Preliminary Study on Reinstatement Labelling for Weighted Argumentation Frameworks*

Stefano Bistarelli[1][0000−0001−7411−9678] and Carlo Taticchi[2][10000−0003−1260−4672]

[1] Università degli Studi di Perugia, Italy
stefano.bistarelli@unipg.it
[2] Gran Sasso Science Institute, Italy
carlo.taticchi@gssi.it

Abstract. Argumentation Theory allows for reasoning with uncertain and controversial information, and provides tools for deciding which arguments (for instance, of a debate) can be accepted together. The strength of an argument and its attacks can be expressed through weighted argumentation frameworks; in this case, the selection criteria, called semantics, used to identify the sets of acceptable arguments, need to take into account the information given by the weights. In this paper, we conduct an initial study on a novel labelling semantics for weighted argumentation frameworks, extending and generalising the crisp one.

Keywords: Argumentation · Weighted Frameworks · Labelling.

1 Introduction

Generalised notions of defence for weighted AFs are studied in several works, each with a different definition for the notion of weights. In two recent papers [3,4] the attacks from an argument to a set of arguments are grouped together as if they were a unique attack; such a weighted notion of defence also generalises the approaches of [7] and [9]. Through the reasoning on the acceptability of the arguments according to a notion of defence, one can divide the set of arguments into two separated subsets, respectively containing accepted and non-accepted arguments. However, for certain applications of argumentation (especially those in which defeating an argument leads to the reinstatement of another one [6]), it is convenient to consider more degrees of acceptability in order for one to be able to further differentiate among arguments. The labelling, defined in [6], refines the concept of acceptable argument and builds on the classical semantics for providing an additional acceptance status.

In this work, we extend the notion of labelling to weighted argumentation frameworks and we provide a definition that is parametric to a chosen notion of defence [3,7,9] and that corresponds to the original labelling [6] when a boolean semiring is used. For each weighted semantics, we give the conditions under which a labelling corresponds to a set of extensions.

* This work was partially supported by "Argumentation 360" (Ricerca di Base 2017-2019) and "RACRA18" (Ricerca di Base 2018-2020).

2 Background

An abstract argumentation framework [8] is a pair $\langle \mathcal{A}, \mathcal{R} \rangle$ where $\mathcal{A}$ is a set of arguments and $\mathcal{R}$ is a binary relation on $\mathcal{A}$. Consider two arguments a, b belonging to an AF. We denote with $(a, b) \in \mathcal{R}$ (or simply $a \to b$) an attack from a to b; we can also say that b is defeated by a. We define the sets of arguments that attack (and that are attacked by) another argument as follows.

Definition 1 (Attacks). *Let $F = \langle \mathcal{A}, \mathcal{R} \rangle$ be an AF, $a \in \mathcal{A}$ and $A \subseteq \mathcal{A}$. We define the sets $a^+ = \{b \in \mathcal{A} \mid a \to b\}$, $a^- = \{b \in \mathcal{A} \mid b \to a\}$, $A^+ = \cup\{a^+ \mid a \in A\}$ and $A^- = \cup\{a^- \mid a \in A\}$.*

In order for b to be acceptable, we require that every argument that defeats b is defeated in turn by some other argument of the AF.

Definition 2 (Acceptable argument [8]). *Given an AF $F = \langle \mathcal{A}, \mathcal{R} \rangle$, an argument $a \in \mathcal{A}$ is acceptable with respect to $D \subseteq \mathcal{A}$ iff $\forall b \in \mathcal{A}$ such that $(b, a) \in \mathcal{R}$, $\exists c \in D$ such that $(c, b) \in \mathcal{R}$, and we say that b is **defended** by D.*

By using the notion of defence as a criterion for distinguishing acceptable arguments, one can further refine the set of selected arguments through semantics.

Definition 3 (Extension-based semantics). *Let $F = \langle \mathcal{A}, \mathcal{R} \rangle$ be an AF. A set $E \subseteq \mathcal{A}$ is conflict-free in F iff there are no $a, b \in \mathcal{A}$ such that $(a, b) \in \mathcal{R}$. A conflict-free subset E is then: i) admissible if each $a \in E$ is defended by E; ii) complete if it is admissible and $\forall a \in \mathcal{A}$ defended by E, $a \in E$.*

2.1 Labelling-Based Semantics

The authors in [1] and [6] describe how to assign labels to the arguments of an AF in such a way that the set of arguments is partitioned in three subsets, each representing a different degree of acceptance.

Definition 4 (Labelling). *Let $F = \{\mathcal{A}, \mathcal{R}\}$ be an AF. A labelling is a total function $L : \mathcal{A} \to \{\texttt{IN}, \texttt{OUT}, \texttt{UNDEC}\}$. For any $A \subseteq \mathcal{A}$, we denote $A|_{\texttt{IN}}$, $A|_{\texttt{OUT}}$ and $A|_{\texttt{UNDEC}}$ the set of all the arguments labelled `IN`, `OUT` and `UNDEC` by L, respectively.*

Given a labelling L, it is also possible to identify a correspondence with the extension-based semantics [1]. More in detail, the set of `IN` arguments coincides with an extension. We rephrase the semantics of Definition 3 as follows.

Definition 5 (Labelling-based semantics). *Let L be a labelling of an AF $F = \langle \mathcal{A}, \mathcal{R} \rangle$ and $a \in \mathcal{A}$. Then L is:*

- *conflict-free iff* $L(a) = \texttt{IN} \implies a^-|_{\texttt{IN}} = \emptyset \ \wedge \ L(a) = \texttt{OUT} \implies a^-|_{\texttt{IN}} \neq \emptyset$
- *admissible iff* $L(a) = \texttt{IN} \implies a^- = a^-|_{\texttt{OUT}} \ \wedge \ L(a) = \texttt{OUT} \implies a^-|_{\texttt{IN}} \neq \emptyset$
- *complete iff* $L(a) = \texttt{IN} \iff a^- = a^-|_{\texttt{OUT}} \ \wedge \ L(a) = \texttt{OUT} \iff a^-|_{\texttt{IN}} \neq \emptyset$

Note that a complete labelling coincides with the definition of reinstatement labelling given in [6].

2.2 Weighted Argumentation Frameworks

Dung's argumentation frameworks can be extended by associating the attacks with a weight that represents the support of the relation.

Definition 6 (WAF). *A weighted argumentation framework is a triple $W = \langle \mathcal{A}, \mathcal{R}, w \rangle$ where $\langle \mathcal{A}, \mathcal{R} \rangle$ is a Dung-style abstract argumentation framework and $w : \mathcal{R} \to \mathbb{R}^+$ is a function assigning positive real valued weights to attacks.*

Notions of weighted defence have been proposed in [4,7,9]. In this paper, we use the interpretation of [4] (which generalises the other two definitions [3]), where WAFs are equipped with a commutative semiring that provides the operation for composing the weights. The acceptability of an argument is determined by comparing the compositions of attacks with the composition of defences.

Definition 7 (c-semirings). *A commutative semiring is a tuple $\mathbb{S} = \langle S, \oplus, \otimes, \bot, \top \rangle$ such that S is a set, $\top, \bot \in S$, and $\oplus, \otimes : S \times S \to S$ are binary operators making the triples $\langle S, \oplus, \bot \rangle$ and $\langle S, \otimes, \top \rangle$ commutative monoids, satisfying* i) *$\forall s, t, u \in S. s \otimes (t \oplus u) = (s \otimes t) \oplus (s \otimes u)$, and* ii) *$\forall s \in S. s \otimes \bot = \bot$.*

Different c-semirings can represent different notions of defence for WAFs, by using the operators $\oplus$ and $\otimes$ for obtaining an ordering among the weights in S. Common instances of c-semirings are $\mathbb{S}_{boolean} = \langle \{false, true\}, \vee, \wedge, false, true \rangle$ and $\mathbb{S}_{weighted} = \langle \mathbb{R}^+ \cup \{+\infty\}, min, +, +\infty, 0 \rangle$. We denote with WAF$_{\mathbb{S}}$ a WAF endowed with a c-semirings and we call it a semiring-based WAF.

Definition 8 (WAF$_{\mathbb{S}}$). *A semiring-based WAF is a quadruple $\langle \mathcal{A}, \mathcal{R}, W, \mathbb{S} \rangle$, where $\mathbb{S}$ is a c-semiring $\langle S, \oplus, \otimes, \bot, \top \rangle$, $\mathcal{A}$ is a set of arguments, R the attack binary-relation on $\mathcal{A}$, and $W : \mathcal{A} \times \mathcal{A} \longrightarrow S$ is a binary function. Given $a, b \in \mathcal{A}$ and $R(a, b)$, then $W(a, b) = s$ means that a attacks b with a weight $s \in S$. Moreover, we require that $R(a, b)$ iff $W(a, b) <_{\mathbb{S}} \top$.*

Given a WAF$_{\mathbb{S}}$, we can evaluate the overall weight of all the attacks from a set of arguments towards another set through the **composition** operator $\otimes$ [3]. In particular, we say that a set of arguments $\mathcal{B}$ attacks a set of arguments $\mathcal{D}$, and the weight of such attack is $k \in S$, if $W(\mathcal{B}, \mathcal{D}) = \bigotimes_{b \in \mathcal{B}, d \in \mathcal{D}} W(b, d) = k$. The notion of weighted defence (or w-defence) can then be expressed as follows.

Definition 9 (w-defence). *Let $F = \langle \mathcal{A}, \mathcal{R}, W, \mathbb{S} \rangle$ be a WAF$_{\mathbb{S}}$. $\mathcal{B} \subseteq \mathcal{A}$ w-defends $b \in \mathcal{A}$ iff $\forall a \in \mathcal{A}$ such that $R(a, b)$, we have $W(a, \mathcal{B} \cup \{b\}) \geq_{\mathbb{S}} W(\mathcal{B}, a)$.*

According to [4], by using the notion of w-defence for checking the acceptability of the arguments in the weighted framework, it is possible to redefine all the extension-based semantics presented in Definition 3.

Definition 10 (Extension-based semantics for WAF$_{\mathbb{S}}$). *Given a WAF$_{\mathbb{S}}$ $F = \langle \mathcal{A}, \mathcal{R}, W, \mathbb{S} \rangle$, a subset $\mathcal{B} \subseteq \mathcal{A}$ is w-conflict-free if $W(\mathcal{B}, \mathcal{B}) = \top$. Then $\mathcal{B}$ is: i) w-admissible iff it is w-conflict-free and the arguments in $\mathcal{B}$ are w-defended by $\mathcal{B}$ from the arguments in $\mathcal{A} \setminus \mathcal{B}$; ii) w-complete iff it is a w-admissible extension and each argument $b \in \mathcal{A}$ such that $\mathcal{B} \cup \{b\}$ is w-admissible belongs to $\mathcal{B}$;*

3 Labelling for Weighted AFs

Contrary to classical AFs, for which we can use the procedure in [6] for assigning labels to the arguments in such a way that there is a correspondence between the labelling and the set of extension, no work on this direction has been done for what concerns the weighted case. We extend the notion of labelling introduced in [6] to a $\mathrm{WAF}_{\mathbb{S}}$ and we give the conditions for determining whether a labelling corresponds to a certain extension. In order to incorporate the notion of weighted defence in the labelling, we need to take into account the strength of the attack relations: according to the definition of collective defence [3] we need to know the resulting strength of the composition of all the attacks towards an argument.

Definition 11 (Labelling for WAFs). *Let $F = \langle \mathcal{A}, \mathcal{R}, W, \mathbb{S} \rangle$ be a $WAF_{\mathbb{S}}$. A labelling L of F is a total function $L : \mathcal{A} \to \{\mathtt{IN}, \mathtt{OUT}, \mathtt{UNDEC}\}$. We also define, for each argument, the weight of attacks, incoming into and outgoing from an argument, as $w_{a^-|_{\mathtt{IN}}} = W(a^-|_{\mathtt{IN}}, a)$ and $w_{a^+|_{\mathtt{IN}}} = W(a, a^+|_{\mathtt{IN}})$.*

In our system, `OUT` arguments are associated with the $\bigotimes$ of the incoming attacks. An argument a with label `OUT` is attacked by the arguments in $a^-|_{\mathtt{IN}}$ with a total strength that is expressed by w_a. In the following, we give a characterisation of the weighted semantics through the notion of labelling of WAFs. The intuition behind this representation is that when an argument a attacked by an `OUT` b cannot be labelled `IN` because of another `IN` argument that is "consuming" the attacks of the defending arguments towards b, then a is labelled `UNDEC`. The w-conflict-free labelling coincides with the conflict-free labelling. Indeed, since no attacks are allowed within a conflict-free set of arguments, one does not need to consider the weights.

Definition 12 (Labelling-based semantics for WAFs). *Let L be a labelling of a $WAF_{\mathbb{S}}$ $F = \langle \mathcal{A}, \mathcal{R}, W, \mathbb{S} \rangle$ and $a \in \mathcal{A}$. Then*

- *L is a w-admissible labelling for F iff:*
  - $L(a) = \mathtt{IN} \implies a^- = a^-|_{\mathtt{OUT}} \ \wedge \ \forall b \in a^-.\ w_{b^-|_{\mathtt{IN}}} \leq_{\mathbb{S}} w_{b^+|_{\mathtt{IN}}}$
  - $L(a) = \mathtt{OUT} \implies w_{a^-|_{\mathtt{IN}}} <_{\mathbb{S}} \top$
- *L is a w-complete labelling for F iff:*
  - $L(a) = \mathtt{IN} \iff a^- = a^-|_{\mathtt{OUT}} \ \wedge \ \forall b \in a^-.\ w_{b^-|_{\mathtt{IN}}} \leq_{\mathbb{S}} w_{b^+|_{\mathtt{IN}}}$
  - $L(a) = \mathtt{OUT} \iff w_{a^-|_{\mathtt{IN}}} <_{\mathbb{S}} \top$

The condition $w_{b^-|_{\mathtt{IN}}} \leq_{\mathbb{S}} w_{b^+|_{\mathtt{IN}}}$ for the w-admissible labelling makes sure that the composition of the attacks of the arguments defending a is stronger than the attack of b. For an argument to be `OUT`, we require $w_{a^-|_{\mathtt{IN}}} <_{\mathbb{S}} \top$, that is to say that there must exist at least an attack coming from an `IN` argument (as for the classical admissible labelling). The definition of the w-complete labelling is similar to the w-admissible one, with the exception that the conditions given for `IN` arguments are both necessary and sufficient. The weighted semantics generalise the classical case: all the labellings for $\mathrm{WAF}_{\mathbb{S}}$ correspond to the respective Dung semantics when the framework is instantiated with a boolean semiring.

Remark 1. Let $F = \langle \mathcal{A}, \mathcal{R}, W, \mathbb{S} \rangle$ be a $\mathrm{WAF}_{\mathbb{S}}$ where $\mathbb{S}$ is boolean, and L a labelling of F. If L is a w-admissible (respectively w-complete) labelling, then L is an admissible (respectively complete) labelling of F.

4 Conclusion and Future Work

The definitions of the labelling-based semantics for WAFs, that we give in Section 3, do not include conditions for the `UNDEC` since they are obtained from `IN` and `OUT` arguments. In this sense, we would like to investigate the possible advantages of giving explicit conditions for labelling the `UNDEC` arguments, similarly to what is done in [10] for classical AFs. We also plan to extend the presented labelling to all the semantics (including preferred, stable and grounded). An interesting study could then be carried out on the *don't care* and *don't know* labels, that are used in [2] as further differentiation of `UNDEC` arguments. In our context, the difference between the two labels could be made more continuous by considering the weight on the attack relations. Moreover, the weights $w_{a^-|_{\texttt{IN}}}$ and $w_{a^+|_{\texttt{IN}}}$ of the attacks towards and from each argument (other than the respective weights for `OUT` and `UNDEC`) could be used for defining a ranking-based semantics for WAFs. Finally, we want to implement the labelling-based semantics for WAFs in ConArg [5], an online tool for dealing with argumentation problems.

References

1. Baroni, P., Caminada, M., Giacomin, M.: An introduction to argumentation semantics. Knowledge Eng. Review **26**(4), 365–410 (2011)
2. Baroni, P., Giacomin, M., Liao, B.S.: I don't care, I don't know ... I know too much! On Incompleteness and Undecidedness in Abstract Argumentation. In: Advances in Knowledge Representation, Logic Programming, and Abstract Argumentation. Lecture Notes in Computer Science, vol. 9060, pp. 265–280. Springer (2015)
3. Bistarelli, S., Rossi, F., Santini, F.: A Collective Defence Against Grouped Attacks for Weighted Abstract Argumentation Frameworks. In: Proceedings of FLAIRS 2016. pp. 638–643. AAAI Press (2016)
4. Bistarelli, S., Rossi, F., Santini, F.: A novel weighted defence and its relaxation in abstract argumentation. Int. J. Approx. Reasoning **92**, 66–86 (2018)
5. Bistarelli, S., Santini, F.: Modeling and solving afs with a constraint-based tool: Conarg. In: Theorie and Applications of Formal Argumentation - First International Workshop, TAFA 2011. Revised Selected Papers. Lecture Notes in Computer Science, vol. 7132, pp. 99–116. Springer (2011)
6. Caminada, M.: On the Issue of Reinstatement in Argumentation. In: Proceedings of JELIA 2006. Lecture Notes in Computer Science, vol. 4160, pp. 111–123. Springer (2006)
7. Coste-Marquis, S., Konieczny, S., Marquis, P., Ouali, M.A.: Weighted Attacks in Argumentation Frameworks. In: Principles of Knowledge Representation and Reasoning. Proceedings of KR 2012. AAAI Press (2012)
8. Dung, P.M.: On the acceptability of arguments and its fundamental role in nonmonotonic reasoning, logic programming and n-person games. Artificial Intelligence **77**(2), 321–357 (Sep 1995)
9. Martínez, D.C., García, A.J., Simari, G.R.: An Abstract Argumentation Framework with Varied-Strength Attacks. In: Principles of Knowledge Representation and Reasoning: Proceedings of KR 2008. pp. 135–144. AAAI Press (2008)
10. Modgil, S., Caminada, M.: Proof Theories and Algorithms for Abstract Argumentation Frameworks. In: Argumentation in Artificial Intelligence, pp. 105–129. Springer (2009)