

Extending Labelling Semantics to Weighted Argumentation Frameworks

Stefano Bistarelli

University of Perugia, Italy
stefano.bistarelli@unipg.it

Carlo Taticchi

Gran Sasso Science Institute, Italy
carlo.taticchi@gssi.it

Abstract

Argumentation Theory provides tools for both modelling and reasoning with controversial information and is a methodology that is often used as a way to give explanations to results provided using machine learning techniques. In this context, labelling-based semantics for Abstract Argumentation Frameworks (AFs) allow for establishing the acceptability of sets of arguments, dividing them into three partitions: acceptable, rejected and undecidable (instead of classical Dung two sets IN and OUT partitions). This kind of semantics have been studied only for classical AFs, whilst the more powerful weighted and preference-based framework has been not studied yet. In this paper, we define a novel labelling semantics for Weighted Argumentation Frameworks, extending and generalising the crisp one.

Introduction

Argumentation and its applications are receiving increasing interest in many fields of AI. For instance, argumentative processes are used in (Lawrence et al. 2017) to interpret online debates, while in (Walton and Koszowy 2017) an argumentation system is devised to support expert opinion. Argumentation is also used to aid machine learning (see (Coarascu and Toni 2016) for a survey) for both improving performances (e.g., classification accuracy) and providing explanations to the results. Argumentation problems are modelled through Abstract Argumentation Frameworks (AFs in short) (Dung 1995), that consist of directed graphs in which the nodes are arguments that contain abstract information and the edges represent attack relations.

The acceptability of an argument of an AF can then be established following different criteria, formalised through the extension-based (Dung 1995) and the labelling-based (Caminada 2006) semantics. 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 acceptable and non-acceptable arguments. However, for certain applications (especially those in which defeating an argument leads to the reinstatement of another one) it is convenient to consider

more degrees of acceptability in order to be able to further differentiate among arguments. The labelling defined in (Caminada 2006) refines the concept of acceptable argument and builds on the classical semantics for providing an additional acceptance status through the assignment of labels to the arguments.

To increase the expressiveness of AFs, attack relations between arguments can be endowed with a value (a weight) which indicates the strength of the attacks themselves. In this kind of frameworks, called weighted AFs, the acceptability criteria for the arguments also need to consider the weight of incoming and outgoing attacks. Three main approaches have been proposed in the literature: in (Martínez, García, and Simari 2008) the attacks are considered individually and the acceptability status of an argument is determined through a one by one comparison on the strengths of the relations; in (Coste-Marquis et al. 2012) each attack towards an argument can be defended from a group of arguments, with an overall strength obtained by aggregating the single strengths of the counter-attacks coming from that group; the method used in (Bistarelli, Rossi, and Santini 2018), instead, aggregates the strengths of both the attacks conducted towards an argument and the defences for that argument. In all these works, extension-based semantics have been used to identify sets of acceptable arguments. A first study of labelling semantics for weighted AFs was conducted in (Bistarelli and Taticchi 2020), solely relying on the notion of weighted defence given in (Bistarelli, Rossi, and Santini 2018). In this work, we study labelling semantics for weighted AFs also considering the other two above-mentioned collective defences, and we give the conditions under which a labelling corresponds to a set of extensions.

Preliminaries

In this section we recall the formal definition of AF and the related semantics introduced by Dung (Dung 1995), together with the main definitions for Weighted AFs and the different notions proposed for the acceptability of arguments.

Abstract Argumentation Frameworks

First of all, we recall the formal definition for an AF (Dung 1995).

Definition 1 (AFs). An *Abstract Argumentation Framework* 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 $a \rightarrow b$) 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 2 (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 $a^+ = \{b \in \mathcal{A} \mid a \rightarrow b\}$, $a^- = \{b \in \mathcal{A} \mid b \rightarrow a\}$, $A^+ = \cup\{a^+ \mid a \in A\}$ 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 3 (Acceptable argument). Given an AF $F = \langle \mathcal{A}, \mathcal{R} \rangle$, an argument $a \in \mathcal{A}$ is acceptable with respect to $D \subseteq \mathcal{A}$ if and only if $\forall b \in \mathcal{A}$ such that $b \in a^-$, $\exists c \in D \cap b^-$, and we say that a is **defended** by D .

Using the notion of defence as a criterion for distinguishing acceptable arguments in the framework, one can further refine the set of selected “good” arguments.

Definition 4 (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 if and only if there are no $a, b \in \mathcal{A}$ such that $(a, b) \in \mathcal{R}$. A *conflict-free subset* E is then

- *admissible*, if each $a \in E$ is defended by E ;
- *complete*, if it is admissible and $\forall a \in \mathcal{A}$ defended by E , $a \in E$;
- *stable*, if $E \cup E^+ = \mathcal{A}$;
- *preferred*, if it is admissible and it is maximal (with respect to set inclusion);
- *grounded*, if it is complete and it is minimal (with respect to set inclusion).

Weighted Argumentation Frameworks

In classical AFs it is not possible to further diversify the relations among arguments, and every attack has the same “strength”, that is, the existence or not of an attack is the only thing that matters in determining the semantics. To overcome this limit, Dung’s AFs have been extended to Weighted AFs (WAFs) by associating the attacks with a weight that represents the support of the relation (Dunne et al. 2011).

To analyse a WAF in terms of sets of extensions, a definition of defence is required that encompasses the notion of weighted attack relations. The acceptability of an argument is then determined by comparing (the compositions of) the attacks with (the composition of) the defences.

Different possible definitions can be considered: for instance, in (Martínez, García, and Simari 2008) the relative strength of the attacks is used to determine if some defence constraints are satisfied. On the other hand, the authors of (Coste-Marquis et al. 2012) define a WAF as a triple $\langle \mathcal{A}, \mathcal{R}, w \rangle$, where $w : \mathcal{A} \times \mathcal{A} \rightarrow \mathbb{N}$ is a function assigning a natural number (representing a weight) to each attack. Finally, in (Bistarelli, Rossi, and Santini 2018) the framework

is equipped with a c-semiring (Bistarelli and Gadducci 2006; Bistarelli, Montanari, and Rossi 1997) that provides the operation for composing the weights in order to estimate the effectiveness of a defence. The use of a generic semiring permits to instantiate the aggregation operator and capture any important properties (Bistarelli, Rossi, and Santini 2018)

Definition 5 (c-semirings). A *c-semiring* is a tuple $\mathbb{S} = \langle S, \oplus, \otimes, \perp, \top \rangle$ such that S is a set, $\top, \perp \in S$, and $\oplus, \otimes : S \times S \rightarrow S$ are binary operators making the triples $\langle S, \oplus, \perp \rangle$ and $\langle S, \otimes, \top \rangle$ commutative monoids (semi-groups with identity), satisfying i) $\forall s, t, u \in S. s \otimes (t \oplus u) = (s \otimes t) \oplus (s \otimes u)$ (distributivity), and ii) $\forall s \in S. s \otimes \perp = \perp$ (annihilator). Moreover, we have that $\forall s, t \in S. s \oplus (s \otimes t) = s$ (absorptiveness). The operator \oplus also defines a preference relation $\leq_{\mathbb{S}}$ over the set S , such that $a \leq_{\mathbb{S}} b \iff a \oplus b = b$, for $a, b \in S$.

Some common instances of c-semirings are:

- $\mathbb{S}_{\text{boolean}} = \langle \{\text{false}, \text{true}\}, \vee, \wedge, \text{false}, \text{true} \rangle$
- $\mathbb{S}_{\text{fuzzy}} = \langle [0, 1], \max, \min, 0, 1 \rangle$
- $\mathbb{S}_{\text{weighted}} = \langle \mathbb{R}^+ \cup \{+\infty\}, \min, +, +\infty, 0 \rangle$

Different c-semirings can represent different notions of defence for WAF, by using the operators \oplus and \otimes for obtaining an ordering among the values in S . For simplicity, we refer to these values as weights. Note that the element \top of the c-semiring (e.g., 0 for the weighted and *true* for the boolean) coincides with having no relation between two arguments. We denote with $\text{WAF}_{\mathbb{S}}$ a WAF endowed with a c-semirings \mathbb{S} and we call it a semiring-based WAF.

Definition 6 ($\text{WAF}_{\mathbb{S}}$). A *semiring-based WAF* is a quadruple $\langle \mathcal{A}, \mathcal{R}, W, \mathbb{S} \rangle$, where \mathbb{S} is a c-semiring $\langle S, \oplus, \otimes, \perp, \top \rangle$, \mathcal{A} is a set of arguments, \mathcal{R} the attack binary-relation on \mathcal{A} , and $W : \mathcal{A} \times \mathcal{A} \rightarrow S$ 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 $R(a, b)$ if and only if $W(a, b) <_{\mathbb{S}} \top$.

Given a $\text{WAF}_{\mathbb{S}}$, we can evaluate the overall weight of all the attacks from a set of arguments towards another set through the **composition** operator \otimes of the c-semiring \mathbb{S} (Bistarelli, Rossi, and Santini 2016; Bistarelli and Santini 2017). In particular, we use \otimes to indicate the \otimes operator on a set of values (indeed \otimes is a binary operator that composes two weights).

Definition 7 (Attacks). Let $F = \langle \mathcal{A}, \mathcal{R}, W, \mathbb{S} \rangle$ be a $\text{WAF}_{\mathbb{S}}$. 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 previous definition also allows composing the attacks from a set of arguments to another single argument, and from a single argument towards a set of arguments. The frameworks in (Martínez, García, and Simari 2008) and (Coste-Marquis et al. 2012) can be then described as instances of a $\text{WAF}_{\mathbb{S}}$. For instance, the attack strength used in (Martínez, García, and Simari 2008) corresponds to the strongest weight among all the counter-attacks and can be

obtained through a fuzzy semiring. The approach in (Coste-Marquis et al. 2012), instead, uses an aggregation function (e.g., + or max) to obtain the overall strength of the attacks coming from the defending arguments; also in this case, a semiring can be selected according to the used aggregation function.

The notion of weighted defence (or w -defence) can then be expressed in terms of preferences over the weighted attack relations. In particular, if we consider the defence of (Martínez, García, and Simari 2008), we obtain the defence \mathbb{D}_1 ; alternatively, we can use the definition in (Coste-Marquis et al. 2012) to obtain \mathbb{D}_2 ; the notion introduced in (Bistarelli, Rossi, and Santini 2016), finally, generalises the other two approaches and provides the defence \mathbb{D}_3 .

Definition 8 (w -defence). *Let $F = \langle \mathcal{A}, \mathcal{R}, W, \mathbb{S} \rangle$ be a $WAF_{\mathbb{S}}$. Then $\mathcal{B} \subseteq \mathcal{A}$ w -defends $b \in \mathcal{A}$ if and only if $\forall a \in \mathcal{A}$ such that $R(a, b)$,*

- \mathbb{D}_1 : $\exists c \in \mathcal{B} \mid W(a, b) \geq_{\mathbb{S}} W(c, a)$, or
- \mathbb{D}_2 : $W(a, b) \geq_{\mathbb{S}} W(\mathcal{B}, a)$, or
- \mathbb{D}_3 : $W(a, \mathcal{B} \cup \{b\}) \geq_{\mathbb{S}} W(\mathcal{B}, a)$.

By using one among \mathbb{D}_1 , \mathbb{D}_2 and \mathbb{D}_3 for checking the acceptability of the arguments in the weighted framework, it is possible to redefine all the extension-based semantics presented in Definition 4.

Definition 9 (Extension-based semantics for $WAF_{\mathbb{S}}$). *Given a $WAF_{\mathbb{S}}$ $F = \langle \mathcal{A}, \mathcal{R}, W, \mathbb{S} \rangle$, a subset of arguments $\mathcal{B} \subseteq \mathcal{A}$ is w -conflict-free if $W(\mathcal{B}, \mathcal{B}) = \top$. A w -conflict-free subset \mathcal{B} is then:*

- w -admissible, if \mathcal{B} w -defends all its elements from the arguments in $\mathcal{A} \setminus \mathcal{B}$;
- w -complete, if it is w -admissible and each argument $b \in \mathcal{A}$ such that $\mathcal{B} \cup \{b\}$ is w -admissible belongs to \mathcal{B} ;
- w -stable, if it is w -admissible and $\mathcal{B} \cup \mathcal{B}^+ = \mathcal{A}$;
- w -preferred, if it is a maximal (with respect to set inclusion) w -admissible subset of \mathcal{A} ;
- w -grounded, if it is the maximal (with respect to set inclusion) w -admissible extension included in the intersection of w -complete extensions.

Labelling for Weighted AFs

The work in (Caminada 2006) describes how to assign labels to the arguments of an AF in such a way that the set of arguments is partitioned into three subsets, each representing a different degree of acceptance: IN, OUT and UNDEC. Given a labelling L , it is possible to identify a correspondence with the extension-based semantics (Baroni, Caminada, and Giacomin 2011; 2018): for instance, the set of IN arguments coincides with an extension of the acceptable semantics. In every labelling of the various semantics, arguments for which not every attacker is labelled OUT and no attacker is labelled IN are labelled UNDEC. The admissible labelling that we consider coincides with the interpretation given in (Caminada 2014), where IN arguments can attack both OUT and UNDEC arguments. Different definitions of labelling (as for instance the one given in (Jakobovits and Vermeir 1999)) force arguments attacked by an IN to be OUT.

However, nothing changes in terms of extensions, since the set of IN arguments remains the same.

We extend the notion of labelling introduced in (Caminada 2006) to weighted AFs. In particular, we consider a $WAF_{\mathbb{S}}$ and we provide a definition for the labelling. Furthermore, we give the conditions for determining whether a labelling corresponds to a certain extension.

Definition 10 (Labelling for $WAF_{\mathbb{S}}$). *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 $L : \mathcal{A} \rightarrow \{IN, OUT, UNDEC\}$. For any $A \subseteq \mathcal{A}$, we denote $A|_{IN}$, $A|_{OUT}$ and $A|_{UNDEC}$ the set of all the arguments labelled IN, OUT and UNDEC by L , respectively.*

$\mathbb{D}_1, \mathbb{D}_2$ and \mathbb{D}_3 demand specific considerations for computing the (overall) weight of the attack relations that, as we can see in Figure 1, lead to different outcomes in terms of acceptable arguments.

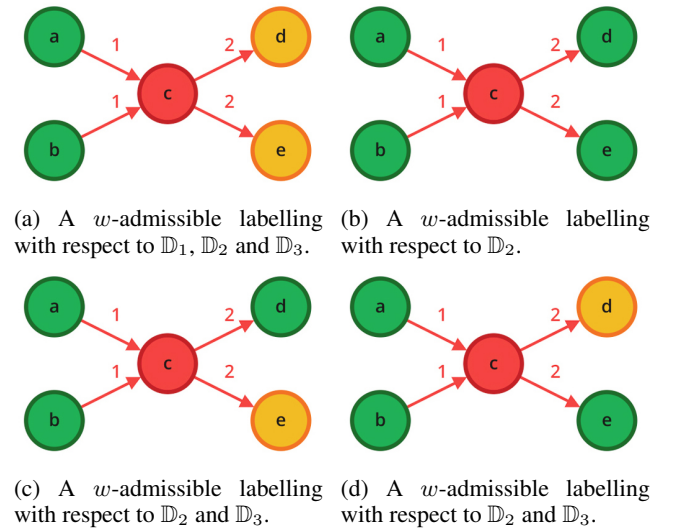


Figure 1: Example of labellings on a $WAF_{\mathbb{S}}$ with a weighted semiring where IN arguments are highlighted in green, UNDEC in yellow, and OUT in red.

The attacks in (Martínez, García, and Simari 2008) are ordered by their strength and it is sufficient to compare the weight of two attacks to establish whether an argument is defended (and so labelled IN) or not. Following (Coste-Marquis et al. 2012), we need to know the strength resulting from the composition \otimes of all the attacks coming from the defending arguments towards the attacker. In particular, an argument b with label OUT is attacked by the arguments in $b^-|_{IN}$ with a total strength that is expressed by $W(b^-|_{IN}, b)$. According to the definition of collective weighted defence given in (Bistarelli, Rossi, and Santini 2016), a set of argument is defended from an attacker b only if the \otimes of all the defending arguments is stronger than the \otimes of the attacks coming from b . This means that the strength of the attacks of the defending arguments is distributed among the defended arguments, so it is not guaranteed for two arguments that are separately w -defended to still be w -defended when considered together (this is what happens in the example in Fig-

ure 1c and 1d with arguments d and e). The intuition behind this representation is that when an argument a is attacked by an OUT b and cannot be labelled IN because another IN argument is “consuming” the attacks of the defending arguments towards b , then a is labelled UNDEC.

In the following, we give a characterisation of the weighted semantics through the notion of labelling of $WAF_{\mathbb{S}}$. We use the term “ w -defends” to generally refer to any of the conditions of Definition 8 (\mathbb{D}_1 , \mathbb{D}_2 and \mathbb{D}_3). The w -conflict-free labelling coincides with the classical conflict-free labelling: since attacks are not allowed within a conflict-free set of arguments, weights are not relevant.

Definition 11 (w -conflict-free labelling). *Let L be a labelling of a $WAF_{\mathbb{S}}$ $F = \langle \mathcal{A}, \mathcal{R}, W, \mathbb{S} \rangle$ and $a \in \mathcal{A}$. L is a conflict-free labelling for F if and only if:*

- $L(a) = IN \implies a^-|_{IN} = \emptyset$, and
- $L(a) = OUT \implies a^-|_{IN} \neq \emptyset$

Concerning w -admissible labellings, for an argument a to be OUT there must exist at least an attack coming from an IN argument, so we require $W(a^-|_{IN}, a) <_{\mathbb{S}} \top$ (where \top means that there is no attack between two arguments).

Definition 12 (w -admissible labelling). *Let L be a labelling of a $WAF_{\mathbb{S}}$ $F = \langle \mathcal{A}, \mathcal{R}, W, \mathbb{S} \rangle$ and $a \in \mathcal{A}$. L is a w -admissible labelling for F if and only if:*

- $L(a) = IN \implies a^- = a^-|_{OUT} \wedge \mathcal{A}|_{IN}$ w -defends a , and
- $L(a) = OUT \implies W(a^-|_{IN}, a) <_{\mathbb{S}} \top$

The definition of the w -complete labelling derives from the w -admissible one.

Definition 13 (w -complete labelling). *Let L be a labelling of a $WAF_{\mathbb{S}}$ $F = \langle \mathcal{A}, \mathcal{R}, W, \mathbb{S} \rangle$ and $a \in \mathcal{A}$. L is a w -complete labelling for F if and only if:*

- $L(a) = IN \iff a^- = a^-|_{OUT} \wedge \mathcal{A}|_{IN}$ w -defends a , and
- $L(a) = OUT \iff W(a^-|_{IN}, a) <_{\mathbb{S}} \top$

A stable semantics partitions the arguments in two disjoint sets: one contains the arguments that are either not attacked or defended by other acceptable arguments, while the other contains the rest of the arguments (i.e., those that are attacked and not defended).

Definition 14 (w -stable labelling). *Let L be a labelling of a $WAF_{\mathbb{S}}$ $F = \langle \mathcal{A}, \mathcal{R}, W, \mathbb{S} \rangle$. L is a w -stable labelling for F if and only if*

- L is a w -complete labelling, and
- $\mathcal{A}|_{UNDEC} = \emptyset$

We next present the w -preferred and w -grounded labellings for $WAF_{\mathbb{S}}$.

Definition 15 (w -preferred labelling). *Let L be a labelling of a $WAF_{\mathbb{S}}$ $F = \langle \mathcal{A}, \mathcal{R}, W, \mathbb{S} \rangle$. L is a w -preferred labelling for F if and only if*

- L is a w -admissible labelling, and
- $\mathcal{A}|_{IN}$ is maximal among all the w -admissible labellings

Definition 16 (w -grounded labelling). *Let L be a labelling of a $WAF_{\mathbb{S}}$ $F = \langle \mathcal{A}, \mathcal{R}, W, \mathbb{S} \rangle$ and $a \in \mathcal{A}$. L is a w -grounded labelling for F if and only if:*

- $L(a) = IN \iff$ for all w -complete labellings L' , $L'(a) = IN$, and
- $L(a) = OUT \iff W(a^-|_{IN}, a) <_{\mathbb{S}} \top$

Table 1 summarises the conditions specified in Definitions from 11 to 16 for obtaining weighted labellings corresponding to the Dung semantics.

Theorem 1. *A labelling L of a $WAF_{\mathbb{S}}$ $F = \langle \mathcal{A}, \mathcal{R}, W, \mathbb{S} \rangle$ is a w -admissible (respectively w -complete, w -stable, w -preferred, w -grounded) labelling if and only if $\mathcal{A}|_{IN}$ is a w -admissible (respectively w -complete, w -stable, w -preferred, w -grounded) extension of F .*

We sketch the proof regarding w -admissible semantics (the others are obtained through similar reasoning). If L is w -admissible, then the OUT arguments attacking $\mathcal{A}|_{IN}$ are also defeated by $\mathcal{A}|_{IN}$. Thus, $\mathcal{A}|_{IN}$ is w -defended from the attacks coming from $\mathcal{A} \setminus \mathcal{A}|_{IN}$ and so it is a w -admissible extension. In the other case, when $\mathcal{A}|_{IN}$ is a w -admissible extension, we know that it w -defends itself from the attacks coming from OUT arguments in $\mathcal{A} \setminus \mathcal{A}|_{IN}$ and thus L is a w -admissible labelling.

Weighted semantics are a generalization of the classical case, and all the labellings for $WAF_{\mathbb{S}}$ corresponds to the respective classical semantics when the framework is instantiated with a boolean semiring. When the $WAF_{\mathbb{S}}$ is instantiated with a boolean semiring, all the attacks from an argument to another are associated with the value *false* and also $W(a^-|_{IN}, a)$ always corresponds to *false*.

Theorem 2. *The labelling of a $WAF_{\mathbb{S}}$ instantiated with a boolean semiring corresponds to the classical labelling.*

	conditions on IN arguments	conditions on OUT arguments	other conditions
w -cf	$L(a) = IN \implies a^- _{IN} = \emptyset$	$L(a) = OUT \implies a^- _{IN} \neq \emptyset$	
w -adm	$L(a) = IN \implies a^- = a^- _{OUT} \wedge \mathcal{A} _{IN}$ w -defends a	$L(a) = OUT \iff W(a^- _{IN}, a) <_{\mathbb{S}} \top$	
w -com	$L(a) = IN \iff a^- = a^- _{OUT} \wedge \mathcal{A} _{IN}$ w -defends a	$L(a) = OUT \iff W(a^- _{IN}, a) <_{\mathbb{S}} \top$	
w -stb	$L(a) = IN \iff a^- = a^- _{OUT} \wedge \forall b \in a^- . W(b^- _{IN}, b) \leq_{\mathbb{S}} W(b, b^+ _{IN})$	$L(a) = OUT \iff W(a^- _{IN}, a) <_{\mathbb{S}} \top$	$\mathcal{A} _{UNDEC} = \emptyset$
w -pre	$L(a) = IN \implies a^- = a^- _{OUT} \wedge \forall b \in a^- . W(b^- _{IN}, b) \leq_{\mathbb{S}} W(b, b^+ _{IN})$	$L(a) = OUT \implies W(a^- _{IN}, a) <_{\mathbb{S}} \top$	$\mathcal{A} _{IN}$ max w -adm
w -gde	$L(a) = IN \iff \forall L' \text{ } w\text{-com}, L'(a) = IN$	$L(a) = OUT \iff W(a^- _{IN}, a) <_{\mathbb{S}} \top$	

Table 1: Summarisation of the introduced labellings for $WAF_{\mathbb{S}}$.

It follows that if L is a w -admissible (respectively w -complete, w -stable, w -preferred, w -grounded) labelling of a $WAF_{\mathbb{S}}$ F , then L is an admissible (respectively complete, stable, preferred, grounded) labelling of F .

Implementation

To complete our study and facilitate the use of weighted labelling semantics for argumentation-based application, we provide a tool able to represent $WAF_{\mathbb{S}}$ and visualize the computed labellings for various semantics. For this purpose, we extend ConArg¹ (Bistarelli and Santini 2011), a suite of tools for argumentation, with a series of functionalities for handling weighted argumentation problems. At the time of writing, the tool allows to use of the defences \mathbb{D}_1 and \mathbb{D}_3 . The web interface, shown in Figure 2, is implemented in JavaScript and relies on a server-side solver written in C. In the following, we describe an example of use of the tool for weighted argumentation.

First of all, we use panel 4 of Figure 2 to select a semiring: this determines both the representation of the AF and the kind of solution provided by the solver. If “weighted” is chosen, it is possible to specify a $WAF_{\mathbb{S}}$ by either using the input area (panel 5) or directly clicking on the canvas to draw arguments and attacks. The next step is to select the semantics (panel 1) for which we want obtain a labelling. Since we selected the weighted semiring, we will obtain a weighted labelling. The solver computes the sets of IN arguments, that are then displayed in panel 6. The labellings are directly visible on the $WAF_{\mathbb{S}}$ through the usual colour scheme: IN arguments are green, OUT red and UNDEC yellow. In case the solver returns more than one solution for the selected semantics (as happens in Figure 2), we can choose which labelling to visualise by using panel 3.

Related Work

The problem of extending classical AFs with values expressing the strength of arguments and attacks is widely studied,

¹ConArg website: <http://dmi.unipg.it/conarg>.

and many different approaches have been presented in the literature. In (Amgoud and Cayrol 1998), the authors take into account preference orderings for comparing arguments, while in (Bench-Capon 2003) the success of an attack conducted by an argument toward another one depends on an ordering among the “values” promoted by each argument. A study on bipolar WAFs is conducted in (Pazienza, Ferilli, and Esposito 2017), where the authors present an extension for weighted frameworks that takes into account attack and support relations. Another formalism based on a notion of strength is given in (Baroni et al. 2015), where arguments are evaluated through a score system. The main difference with our work lies in the fact that we take into account the basic definition of WAFs (Dunne et al. 2011), without further refinements on the framework level. Moreover, our study is focused on the interpretation of the labelling in the weighted case.

Conclusion and Future Work

We introduce a labelling for WAFs, together with a set of labelling conditions corresponding to extensions for some semantics. We also show that our labelling function generalises the classical approach for the non-weighted case. Finally, we present an online implementation of the labelling for WAFs. We have considered three different definitions of collective defence provided in the literature.

As future work, we plan to extend this work in different directions. The definitions of the labelling-based semantics for WAFs 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 (Modgil and Caminada 2009) for classical AFs. An interesting study could then be carried out using a four-states labelling semantics (Baroni, Giacomini, and Liao 2015; Arieli 2016) which further differentiate UNDEC arguments. In our context, the difference between labels could be made more continuous by considering the weight on the attack relations. continuous by considering the weight on the attack relations. Accordingly,

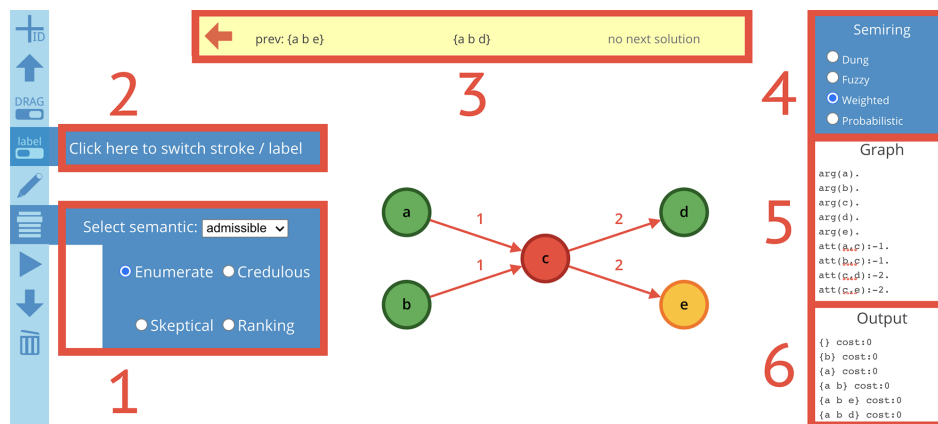


Figure 2: ConArg web interface displaying a weighted labelling for a $WAF_{\mathbb{S}}$. The highlighted areas corresponds to: 1) semantics selection, 2) representation of weights by stroke/label, 3) solution selection, 4) semiring selection, 5) input area, 6) output area.

we want to provide real world examples where the use of three (or four) labels is more convenient than the classical accepted/rejected partitioning. We also plan to give a definition of w -strongly admissible extension (generalising the one provided in (Baroni and Giacomin 2007) for the crisp case) and introduce the respective labelling. Finally, given the close correspondence between argumentation semantics and the stable model semantics for logic programming we would like to investigate possible connections between weighted argumentation and multi-valued logic programming (Mobasher, Pigozzi, and Slutzki 1997).

References

- Amgoud, L., and Cayrol, C. 1998. On the acceptability of arguments in preference-based argumentation. In *UAI*, 1–7. Morgan Kaufmann.
- Arieli, O. 2016. On the acceptance of loops in argumentation frameworks. *J. Log. Comput.* 26(4):1203–1234.
- Baroni, P., and Giacomin, M. 2007. On principle-based evaluation of extension-based argumentation semantics. *Artif. Intell.* 171(10-15):675–700.
- Baroni, P.; Romano, M.; Toni, F.; Aurisicchio, M.; and Bertanza, G. 2015. Automatic evaluation of design alternatives with quantitative argumentation. *Argument Comput.* 6(1):24–49.
- Baroni, P.; Caminada, M.; and Giacomin, M. 2011. An introduction to argumentation semantics. *Knowl. Eng. Rev.* 26(4):365–410.
- Baroni, P.; Caminada, M.; and Giacomin, M. 2018. Abstract argumentation frameworks and their semantics. *Handbook of formal argumentation* 1:157–234.
- Baroni, P.; Giacomin, M.; and Liao, B. S. 2015. 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*, volume 9060 of *Lecture Notes in Computer Science*, 265–280. Springer.
- Bench-Capon, T. J. M. 2003. Persuasion in practical argument using value-based argumentation frameworks. *J. Log. Comput.* 13(3):429–448.
- Bistarelli, S., and Gadducci, F. 2006. Enhancing constraints manipulation in semiring-based formalisms. In *ECAI*, volume 141 of *Frontiers in Artificial Intelligence and Applications*, 63–67. IOS Press.
- Bistarelli, S., and Santini, F. 2011. Conarg: A constraint-based computational framework for argumentation systems. In *ICTAI*, 605–612. IEEE Computer Society.
- Bistarelli, S., and Santini, F. 2017. A hasse diagram for weighted sceptical semantics with a unique-status grounded semantics. In *LPNMR*, volume 10377 of *Lecture Notes in Computer Science*, 49–56. Springer.
- Bistarelli, S., and Taticchi, C. 2020. A labelling semantics for weighted argumentation frameworks. In *CILC*, volume 2710 of *CEUR Workshop Proceedings*, 263–277. CEUR-WS.org.
- Bistarelli, S.; Montanari, U.; and Rossi, F. 1997. Semiring-based constraint satisfaction and optimization. *J. ACM* 44(2):201–236.
- Bistarelli, S.; Rossi, F.; and Santini, F. 2016. A collective defence against grouped attacks for weighted abstract argumentation frameworks. In *FLAIRS Conference*, 638–643. AAAI Press.
- Bistarelli, S.; Rossi, F.; and Santini, F. 2018. A novel weighted defence and its relaxation in abstract argumentation. *Int. J. Approx. Reason.* 92:66–86.
- Caminada, M. 2006. On the issue of reinstatement in argumentation. In *JELIA*, volume 4160 of *Lecture Notes in Computer Science*, 111–123. Springer.
- Caminada, M. 2014. Strong admissibility revisited. In *COMMA*, volume 266 of *Frontiers in Artificial Intelligence and Applications*, 197–208. IOS Press.
- Cocarascu, O., and Toni, F. 2016. Argumentation for machine learning: A survey. In *COMMA*, volume 287 of *Frontiers in Artificial Intelligence and Applications*, 219–230. IOS Press.
- Coste-Marquis, S.; Konieczny, S.; Marquis, P.; and Ouali, M. A. 2012. Weighted attacks in argumentation frameworks. In *KR*. AAAI Press.
- Dung, P. M. 1995. On the acceptability of arguments and its fundamental role in nonmonotonic reasoning, logic programming and n-person games. *Artif. Intell.* 77(2):321–358.
- Dunne, P. E.; Hunter, A.; McBurney, P.; Parsons, S.; and Wooldridge, M. J. 2011. Weighted argument systems: Basic definitions, algorithms, and complexity results. *Artif. Intell.* 175(2):457–486.
- Jakobovits, H., and Vermeir, D. 1999. Robust semantics for argumentation frameworks. *J. Log. Comput.* 9(2):215–261.
- Lawrence, J.; Park, J.; Budzynska, K.; Cardie, C.; Konat, B.; and Reed, C. 2017. Using argumentative structure to interpret debates in online deliberative democracy and erulemaking. *ACM Trans. Internet Techn.* 17(3):25:1–25:22.
- Martínez, D. C.; García, A. J.; and Simari, G. R. 2008. An abstract argumentation framework with varied-strength attacks. In *KR*, 135–144. AAAI Press.
- Mobasher, B.; Pigozzi, D.; and Slutzki, G. 1997. Multi-valued logic programming semantics: An algebraic approach. *Theor. Comput. Sci.* 171(1-2):77–109.
- Modgil, S., and Caminada, M. 2009. Proof theories and algorithms for abstract argumentation frameworks. In *Argumentation in Artificial Intelligence*. Springer. 105–129.
- Pazienza, A.; Ferilli, S.; and Esposito, F. 2017. Constructing and evaluating bipolar weighted argumentation frameworks for online debating systems. In *AI³@AI*IA*, volume 2012 of *CEUR Workshop Proceedings*, 111–125. CEUR-WS.org.
- Walton, D., and Koszowy, M. 2017. Arguments from authority and expert opinion in computational argumentation systems. *AI Soc.* 32(4):483–496.