Adapting the DF-QuAD Algorithm to Bipolar Argumentation

Antonio RAGO [a,1], Kristijonas ČYRAS [a] and Francesca TONI [a]
[a] *Department of Computing, Imperial College London, UK*

Abstract. We define a quantitative semantics for evaluating the strength of arguments in Bipolar Argumentation frameworks (BAFs) by adapting the Discontinuity-Free QuAD (DF-QuAD) algorithm previously used for evaluating the strength of arguments in Quantitative Argumentation Debates (QuAD) frameworks. We study the relationship between the new semantics and some existing semantics for other argumentation frameworks, as well as some properties of the semantics.

Keywords. Bipolar argumentation, Quantitative semantics, QuAD frameworks

1. Introduction

Bipolar Argumentation Frameworks (BAFs) [1] extend *Abstract Argumentation frameworks* (AFs) [2], consisting of *arguments* and *attacks* between them, to include *supports* between arguments. Formally a BAF is a triple $\langle \mathcal{X}, \mathcal{R}^-, \mathcal{R}^+ \rangle$ consisting of a set $\mathcal{X}$ of arguments, a binary (*attack*) relation $\mathcal{R}^-$ on $\mathcal{X}$ and a binary (*support*) relation $\mathcal{R}^+$ on $\mathcal{X}$.

In recent years, the standard, qualitative *acceptance* semantics first proposed for AFs in [2] and for BAFs in [1] have been supplemented with quantitative semantics, e.g. as in [3,4,5]. These measure arguments on a gradual scale (usually $\mathbb{I} = [0,1]$, as in this paper) so that arguments can be ranked against one another. Some of these quantitative semantics are defined for variants of BAFs, e.g. for the *Quantitative Argumentation Debate* (QuAD) frameworks of [7], corresponding to restricted forms of BAFs, namely (acyclic) trees, extended with *base scores* (i.e. intrinsic strengths) for arguments, the *Discontinuity-Free QuAD* (DF-QuAD) algorithm of [6] is defined. Similarly, the *social models* of [8] are defined for the *Social Abstract Argumentation Frameworks* (SAFs) of [8]. SAFs are cyclic or acyclic AFs (without supports) extended with positive and negative votes for arguments used to determine their *social support* (i.e., again, intrinsic strength). These methods define an overall strength for each argument, based on their intrinsic strength and the strengths of its attackers and, where present, supporters.

In this paper we define a novel quantitative semantics for evaluating the strength of arguments in BAFs by adapting the DF-QuAD algorithm originally defined for QuAD frameworks. We also relate the new semantics and other methods in the literature, namely the *gradual valuation* semantics [9] for BAFs and Social Models for SAFs [8].

[1] Corresponding Author; E-mail: a.rago15@imperial.ac.uk

2. Adapting DF-QuAD

In this section, unless specified otherwise, we will assume as given a generic BAF $\langle \mathcal{X}, \mathcal{R}^-, \mathcal{R}^+ \rangle$. With an abuse of terminology, given any $a \in \mathcal{X}$, we will use $\mathcal{R}^-(a)$ to denote $\{b \in \mathcal{X} | (b,a) \in \mathcal{R}^-\}$, referred to as the set of all *attackers* against a, and $\mathcal{R}^+(a)$ to denote $\{b \in \mathcal{X} | (b,a) \in \mathcal{R}^+\}$, referred to as the set of all *supporters* for a. Also, we will use $\mathbb{I}^* \triangleq \bigcup_{i \geq 0} \mathbb{I}^i$ to denote the set of all possible sequences of elements of $\mathbb{I}$.

We now present our adaptation of the DF-QuAD semantics for BAFs. This determines the strength of an argument based on the aggregated strength of its attackers and supporters. The semantics differs from the original DF-QuAD semantics for QuAD frameworks in that it does not require a base score for arguments. Like the original DF-QuAD semantics for QuAD frameworks though it relies upon a *strength aggregation function*, defined exactly as in [6] as follows:

Definition 1. The *strength aggregation function* is defined as $\sigma : \mathbb{I}^* \to \mathbb{I}$, where for $S = (v_1, \ldots, v_n) \in \mathbb{I}^*$:

$$\begin{aligned}
&\text{if } n = 0 : \sigma(S) = 0 \\
&\text{if } n = 1 : \sigma(S) = v_1 \\
&\text{if } n = 2 : \sigma(S) = f(v_1, v_2) \\
&\text{if } n > 2 : \sigma(S) = f(\sigma(v_1, \ldots, v_{n-1}), v_n)
\end{aligned}$$

with the *base function* $f : \mathbb{I} \times \mathbb{I} \to \mathbb{I}$ defined, for $v_1, v_2 \in \mathbb{I}$, as:

$$f(v_1, v_2) = v_1 + (1 - v_1) \cdot v_2 = v_1 + v_2 - v_1 \cdot v_2$$

Thus, the base function is the building block to handle sequences of strengths of attackers or supporters by σ, proportionally increasing the attacking or supporting arguments' strength towards 1, as in [6].

Once the strengths of attacking (resp. supporting) arguments against (resp. for) an argument have been aggregated separately using σ, the following *mediating function* is used to obtain the overall strength of arguments, using the distance between the strengths of attacking and supporting arguments:

Definition 2. The *mediating function* is defined as $\mu : \mathbb{I} \times \mathbb{I} \to \mathbb{I}$, where for $v_a, v_s \in \mathbb{I}$:

$$\mu(v_a, v_s) = 0.5 + 0.5 \cdot (v_s - v_a) \tag{1}$$

The intuition behind this function is that if the strength of the attacking arguments is larger (resp. smaller) we obtain a score that is below (resp. above) 0.5, using the difference between the aggregated supporting and attacking strengths to obtain a score that is proportionally closer to 0 (resp. 1).

The argument's overall strength is then determined by the following *score function*:

Definition 3. The *score function* is defined as $\mathcal{SF} : \mathcal{X} \to \mathbb{I}$ where, for any $a \in \mathcal{X}$:

$$\mathcal{SF}(a) = \mu(\sigma(SEQ_{\mathcal{SF}}(\mathcal{R}^-(a))), \sigma(SEQ_{\mathcal{SF}}(\mathcal{R}^+(a))))$$

where if $(a_1,\dots,a_n)$ is an arbitrary permutation of the $(n \geq 0)$ attackers in $\mathcal{R}^-(a)$, $SEQ_{\mathcal{SF}}(\mathcal{R}^-(a)) = (\mathcal{SF}(a_1),\dots,\mathcal{SF}(a_n))$ (similarly for supporters).

Note that whenever $\mathcal{X}$ is finite with $|\mathcal{X}| = n$, the output of the score function $\mathcal{SF}$ for $a \in \mathcal{X}$ can be seen as the (appropriate) projection of the output of a function $\mathcal{SF} : \mathbb{I}^n \to \mathbb{I}^n$ which takes a 'valuation' $(v_1,\dots,v_n) \in \mathbb{I}^n$ of all the arguments in $\mathcal{X}$ and maps it to another 'valuation'. In particular, the i^{th} coordinate v_i of $(v_1,\dots,v_n) \in \mathbb{I}^n$ is mapped thus:

$$\mathcal{SF}(v_i) = 0.5 + 0.5\left(\prod_{a_k \in \mathcal{R}^-(a_i)} (1 - v_k) - \prod_{a_m \in \mathcal{R}^+(a_i)} (1 - v_m)\right).$$

If $\mathbb{I}$ is a closed interval $[a,b] \subseteq \mathbb{R}$ on the real line, then a generalised version of Brouwer's fixed-point theorem ensures that $\mathcal{SF}$ has a fixed-point, provided $\mathcal{SF}$ is continuous [10]. Clearly, $\mathcal{SF}$ is continuous, because it is obtained by composition of continuous functions, namely sum, product and projections. Consequently, if $\mathcal{X}$ is finite and $\mathbb{I}$ is a closed interval on the real line, in particular if $\mathbb{I} = [0,1]$, then $\mathcal{SF} : \mathcal{X} \to \mathbb{I}$ has at least one solution, i.e. the arguments' overall strengths are well-defined.

To show that there is a unique solution, one could utilize Banach's fixed-point theorem, by showing that $\mathcal{SF}$ is a contraction mapping, e.g. as in [8]. To this end, note that the Jacobian $\mathcal{J}$ of $\mathcal{SF}$ is continuous and differentiable, and its derivative's $\mathcal{J}'$ matrix norm $\|\mathcal{J}'\|$ given by the maximum absolute value of its entries (since matrix norms on $\mathbb{R}^n$ are equivalent, we can use the most suitable one) satisfies $\|\mathcal{J}'(v_1,\dots,v_n)\| \leqslant 1$ for any $(v_1,\dots,v_n) \in \mathbb{I}^n$. The latter inequality needs to be strict in order to apply Banach's fixed-point theorem, but we conjecture that this is attainable and that with finite $\mathcal{X}$ and $\mathbb{I} = [0,1]$, the arguments' strength is guaranteed to be unique, i.e. equations defined by $\mathcal{SF}$ have a unique solution. Proving this conjecture is left for future work.

For illustration, Figure 1 shows some BAFs with arguments' strengths as indicated.

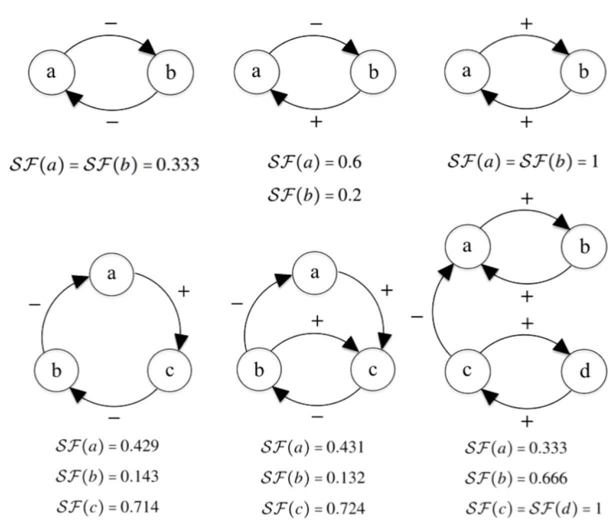

Figure 1. Example BAFs with strengths of arguments as indicated

Most of the properties of the DF-QuAD algorithm for QuAD frameworks, given and proven in [6], also hold for our adaptation of the DF-QuAD algorithm for BAFs. In particular, the following properties hold, mirroring the corresponding properties in [6] (i.e. Proposition 7 and Proposition 9, resp.)

Proposition 1. For any $S = (v_1, \dots, v_n) \in \mathbb{I}^*$ (for $n \geq 0$), $\sigma(S \bigcup (1)) = 1$.

In other words, adding an attacker/supporter with maximum strength to the list of attakers/supporters (resp.) saturates the strength aggregation function.

Proposition 2. For any $a \in \mathcal{X}$ with $\sigma(SEQ_{\mathcal{SF}}(\mathcal{R}^-(a))) = v_a$, $\sigma(SEQ_{\mathcal{SF}}(\mathcal{R}^+(a))) = v_s$:

$$\mathcal{SF}(a) = 0 \text{ iff } v_a = 1 \wedge v_s = 0$$
$$\mathcal{SF}(a) = 1 \text{ iff } v_a = 0 \wedge v_s = 1$$

In other words, the extreme values, 0 and 1, can only be achieved as the strength of an argument if they are somehow already present.

3. Comparison with Other Approaches

QuAD frameworks [7] where all arguments have a base score of 0.5, under the DF-QuAD algorithm of [6], can be mapped to BAF frameworks, under the new semantics given in Section 2. The mapping is defined as follows:

Definition 4. Let $\mathcal{Q} = \langle \mathcal{A}, \mathcal{C}, \mathcal{P}, \mathcal{R}, \mathcal{BS}^{0.5} \rangle$ be a QuAD framework with $\mathcal{BS}^{0.5} \triangleq (\mathcal{A} \times \mathcal{C} \cup \mathcal{P}) \times \{0.5\}$.[2] The *corresponding BAF* is defined as $\mathcal{B} = \langle \mathcal{X}, \mathcal{R}^-, \mathcal{R}^+ \rangle$ such that $\mathcal{X} = \mathcal{A} \cup \mathcal{C} \cup \mathcal{P}$, $\mathcal{R}^- = \mathcal{R} \cap (\mathcal{C} \times \mathcal{X})$ and $\mathcal{R}^+ = \mathcal{R} \cap (\mathcal{P} \times \mathcal{X})$.

Since $\mathcal{R}$ is acyclic, by definition of QuAD framework, the BAF corresponding to a QuAD framework is guaranteed to be acyclic. Thus, QuAD frameworks can be seen as restricted types of BAFs.

We now show that the strength of arguments in a QuAD framework by the DF-QuAD semantics in [6] is the same as the strength of arguments in the corresponding BAF obtained by Definition 3. We have already stated that the strength aggregation function in Definition 1 is borrowed from the DF-QuAD semantics in [6]. The notion of strength in [6] is however defined in terms of a *combination function* (rather than a mediating function as in our case) defined in [6] as $c : \mathbb{I} \times \mathbb{I} \times \mathbb{I} \to \mathbb{I}$, where for $v_0, v_a, v_s \in \mathbb{I}$:

$$c(v_0, v_a, v_s) = v_0 - v_0 \cdot |v_s - v_a| \qquad \text{if } v_a \geq v_s \tag{2}$$
$$c(v_0, v_a, v_s) = v_0 + (1 - v_0) \cdot |v_s - v_a| \qquad \text{if } v_a < v_s \tag{3}$$

The following proposition shows that the mediating function is equivalent to the combination function when the base score (v_0) is set to 0.5:

Proposition 3. For any $v_a, v_s \in \mathbb{I}$, $c(0.5, v_a, v_s) = \mu(v_a, v_s)$.

[2] By definition, $\mathcal{A}$, $\mathcal{C}$ and $\mathcal{P}$ are pairwise disjoint and $\mathcal{R} \subseteq (\mathcal{C} \cup \mathcal{P}) \times (\mathcal{A} \times \mathcal{C} \cup \mathcal{P})$ is acyclic, see [7].

Proof. If $v_a \geq v_s$, by Eq.2 $c(0.5, v_a, v_s) = 0.5 - 0.5 \cdot (v_a - v_s) = 0.5 + 0.5 \cdot (v_s - v_a)$, corresponding to Eq.1. If $v_s > v_a$, by Eq.3 $c(0.5, v_a, v_s) = 0.5 + (1 - 0.5) \cdot (v_s - v_a) = 0.5 + 0.5 \cdot (v_s - v_a)$, again corresponding to Eq.1. □

The next proposition follows immediately, given $\mathcal{SF}_Q$ defined in [6] as $\mathcal{SF}_Q : \mathcal{A} \cup \mathcal{C} \cup \mathcal{P} \to \mathbb{I}$ where, for any $a \in \mathcal{A} \cup \mathcal{C} \cup \mathcal{P}$, $\mathcal{SF}_Q(a) = c(\mathcal{BS}(a), \sigma(SEQ_{\mathcal{SF}_Q}(\mathcal{R}^-(a))), \sigma(SEQ_{\mathcal{SF}_Q}(\mathcal{R}^+(a))))$ where if $(a_1, \ldots, a_n)$ is an arbitrary permutation of the $(n \geq 0)$ attackers in $\mathcal{R}^-(a)$, $SEQ_{\mathcal{SF}_Q}(\mathcal{R}^-(a)) = (\mathcal{SF}_Q(a_1), \ldots, \mathcal{SF}_Q(a_n))$ (similarly for supporters).

Proposition 4. Given a QuAD framework $\mathcal{Q} = \langle \mathcal{A}, \mathcal{C}, \mathcal{P}, \mathcal{R}, \mathcal{BS}^{0.5} \rangle$ and the corresponding BAF $\mathcal{B} = \langle \mathcal{X}, \mathcal{R}^-, \mathcal{R}^+ \rangle$, for $a \in \mathcal{X}$, let $s_{\mathcal{Q}}(a)$ be $\mathcal{SF}_Q(a)$ in $\mathcal{Q}$ and $s_{\mathcal{B}}(a)$ be $\mathcal{SF}(a)$ in $\mathcal{B}$. Then for all $a \in \mathcal{X}$: $s_{\mathcal{Q}}(a) = s_{\mathcal{B}}(a)$.

Thus, our new semantics for BAFs can be seen as a specialisation (setting all base scores to 0.5) as well as a generalisation (to any BAFs, rather than those corresponding to QuAD frameworks) of the original DF-QuAD semantics in [6].

Since BAFs are an extension of AFs [1], social models of SAFs [8] give the same output as our score function when the social support resulting from votes amounts to 0.5 (this equates to an approximately equal number of positive and negative votes, depending on parameter selection). We omit the details of this correspondence for lack of space.

The gradual valuation semantics for BAFs, given in [9], is also restricted to acyclic BAFs. This semantics can be presented as follows, for direct comparison with our semantics. Let the *equivalent strength aggregation function* be defined as $\sigma' : [-1,1]^* \to [0, \infty)$, where for $S = (v_1, \ldots, v_n) \in [-1,1]^*$:

$$\sigma'(S) = \sum_{i=1}^{n} \frac{v_i + 1}{2}$$

Let the *equivalent mediating function* be defined as $\mu' : [0,\infty) \times [0,\infty) \to [-1,1]$, where for $v_a, v_s \in [0, \infty)$:

$$\mu'(v_a, v_s) = \frac{1}{1 + v_a} - \frac{1}{1 + v_s}$$

Finally, let the *equivalent score function* be defined as our score function, but using the equivalent strength aggregation and mediating function, namely as $\mathcal{SF}' : \mathcal{X} \to \mathbb{I}$ where, for any $a \in \mathcal{X}$:

$$\mathcal{SF}'(a) = \mu'(\sigma'(SEQ_{\mathcal{SF}'}(\mathcal{R}^-(a))), \sigma'(SEQ_{\mathcal{SF}'}(\mathcal{R}^+(a))))$$

where if $(a_1, \ldots, a_n)$ is an arbitrary permutation of the $(n \geq 0)$ attackers in $\mathcal{R}^-(a)$, $SEQ_{\mathcal{SF}'}(\mathcal{R}^-(a)) = (\mathcal{SF}'(a_1), \ldots, \mathcal{SF}'(a_n))$ (similarly for supporters).

For lack of space we omit to prove that this formulation of gradual valuation is indeed equivalent to the original formulation in [9]. Here, we note however that our semantics differs from this formulation of gradual valuation. Indeed, the equivalent of Proposition 1 is not held by σ', because each argument with maximum strength incrementally increases the aggregated attacking/supporting strength towards infinity, and the equivalent of Proposition 2 is not held by $\mathcal{SF}'$, since, then minimum (resp. maximum) value, -1 (resp. 1), can only be achieved if the attacking (resp. supporting) component is ∞ and the supporting (resp. attacking) component is 0, requiring an infinite sequence of attackers (resp. supporters) and no supporters (resp. attackers).

4. Conclusion

We have defined a quantitative semantics for evaluating the strength of arguments in (possibly non-acyclic) BAFs (including both attacks and supports). The semantics is defined as an adaptation of the DF-QuAD algorithm [6] assuming that all arguments have a base score of 0.5 and, differently from the original algorithm, it can deal with any BAFs rather than restricted (acyclic) BAFs corresponding to QuAD frameworks [7]. The semantics generalises, by dealing with supports too, the social model semantics of [8] when arguments' social support is 0.5. The semantics differs from other semantics for BAFs, and in particular the gradual valuation of [9].

We have sketched and conjectured, resp., existence and uniqueness of strength of arguments obtained by our semantics. We leave formal proofs to future work, alongside proving other properties of the semantics and relationship to other existing semantics.

The semantics defines the strength of arguments, within the $[0,1]$ interval of the real numbers, as solutions of non-linear equations, and can be computed, algorithmically, by any solver for systems of non-linear equations, e.g. via Newton's method. The definition of a system supporting the computation is left as future work.

References

[1] C. Cayrol, & M. Lagasquie-Schiex, On the Acceptability of Arguments in Bipolar Argumentation Frameworks. *Symbolic and Quantitative Approaches to Reasoning with Uncertainty, 8th European Conf., ECSQARU 2005, Barcelona, Spain, July 6-8, 2005, Proc.*, 378-389.

[2] P.M. Dung, On the acceptability of arguments and its fundamental role in nonmonotonic reasoning, logic programming and n-person games. *Artif. Intell.*, **77(2)** (1995). 321-358.

[3] P.A. Matt & F. Toni, A Game-Theoretic Measure of Argument Strength for Abstract Argumentation. *Proc. of 11th European Conf. on Logics in Artificial Intelligence (JELIA 2008)*, 285-297.

[4] A. Hunter & M. Thimm, On Partial Information and Contradictions in Probabilistic Abstract Argumentation. *In Principles of Knowledge Representation and Reasoning: Proc. of the 15th Int. Conf., KR 2016, Cape Town, South Africa, April 25-29, 2016.*, 53-62.

[5] L. Amgoud, J. Ben-Naim, D. Doder & S. Vesic, Ranking Arguments With Compensation-Based Semantics. *In Principles of Knowledge Representation and Reasoning: Proc. of the 15th Int. Conf., KR 2016, Cape Town, South Africa, April 25-29, 2016.*, 12-21.

[6] A. Rago, F. Toni, M. Aurisicchio & P. Baroni, Discontinuity-Free Decision Support with Quantitative Argumentation Debates. *In Principles of Knowledge Representation and Reasoning: Proc. of the 15th Int. Conf., KR 2016, Cape Town, South Africa, April 25-29, 2016.*, 63-73.

[7] P. Baroni, M. Romano, F. Toni, M. Aurisicchio & G. Bertanza, Automatic Evaluation of Design Alternatives with Quantitative Argumentation. *Argument & Computation*, **6(1)** (2015), 24-49.

[8] J. Leite & J. Martins, Social Abstract Argumentation. *Proc. of the 22nd Int. Joint Conf. on Artificial Intelligence, IJCAI 2011*, 2287-2292.

[9] L. Amgoud, C. Cayrol, M. Lagasquie-Schiex & P. Livet, On bipolarity in argumentation frameworks. *Int. J. Intell. Syst.*, **23(10)** (2008), 1062-1093.

[10] E. Dyer, A Fixed Point Theorem. *Proc. of the American Mathematical Society, 1956*, 7(4), 662-672.