solution to the  $Eq_{inv}$  equations for  $\mathcal{N}_U$ . Let  $\leq_f$  be an ordering on  $\mathcal{A}$  defined by  $X \leq_f Y$  iff  $\mathbf{f}(X) \leq \mathbf{f}(Y)$ . Then  $\leq$  induces an ordering on the sets  $IN = \mathcal{E}$  (“in”);  $OUT = \{X \in \mathcal{A} \mid \exists Y \in \mathcal{E} \text{ such that } (Y, X) \in \mathcal{R}\}$  (“out”); and  $UND = \mathcal{A} \setminus (\mathcal{E} \cup OUT)$  (“undecided”), giving a degree scale in each category.

Note that Definition 3.3 offers a geometrical ranking of the nodes in a network ( $\leq_f$ ) which is independent of the notion of extension but can be used in conjunction with an extension to distinguish the nodes in the categories “in”, “out” and “undecided” with respect to that extension.

**Remark 3.1** It should be clear that if the network  $\mathcal{N} = \langle \mathcal{A}, \mathcal{R} \rangle$  is acyclic, then a solution to the system of  $Eq_{inv}$  equations to the  $U$ -augmentation of  $\mathcal{N}$  exists and is unique. To see this, we simply order the equations in ascending order of the longest chain of attack of each node and solve them in this order.  $U$  will solve to  $\frac{1}{2}$  as well as every source node in  $\mathcal{N}$ . We then propagate these values until all node values are calculated. This will form the unique solution  $\mathbf{f}$ .

<sup>3</sup>In light of Theorem 2.2, all nodes in a preferred extension have value 1 and all nodes attacked by the extension have value 0.

<sup>4</sup>As nicely put by one of the reviewers of this paper, this hypothetical node  $U$  could represent an unforeseen future argument attacking all nodes.

In the general case, we believe the solution  $\mathbf{f}$  of Definition 3.3 exists and is unique. Empirical results also suggest this. The proof would be similar to the proof of uniqueness in [3], but it has to be written down fully to confirm. Note that because of the introduction of the new node  $U$ , all equations involved become contractions.

We now show a number of properties of the solutions to the system of equations. The first one has to do with the upper bound of the value of a node and the second one with the effect of attacks on nodes.

**Proposition 3.1** Let  $\mathcal{N} = \langle \mathcal{A}, \mathcal{R} \rangle$  be an argumentation network and  $\mathbf{f}$  a solution to the  $Eq_{inv}$  equations written for the  $U$ -augmentation of  $\mathcal{N}$ . For all  $A \in \mathcal{A}$ ,  $\mathbf{f}(A) \leq \frac{1}{2}$ .

The above proposition shows that 1)  $\frac{1}{2}$  is the upper bound for the values of an  $U$ -augmented network; and 2) source nodes get maximum value, i.e.,  $\frac{1}{2}$ . It is also easy to see that the values of nodes decrease proportionally to the number of attackers. In particular:

**Proposition 3.2** Let  $\mathcal{N} = \langle \mathcal{A}, \mathcal{R} \rangle$  be an argumentation network and  $\mathbf{f}$  a solution to the  $Eq_{inv}$  equations written for the  $U$ -augmentation of  $\mathcal{N}$ . Let  $X, Y \in \mathcal{A}$ , such that  $Att(X) \subseteq Att(Y)$ , then  $\mathbf{f}(Y) \leq \mathbf{f}(X)$ .

**Proposition 3.3** Let  $\mathcal{N} = \langle \mathcal{A}, \mathcal{R} \rangle$  be a finite argumentation network and  $\mathbf{f}$  a solution to the  $Eq_{inv}$  equations written for the  $U$ -augmentation of  $\mathcal{N}$ . Take  $X \in \mathcal{A}$  and let  $|Att(X)| = k$ . Then  $\mathbf{f}(X) \geq \frac{1}{2^{k+1}}$ .

#### 4. Comparisons with Other Work

We start our comparisons with the approach in [3] in which the relative strengths of arguments in a graph are indirectly calculated in terms of the relative *burden number* of these arguments. Essentially, this technique assigns a unique rank to every node which can be compared with the relative ranking our geometrical interpretation gives:

**Definition 4.1** ( $s_\alpha$ , [3]) Let  $\alpha \in (0, +\infty)$  and  $\mathcal{F} = \langle \mathcal{A}, \mathcal{R} \rangle$  be an argumentation graph. We define the function  $s_\alpha$  as follows:  $s_\alpha : \mathcal{A} \mapsto [1, +\infty)$  such that  $\forall a \in \mathcal{A}$ ,

$$s_\alpha(a) = 1 + \left( \sum_{b \in Att(a)} \frac{1}{(s_\alpha(b))^\alpha} \right)^{\frac{1}{\alpha}}$$

If  $Att(a) = \emptyset$ , then  $s_\alpha(a) = 1$ .  $s_\alpha(a)$  is called the *burden number* of  $a$ .

It is easy to see that, as is the case in our technique, the value of  $s_\alpha$  takes into account both the number of attackers of an argument as well as the relative strength of these attackers. An argument with a small burden number is deemed more acceptable than an argument with a greater burden providing what was called a *compensation-based semantics*.

In order to compare the results we will use the sample networks in Figure 2 taken from [3]. The computed  $s_\alpha$  values for  $\alpha = 1$  of all nodes in the networks are given in Table 1 along with the solutions for the  $U$ -augmentation of the networks.

It is easy to see that although the values differ, the rankings of arguments in networks  $\mathbf{N}_1, \mathbf{N}_2, \mathbf{N}_3$  and  $\mathbf{N}_5$  are exactly the same. However, in network  $\mathbf{N}_4$ , compensation-based semantics fails to distinguish between arguments  $P, Q$  and  $A$  when  $\alpha = 1$ , whereas our formalism considers  $P$  and  $Q$  equivalent *but* strictly weaker than  $A$  (remember  $P$  are  $Q$

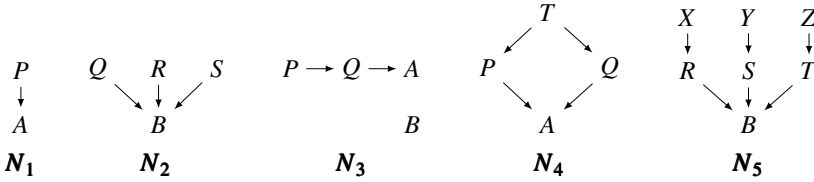


Figure 2. Some networks taken from [3] for comparison.

$N_1$ :	$P$	$A$	$N_2$ :	$Q$	$R$	$S$	$B$	$N_3$ :	$P$	$B$	$A$	$Q$
$\alpha = 1$	1	2		1	1	1	4		1	1	$\frac{3}{2}$	2
$U(\frac{1}{2})$	$\frac{1}{2}$	$\frac{1}{4}$		$\frac{1}{2}$	$\frac{1}{2}$	$\frac{1}{2}$	$\frac{1}{16}$		$\frac{1}{2}$	$\frac{1}{2}$	$\frac{3}{8}$	$\frac{1}{4}$
$N_4$ :	$T$	$A$	$P$	$Q$	$N_5$ :	$X$	$Y$	$Z$	$R$	$S$	$T$	$B$
$\alpha = 1$	1	2	2	2		1	1	1	2	2	2	$\frac{5}{2}$
$U(\frac{1}{2})$	$\frac{1}{2}$	$\frac{9}{32}$	$\frac{1}{4}$	$\frac{1}{4}$		$\frac{1}{2}$	$\frac{1}{2}$	$\frac{1}{2}$	$\frac{1}{4}$	$\frac{1}{4}$	$\frac{1}{4}$	$\frac{27}{128}$

Table 1. Summary of comparison with compensation-based semantics rankings [3].

are both “out” whereas  $A$  is “in” in any traditional semantics). We argue that in  $N_4$   $A$  should be more acceptable than both  $P$  and  $Q$ , because although it has two attackers, these are both defeated by  $T$ . Note that in theirs and ours we still obtain that  $A$  is less acceptable than  $T$  itself, as expected. In general, we get the following values for the nodes in  $N_4$  in the compensation-based semantics:  $s_\alpha(T) = 1$ ,  $s_\alpha(P) = s_\alpha(Q) = 2$  (for  $\alpha \geq 1$ ); and  $s_\alpha(A) = 2$ , if  $\alpha = 1$ ; and  $1 \leq s_\alpha(A) < 2$ , if  $\alpha > 1$ . This shows that within the single network  $N_4$  we can get different rankings for  $A$ ,  $P$  and  $Q$  depending on the value of  $\alpha$  used. The idea behind this is to fine-tune “the influence of the quality of the attackers”. This has two main problems. 1) It is difficult to know *in advance* which value of  $\alpha$  to use. Consider the network  $N_6$  formed by the aggregation of  $N_1$  and  $N_5$  (this is network  $\mathcal{F}_4$  in [3]). In  $N_6$  as well, the relative acceptability ranking between  $A$  and  $B$  will vary depending on the value of  $\alpha$  chosen.<sup>5</sup> 2) Fixing the relative ranking of some nodes by employing a certain value of  $\alpha$  may inadvertently cause the ranking of other nodes to change. In  $N_4$ , if  $\alpha > 1$ , then again our formalism and theirs will agree on the ranking of all arguments in  $N_4$  except in the limit  $\alpha \rightarrow \infty$ . We argue that it is simpler to adjust the impact attacks have on the values of the nodes according to the application via the equational approach, because we can separate out the necessary components via an appropriate equation schema. In the case of  $Eq_{inv}$  there are two separate components dealing with attacks: the complement-to-1 function (for attack itself) and product (for their aggregation). More sophisticated  $t$ -norms can be used instead of product. In Definition 4.1 the two components are intertwined.

In [6], Cayrol and Lagasque-Schiex propose two approaches to evaluate the value of an argument based solely on the attack relation of the argumentation framework to which it belongs. The first approach calculates the value of an argument using only the values of its direct attackers and is called *local*; the second approach takes into account the set of all ancestors of the argument in the attack relation and is called *global*.<sup>6</sup>

In the local approach, the value  $v(X)$  of an argument  $X$  is obtained via the composition of two functions  $h$  and  $g$ :  $h$  calculates the value of each attacker of  $X$ ; and  $g$  com-

<sup>5</sup>For  $\alpha \approx 1.585$ ,  $s_\alpha(A) = s_\alpha(B)$ .

<sup>6</sup>As it turns out, the values of the attackers of an argument are calculated recursively so in effect the local approach also takes into account all ancestors of an argument.

puts the effect on  $X$  of the aggregation of the attacks on it. The values of the nodes as calculated by the schema  $Eq_{inv}$  can be seen as a *local valuation* of the nodes in the same way as Besnard and Hunter’s  $h$ -categoriser valuation [7]. In our case, the function  $h$  is the complement to 1 and  $g$  is half of the product of the attacks.

In [8], Modgil and Grossi proposed a framework to take into account the degree of justification of the arguments of an argumentation framework. The central idea is the notion of *graded defense* which counts the number of attackers and defenders of a node. This is used to define *graded extensions*, which are parametrised by two integers  $m$  for attacks and  $n$  for defenders, and in essence select arguments with a particular configuration of attackers and defenders. So the motivation is the same, but the approach is completely different to ours.

Finally, in [9], Thimm and Kern-Isberner propose a *stratified semantics*, which assigns different ranks for the arguments of a network. The ranks are constructed by successively taking the accepted arguments of a network according to a given semantics, assigning them a rank, then considering the network resulting from the removal of such nodes and then re-calculating the nodes in the next rank until all nodes are ranked. This fails to distinguish between the accepted arguments in each rank, but agrees in spirit with our treatment of extensions in that it follows the directionality of attacks.

## 5. Conclusions, Discussion and Future Work

Given a network  $\langle \mathcal{A}, \mathcal{R} \rangle$  and a complete extension  $\mathcal{E} \subseteq \mathcal{A}$ , our objective was to provide a ranking of the arguments in  $\mathcal{A}$  given  $\mathcal{E}$  taking into account the geometry of  $\langle \mathcal{A}, \mathcal{R} \rangle$ .

We offered a solution to the following meta-level problem  $\Pi$ : given an argumentation network  $\langle \mathcal{A}, \mathcal{R} \rangle$ , we note that some nodes are geometrically attacked more than others or are more “loopy” than others. Can we make this observation more quantitative? Our solution used the *equational approach* in an augmented network with set of nodes  $\mathcal{A}^* = \mathcal{A} \cup \{U\}$ .

We now discuss the methodological aspect of our solution. The problem  $\Pi$  above is a special case of a general problem: given an object-level system  $\mathcal{S}$  and a meta-level property  $\mathbb{P}$  of  $\mathcal{S}$  how can we express/discuss/quantify this property for  $\mathcal{S}$ ?

There are two ways: 1) Construct a meta-level system  $\mathbb{P}(\mathcal{S})$  to describe/discuss  $\mathbb{P}$ ; and 2) Construct a new system  $\mathcal{S}^*$  out of  $\mathcal{S}$ , and within  $\mathcal{S}^*$ , the property  $\mathbb{P}$  can be highlighted. Method 2) is better for the following reasons. It is simpler, using the same machinery used in  $\mathcal{S}$ . It is also more robust. If we modify, generalise or apply  $\mathcal{S}$ , we do the same for  $\mathcal{S}^*$  and thus carry the results for the property  $\mathbb{P}$ . If we use  $\mathbb{P}(\mathcal{S})$  we may not know what to do for  $\mathbb{P}(\mathcal{S}^*)$ .

Our approach followed method 2) above.  $\mathcal{S}^*$  is simply our  $U$ -augmented network and we use for  $\mathcal{S}^*$  the same machinery for finding extensions that we use for  $\mathcal{S}$ .

Bearing the above considerations in mind, let us summarise what we did. We know that the solutions to classes of equations written for an argumentation framework have a correspondence with the set of extensions of that network. In the case of the equation schema  $Eq_{max}$  the totality of the solutions to the system of equations corresponds to the totality of complete extensions of the network. In the case of the schema  $Eq_{inv}$ , the solutions only yield preferred extensions. However, for the quantitative measurement of attack we wanted to consider in this paper,  $Eq_{inv}$  has a significant advantage over  $Eq_{max}$ , because  $Eq_{max}$  simply takes the maximum value of the attacks, whereas  $Eq_{inv}$  provides a number reflecting the effect of the aggregation of all attacks on a node.

Because of  $Eq_{inv}$ 's correspondence with the class of preferred extensions, we cannot simply use it directly, since it will neither differentiate between the nodes in a preferred extension (i.e., the ones with value 1) nor will it differentiate between the nodes attacked by a node in the extension (i.e., the ones with value 0).  $Eq_{inv}$  will however differentiate between the nodes in the “undecided” range. So our simple solution is to force all nodes into the “undecided” range and then use the relative ranking of the nodes thus obtained to distinguish between the nodes given the *original* extension. Conceptually, this can be done simply by considering a modified network with a new undecided node attacking all original nodes. Mathematically, this can be seen as yielding a new schema of equations, call it,  $Eq_{deg}$  such that the value of a node  $X$  is defined as  $f(X) = \varepsilon \times \prod_{Y \in Att(X)} (1 - f(Y))$ .

We took  $\varepsilon$  to be  $\frac{1}{2}$ . This is not simply the same as multiplying a solution to  $Eq_{inv}$  by  $\varepsilon = \frac{1}{2}$ . This is indeed a new class of equations. The reader might then ask why  $\varepsilon = \frac{1}{2}$ ? Conceptually, it makes sense to use  $\frac{1}{2}$  as it is arguably the most “undecided” value and given its connection with the  $U$ -augmentation it is what a node with a single self-attack resolves to. Some further comparisons with other work on ranking semantics (e.g., [10,11,12]) as well as the effects of using different values of  $\varepsilon \in (0, 1)$  is left for a full version of this paper.

## References

- [1] C. da C. Pereira, A. G. B. Tettamanzi, and S. Villata. Changing one’s mind: erase or rewind? possibilistic belief revision with fuzzy argumentation based on trust. In *Proceedings of the 22nd International joint conference on artificial intelligence : IJCAI’11*, pages 164 – 171, Menlo Park, 2011. AAAI Press.
- [2] P. Dondio. Multi-valued argumentation frameworks. In Antonis Bikakis, Paul Fodor, and Dumitru Roman, editors, *Rules on the Web. From Theory to Applications*, volume 8620 of *Lecture Notes in Computer Science*, pages 142–156. Springer International Publishing, 2014.
- [3] L. Amgoud, J. Ben-Naim, D. Doder, and S. Vesic. Ranking arguments with compensation-based semantics. In *Proceedings of the 15th International Conference on Principles of Knowledge Representation and Reasoning, KR 2016*, pages 12–21, 2016.
- [4] D. M. Gabbay. Equational approach to argumentation networks. *Argument and Computation*, 3:87–142, 2012. DOI: 10.1080/19462166.2012.704398.
- [5] D. M. Gabbay and O. Rodrigues. Probabilistic argumentation: An equational approach. *Logica Universalis*, 9(3):345–382, 2015.
- [6] C. Cayrol and M.-C. Lagasque-Schiex. Graduality in argumentation. *Journal of Artificial Intelligence Research*, 23:245–297, 2005.
- [7] P. Besnard and A. Hunter. A logic-based theory of deductive arguments. *Artificial Intelligence*, 128(1-2):203 – 235, 2001.
- [8] D. Grossi and S. Modgil. On the graded acceptability of arguments. In *Proceedings of the Twenty-Fourth International Joint Conference on Artificial Intelligence, IJCAI 2015*, pages 868–874, 2015.
- [9] M. Thimm and G. Kern-Isberner. On controversiality of arguments and stratified labelings. In *Proceedings of the Fifth International Conference on Computational Models of Argumentation (COMMA’14)*, September 2014.
- [10] E. Bonzon, J. Delobelle, S. Konieczny, and N. Maudet. A comparative study of ranking-based semantics for abstract argumentation. In *Proceedings of the Thirtieth AAAI Conference on Artificial Intelligence, February 12-17, 2016, Phoenix, Arizona, USA.*, pages 914–920, 2016.
- [11] L. Amgoud and J. Ben-Naim. Ranking-based semantics for argumentation frameworks. In *Scalable Uncertainty Management - 7th International Conference, SUM 2013, Washington, DC, USA, September 16-18, 2013. Proceedings*, pages 134–147, 2013.
- [12] P.-A. Matt and F. Toni. A game-theoretic measure of argument strength for abstract argumentation. In *Logics in Artificial Intelligence, 11th European Conference, JELIA 2008, Dresden, Germany, September 28 - October 1, 2008. Proceedings*, pages 285–297, 2008.