Proceedings of the Twenty-Ninth International
Florida Artificial Intelligence Research Society Conference

A Collective Defence Against Grouped Attacks for Weighted Abstract Argumentation Frameworks*

Stefano Bistarelli and **Fabio Rossi** and **Francesco Santini**
Dipartimento di Matematica e Informatica, Università di Perugia
Via Vanvitelli 1, 06123, Perugia, Italy
[bista,rossi,francesco.santini]@dmi.unipg.it

Abstract

Adding weights or preferences to Abstract Argumentation Frameworks can help disentangle semantics from otherwise all-equivalent attacks. Having such information makes possible to distil the set of found extensions by reducing their number. In this work we provide a new definition of weighted defence: according to it, all the attacks from an argument to a set of arguments are considered with a single global weight, i.e., attacks are grouped together. This provides a coherent view w.r.t. defence, which is usually "collective" in the literature. Moreover, we model weighted defences from related works in the same algebraic framework: this allows us to compare all the different proposals together.

1 Introduction

An *Abstract Argumentation Framework* (*AAF*) (Dung 1995) is essentially a pair $\langle \mathcal{A}_{rgs}, R\rangle$ consisting of a set of arguments and a binary relationship of attack defined among them. Given a framework, it is possible to examine the question on which set(s) of arguments can be accepted, hence surviving the conflict defined by R. Answering this question corresponds to defining an argumentation semantics. The key idea behind *extension-based* semantics is to identify some sets of arguments (called *extensions*) that survive the conflict "together". A very simple example of AAF is $\langle\{a, b\}, \{R(a, b), R(b, a)\}\rangle$, where two arguments a and b attack each other. In this case, each of the two positions represented by either $\{a\}$ or $\{b\}$ can be intuitively valid, since no additional information is provided on which of the two attacks is stronger. For instance, in case the attack $R(a, b)$ is stronger than $R(b, a)$, taking the position defined by a may result in a better choice (*e.g.*, , for an intelligent agent).

Several notions of weighted defence have been defined in the literature. Attacks are associated with a weight indicating a "strength" value of an attack, thus they represent an additional quantitative-information. Examples are Preference-based AAFs (*PAFs*) (Amgoud and Cayrol 1998), Value-based AAFs (*VAFs*) (Bench-Capon 2003), frameworks considering a probability or uncertainty score of attacks (Li, Oren, and Norman 2011), a fuzzy measure of their strength (Janssen, Cock, and Vermeir 2008). Preferences can also be given qualitatively (Martínez, García, and Simari 2008; Modgil 2009).

Having such weights naturally brings to generalise the notion of AAF into *Weighted AAF* (*WAAF*) (Dunne et al. 2009; Bistarelli and Santini 2010; Dunne et al. 2011). The aim of this paper is to provide a new definition of defence for WAAFs, here called w-defence, which encompasses also weights in the style of similar works, as (Martínez, García, and Simari 2008) and (Coste-Marquis et al. 2012). In our proposal, an extension $\mathcal{B} \subseteq \mathcal{A}_{rgs}$ defends an argument $b \in \mathcal{A}_{rgs}$ from $a \in \mathcal{A}_{rgs}$, if the "sum" (a parametric $\times$ operation from a *c-semiring* structure (Bistarelli, Montanari, and Rossi 1997)) of all the attack weights from $\mathcal{B}$ to a is stronger than the "sum" of *all* the attacks from a to $\mathcal{B} \cup \{b\}$. Differently from (Coste-Marquis et al. 2012), where the arithmetic sum of all attack weights from $\mathcal{B}$ to a needs to be only stronger than the attack from a to b, we also consider the set of attacks from a to the indented defender $\mathcal{B}$. Therefore, both our proposal and the one given by Coste-Marquis et al. suggest a collective defence from $\mathcal{B}$ to a, but, differently, in this paper we consider the group of attacks from a to b and $\mathcal{B}$ as a single entity, i.e., with a single global weight.

We believe our choice provides a more coherent view: in the literature defence is usually checked by considering all the counter-attacks from a set $\mathcal{B}$ to a (*e.g.*, in order to satisfy the admissible semantics), but each attack from a to $\mathcal{B}$ is treated separately. Our intent is to normalise such dishomogeneity.

As a second result, w-defence and the defences defined by Coste-Marquis et al. and Martínez, García, and Simari are framed into the same parametric algebraic-framework that exploits semirings, with the possibility to consider different instantiations of aggregation operators. As a consequence, we can compare all such proposals together, *e.g.*, w-defence implies the defence in (Coste-Marquis et al. 2012).

The paper is structured as follows: in Sec. 2.1 we present c-semirings, while Sec. 2.2 recollects the basic definitions of (unweighted) AAF given by Dung. Section 3 presents WAAFs and w-defence. Afterwards, Sec. 4 exploits w-defence to redefine the classical semantics into new w-semantics. In Sec. 5 we describe in detail how this proposal is placed with respect to similar works in the literature. Fi-

*This work is supported by GNCS-INDAM "Efficient Methods for Argumentation-based Decision Support".

nally, Sec. 6 concludes the work by summarising the main ideas and introducing future work.

2 Background

In the following of this section we first introduce c-semirings (Sec. 2.1), and then (Sec. 2.2) we recollect the main definitions at the basis of AAFs (Dung 1995). C-semirings represent a parametric framework where to measure and compose attack-weights. By changing the underlying c-semiring instantiation, it is possible to capture different metrics (*e.g.*, fuzzy or probabilistic ones).

2.1 Semirings

Definition 1 (Semirings)**.** *A commutative semiring is a five-tuple $\mathbb{S} = \langle S, +, \times, \bot, \top \rangle$ such that S is a set, $\top, \bot \in S$, and $+, \times : S \times S \to S$ are binary operators making the triples $\langle S, +, \bot \rangle$ and $\langle S, \times, \top \rangle$ commutative monoids (semigroups with identity), satisfying*

- *(distributivity)* $\forall a, b, c \in S.a \times (b+c) = (a \times b) + (a \times c)$.
- *(annihilator)* $\forall a \in A.a \times \bot = \bot$.

Definition 2 (Absorptive semirings)**.** *Let $\mathbb{S}$ be a commutative semiring. An absorptive semiring verifies the absorptiveness property: $\forall a, b \in S.a + (a \times b) = a$, which is equivalent to $\forall a \in S.a + \top = \top$.*

Absorptive semirings are referred also as *simple*, and their $+$ operator is necessarily idempotent. Semirings where $+$ is idempotent are defined as *tropical* semirings, or *diods*.

Definition 3 (C-semirings (Bistarelli, Montanari, and Rossi 1997))**.** *C-semirings are commutative and absorptive semirings. Therefore, c-semirings are tropical semirings where $\top$ is an absorbing element for $+$.*

The idempotency of $+$ leads to the definition of a partial ordering $\leq_{\mathbb{S}}$ over the set S (S is a poset). Such partial order is defined as $a \leq_{\mathbb{S}} b$ if and only if $a + b = b$, and $+$ becomes the *least upper bound* of the lattice $\langle S, \leq_{\mathbb{S}} \rangle$. This intuitively means that b is "better" than a. As a consequence, we can use $+$ as an optimisation operator and always choose the best available solution. Some more properties can be derived on c-semirings (Bistarelli, Montanari, and Rossi 1997): *i)* both $+$ and $\times$ are monotone over $\leq_{\mathbb{S}}$, *ii)* $\times$ is intensive (i.e., $a \times b \leq_{\mathbb{S}} a$), and *iii)* $\langle S, \leq_{\mathbb{S}} \rangle$ is a complete lattice. $\bot$ and $\top$ are respectively the bottom and top elements of such lattice. When also $\times$ is idempotent, *i)* $+$ distributes over $\times$, *ii)* $\times$ is the *greater lower bound* of the lattice, and *iii)* $\langle S, \leq_{\mathbb{S}} \rangle$ is a distributive lattice.

Some c-semiring instances are: *boolean* $\langle \{F, T\}, \vee, \wedge, F, T \rangle$[1], *fuzzy* $\langle [0,1], \max, \min, 0, 1 \rangle$, *bottleneck* $\langle \mathbb{R}^+ \cup \{+\infty\}, \max, \min, 0, \infty \rangle$, *probabilistic* $\langle [0,1], \max, \hat{\times}, 0, 1 \rangle$ (known as the Viterbi semiring), *weighted* $\langle \mathbb{R}^+ \cup \{+\infty\}, min, \hat{+}, +\infty, 0 \rangle$. Capped operators stand for their arithmetic equivalent to distinguish them from $+$ and $\times$.

Furthermore, it is also possible to consider several optimisation criteria at the same time: the cartesian product of semirings is still a semiring. Clearly, in this case the ordering induced by $+$ is partial, *e.g.*, when we have $\langle k_1, k_2 \rangle$ and $\langle k_3, k_4 \rangle$, and $k_1 \leq k_3$ while $k_2 \geq k_4$.

[1] Boolean c-semirings can be used to model crisp problems.

2.2 Argument Systems

In his pioneering work (Dung 1995), Dung proposed *Abstract Frameworks* for Argumentation, where an argument is an abstract entity whose role is solely determined by its relations to other arguments:

Definition 4. *An Abstract Argumentation Framework (AAF) is a pair $\langle \mathcal{A}_{rgs}, R \rangle$ of a set $\mathcal{A}_{rgs}$ of arguments and a binary relation R on $\mathcal{A}_{rgs}$ called the (asymmetric) attack relation. $\forall a_i, a_j \in \mathcal{A}_{rgs}$, $a_i R\, a_j$ (or $R(a_i, a_j)$) means that a_i attacks a_j (R is asymmetric).*

An example is given in Fig. 1. An *argumentation semantics* is the formal definition of a method (either declarative or procedural) ruling the argument evaluation process. In the *extension*-based approach, a semantics definition specifies how to derive from an AAF a set of extensions, where an extension $\mathcal{B}$ of an AAF $\langle \mathcal{A}_{rgs}, R \rangle$ is simply a subset of $\mathcal{A}_{rgs}$. In Def. 5 we define the first semantics, which is at hearth of all the other ones:

Definition 5 (Conflict-free)**.** *A set $\mathcal{B} \subseteq \mathcal{A}_{rgs}$ is conflict-free iff no two arguments a and b in $\mathcal{B}$ exist such that a attacks b.*

All the other semantics presented in this section rely (explicitly or implicitly) upon the concept of defence:

Definition 6 (defence ($\mathbb{D}_0$))**.** *An argument b is defended by a set $\mathcal{B} \subseteq \mathcal{A}_{rgs}$ (or $\mathcal{B}$ defends b) iff for any argument $a \in \mathcal{A}_{rgs}$, if a attacks b then $\mathcal{B}$ attacks a.*

An admissible set of arguments according to Dung must be a conflict-free set which defends all its elements. Formally:

Definition 7 (Admissible)**.** *A conflict-free set $\mathcal{B} \subseteq \mathcal{A}_{rgs}$ is admissible iff each argument in $\mathcal{B}$ is defended by $\mathcal{B}$.*

The four classical semantics refining admissibility are defined in the following three definitions:

Definition 8 (Complete)**.** *An admissible extension $\mathcal{B} \subseteq \mathcal{A}_{rgs}$ is a complete extension iff each argument which is defended by $\mathcal{B}$ is in $\mathcal{B}$.*

Definition 9 (Preferred and Grounded)**.** *A preferred extension is a maximal (w.r.t. set inclusion) admissible subset of $\mathcal{A}_{rgs}$. The least (w.r.t. set inclusion) complete extension is the grounded extension.*

Finally, the stable semantics corresponds to the most stringent among all:

Definition 10 (Stable)**.** *A conflict-free set $\mathcal{B} \subseteq \mathcal{A}_{rgs}$ is a stable extension iff for each argument which is not in $\mathcal{B}$, there exists an argument in $\mathcal{B}$ that attacks it.*

If $\sigma = \{adm, com, stb, prf, gde\}$ respectively stand for admissible, complete, stable, preferred, and grounded semantics, we recall that given any framework F, $stb(F) \subseteq$

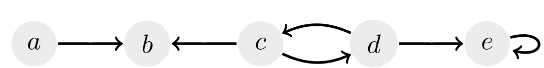

Figure 1: An example of AAF.

$prf(F) \subseteq com(F) \subseteq adm(F)$ always holds. Moreover, for each σ except stb we have $\sigma(F) \neq \emptyset$ holds.

Consider the AAF $F = \langle A, R\rangle$ in Fig. 1, with $A = \{a, b, c, d, e\}$ and $R = \{(a, b), (c, b), (c, d), (d, c), (d, e), (e, e)\}$. We have that $adm(F)$ corresponds to $\{\emptyset, \{a\}, \{c\}, \{d\}, \{a, c\}, \{a, d\}\}$, $com(F) = \{\{a\}, \{a, c\}, \{a, d\}\}$, $prf(F) = \{\{a, c\}, \{a, d\}\}$, $stb(F) = \{\{a, d\}\}$, and $gde(F) = \{a\}$.

3 Weighted Abstract AFs

In the following of this section we rephrase some of the classical definitions given in Sec. 2.2, with the purpose to parametrise them with the notion of weighted attack and c-semiring. Such notions, e.g., the one of w-defence, are the premise behind the new semantics we then propose in Sec. 4. The following definition presents *semiring-based WAAF* (Bistarelli and Santini 2010), called $WAAF_{\mathbb{S}}$:

Definition 11 (c-semiring-based WAAF (Bistarelli and Santini 2010))**.** *A semiring-based Argumentation Framework ($WAAF_{\mathbb{S}}$) is a quadruple $\langle \mathcal{A}_{rgs}, R, W, \mathbb{S}\rangle$, where $\mathbb{S}$ is a semiring $\langle S, +, \times, \bot, \top\rangle$, $\mathcal{A}_{rgs}$ is a set of arguments, R the attack binary-relation on $\mathcal{A}_{rgs}$, and $W : \mathcal{A}_{rgs} \times \mathcal{A}_{rgs} \longrightarrow S$ is a binary function. Given $a, b \in \mathcal{A}_{rgs}$, $\forall (a, b) \in R$, $W(a, b) = s$ means that a attacks b with a weight $s \in S$. Moreover, we require that $R(a, b)$ iff $W(a, b) <_{\mathbb{S}} \top$.*

In Fig. 2 we provide an example of a weighted interaction graph describing the $WAAF_{\mathbb{S}}$ defined by $\mathcal{A}_{rgs} = \{a, b, c, d, e\}$, $R = \{(a, b), (c, b), (c, d), (d, c), (d, e), (e, e)\}$, with $W(a, b) = 7, W(c, b) = 8, W(c, d) = 9, W(d, c) = 8, W(d, e) = 5, W(e, e) = 6$, and $\mathbb{S} = \langle \mathbb{R}^+ \cup \{\infty\}, \min, \hat{+}, \infty, 0\rangle$ (i.e., a weighted semiring).

Therefore, each attack is associated with a semiring value that represents the "strength" of an attack between two arguments. We can consider the weights in Fig. 2 as supports to the associated attack, as similarly suggested in (Dunne et al. 2009) and (Dunne et al. 2011). A semiring value equal to the top element of the c-semiring $\top$ (e.g., 0 for the weighted semiring) represents a no-attack relation between two arguments: for instance, $(a, c) \notin R$ in Fig. 2 corresponds to $W(a, c) = 0$. Note that, when $R(a, b)$, we always have $W(a, b) \neq \top$ (in Fig. 2, *e.g.*, $W(a, b) = 7$). On the other side, the bottom element, i.e., $\bot$ (e.g., ∞ for the weighted semiring), represents the strongest attack possible (e.g., $\forall s \in S$, $s \geq_{\mathbb{S}} \bot$ and $\top \geq_{\mathbb{S}} s$).

In Def. 12 we define the attack strength for a set of arguments that attacks an argument, a different set of arguments, or an argument that attacks a set of arguments; the former and the latter are what we need to define w-defence. In the following, we will use $\prod$ to indicate the $\times$ operator of the c-semiring $\mathbb{S}$ on a set of values:

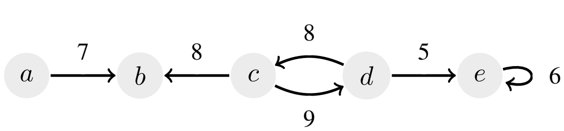

Figure 2: An example of WAAF.

Definition 12 (Attacks to/from sets of arguments)**.** *Given a $WAAF_{\mathbb{S}}$, $WF = \langle \mathcal{A}_{rgs}, R, W, \mathbb{S}\rangle$,*

- *a set of arguments $\mathcal{B}$ attacks an argument a with a weight of $k \in S$ if*
$$W(\mathcal{B}, a) = \prod_{b \in \mathcal{B}} W(b, a) = k$$
- *an argument a attacks a set of arguments $\mathcal{B}$ with a weight of $k \in S$ if*
$$W(a, \mathcal{B}) = \prod_{b \in \mathcal{B}} W(a, b) = k$$
- *a set of arguments $\mathcal{B}$ attacks a set of arguments $\mathcal{D}$ with a weight of $k \in S$ if*
$$W(\mathcal{B}, \mathcal{D}) = \prod_{b \in \mathcal{B}, d \in \mathcal{D}} W(b, d) = k$$

For example, looking at Fig. 2 we have that $W(\{a, c\}, b) = 15$, $W(c, \{b, d\}) = 17$, and $W(\{a, c\}, \{b, d\}) = 24$.

We now ready to define our version of weighted defence, i.e., w-defence:

Definition 13 (w-defence ($\mathbb{D}_w$))**.** *Given a $WAAF_{\mathbb{S}}$, $WF = \langle \mathcal{A}_{rgs}, R, W, S\rangle$, $\mathcal{B} \subseteq \mathcal{A}_{rgs}$ w-defends $b \in \mathcal{A}_{rgs}$ iff, given $a \in \mathcal{A}_{rgs}$ s.t. $R(a, b)$, then $W(a, \mathcal{B} \cup \{b\}) \geq_{\mathbb{S}} W(\mathcal{B}, a)$; $\mathcal{B}$ w-defends b iff it defends b from any a s.t. $R(a, b)$.*

As previously advanced, a set $\mathcal{B} \subseteq \mathcal{A}_{rgs}$ defends an argument b, if the $\times$ of all the attack weights from $\mathcal{B}$ to a (for any a s.t. $R(a, b)$) is worse-equal (w.r.t. $\leq_{\mathbb{S}}$) than the $\times$ of the attacks from a to $\mathcal{B} \cup \{b\}$. For example, the set $\{c\}$ in Fig. 2 defends c from d because $W(d, \{c\}) \geq_{\mathbb{S}} W(\{c\}, d)$, i.e., $(8 \leq 9)$. On the other hand, $\{d\}$ in Fig. 2 does not defend d because $W(c, \{d\}) \not\geq_{\mathbb{S}} W(\{d\}, c)$.

As defined, w-defence implies the classical Dung's defence in Def. 6:

Proposition 1 ($\mathbb{D}_w \Rightarrow \mathbb{D}_0$)**.** *Given a $WAAF_{\mathbb{S}}$, $WF = \langle \mathcal{A}_{rgs}, R, W, S\rangle$, a subset of arguments $\mathcal{B}$, and $b \in \mathcal{A}_{rgs}$, "$\mathcal{B}$ w-defends b" $\Rightarrow$ "$\mathcal{B}$ defends b (Dung 1995)".*

Proof. As hypothesis we have $R(a, b)$ (from Def. 13), then $W(a, \mathcal{B} \cup \{b\}) \neq \top$. Therefore, if $W(a, \mathcal{B} \cup \{b\}) \geq_{\mathbb{S}} W(\mathcal{B}, a)$ is true (i.e., $\mathcal{B}$ w-defends b from a), this implies that $W(\mathcal{B}, a) \neq \top$. This can be also read as "$\mathcal{B}$ attacks a", which exactly corresponds to the original definition of defence (see Def. 6). □

Moreover, the following proposition equates defence and w-defence in case we adopt the boolean c-semiring (see Sec. 2.1):

Proposition 2. *Given a $WAAF_{\mathfrak{B}}$, $WF = \langle \mathcal{A}_{rgs}, R, W, \mathbb{S}\rangle$, where $\mathbb{S} = \langle \{true, false\}, \vee, \wedge, false, true\rangle$ (i.e., the boolean semiring), "$\mathcal{B}$ w-defends a" $\iff$ "$\mathcal{B}$ defends a".*

Proof. This holds because, $\mathcal{B}$ defends b corresponds to, "if $W(a, b) \neq \top$ then $W(\mathcal{B}, a) \neq_{\mathbb{S}} \top$". But, since we are using the boolean semiring, this statement can only correspond to, "if $W(a, b) = false$ then $W(\mathcal{B}, a) = false$", since the set of preferences only contains $\top$ (*true*) and $\bot$ (*false*). Therefore, $W(a, b) \geq_{\mathbb{S}} W(\mathcal{B}, a)$ is always true (in this case, $false \geq_B false$), and $\mathcal{B}$ w-defends b from a. □

4 Dung's Semantics Revisited

Given all the definitions in Sec. 3, it is possible to redefine the extension-based semantics summarised in Sec. 2.2. Since w-defence does not interfere with classical conflict-free semantics (see Def. 6), then, w-defence does not affect w-conflict-free semantics:

Definition 14 (w-conflict-free)**.** *Given a $WAAF_{\mathbb{S}}$, $WF = \langle \mathcal{A}_{rgs}, R, W, S\rangle$, a subset of arguments $\mathcal{B} \subseteq \mathcal{A}_{rgs}$ is w-conflict-free iff $W(\mathcal{B}, \mathcal{B}) = \top$.*

Note that, by allowing $W(\mathcal{B}, \mathcal{B}) \neq \top$ it is possible to tolerate a certain level of conflict inside an extension satisfying such semantics; in this case, this definition reconnects with the work in (Bistarelli and Santini 2010).

The notion of w-defence brings instead to the definition of the w-admissible semantics:

Definition 15 (w-admissible)**.** *Given $WF = \langle \mathcal{A}_{rgs}, R, W, S\rangle$, a w-conflict-free extension $\mathcal{B} \subseteq \mathcal{A}_{rgs}$ is w-admissible iff all the arguments in $\mathcal{B}$ are w-defended by $\mathcal{B}$.*

Considering the framework in Fig. 2 as unweighted (i.e., as the one in Fig. 1), Dung's admissible sets are: $\{a\}, \{c\}, \{d\}, \{a, c\}, \{a, d\}$. w-admissible extensions are $\{a\}$, $\{c\}$, and $\{a, c\}$ instead: $\{a\}$ because is not attacked by any other argument, $\{c\}$ and $\{a, c\}$ because they both w-defends c from the attack performed by d, i.e., $W(d, \{c\}) \geq_{\mathbb{S}} W(\{c\}, d)$ and $W(d, \{a, c\}) \geq_{\mathbb{S}} W(\{a, c\}, d)$ (i.e., for both of them $8 \leq 9$). Therefore, by using w-defence we restrain the set of admissible extensions, eliminating $\{d\}$ and $\{a, d\}$. An alternative definition of the w-admissible semantics is proposed in Prop. 3.

Proposition 3 (Alternative definition)**.** *Given $WF = \langle \mathcal{A}_{rgs}, R, W, S\rangle$, a w-conflict-free extension $\mathcal{B} \subseteq \mathcal{A}_{rgs}$ is w-admissible iff $\forall a \in \mathcal{A}_{rgs} \setminus \mathcal{B}$, $W(a, \mathcal{B}) \geq_{\mathbb{S}} W(\mathcal{B}, a)$.*

Proof. Given $b \in \mathcal{B}$, w-defence, i.e., $W(a, \mathcal{B} \cup \{b\}) \geq_{\mathbb{S}} W(\mathcal{B}, a)$, directly reduces to $W(a, \mathcal{B}) \geq_{\mathbb{S}} W(\mathcal{B}, a)$. □

Four further semantics, which refine the w-admissible one, are introduced from Def. 16 to Def. 18:

Definition 16 (w-complete)**.** *A w-admissible extension $\mathcal{B} \subseteq \mathcal{A}_{rgs}$ is also a w-complete extension iff each argument $b \in \mathcal{A}_{rgs}$ s.t. $\mathcal{B} \cup \{b\}$ is w-admissible belongs to $\mathcal{B}$, i.e., $b \in \mathcal{B}$.*

Definition 17 (w-preferred and w-grounded)**.** *A w-preferred extension is a maximal (with respect to set inclusion) w-admissible subset of $\mathcal{A}_{rgs}$. The least (with respect to set inclusion) w-complete extension is the w-grounded extension.*

Definition 18 (w-stable)**.** *Given $WF = \langle \mathcal{A}_{rgs}, R, W, S\rangle$, a w-admissible set $\mathcal{B}$ is a w-stable extension iff $\forall a \notin \mathcal{B}, \exists b \in \mathcal{B}. W(b, a) \leq_{\mathbb{S}} \top$.*

Note that, w.r.t. Fig. 2, the set of w-complete extensions is $\{\{a\}, \{a, c\}\}$, the only w-preferred extension is $\{\{a, c\}\}$, but there is no w-stable extension, despite $\{\{a, d\}\}$ is a stable extension according to Dung.

Theorems 1 and 2 relate new w-extensions to their counterpart in the original proposal (Dung 1995).

Theorem 1. *Given $F = \langle \mathcal{A}_{rgs}, R\rangle$, and $WF = \langle \mathcal{A}_{rgs}, R, W, \mathbb{S}\rangle$, with $\mathbb{S}$ as desired, then*

$$\begin{tikzcd}
w\text{-stable} \arrow[dr, Rightarrow] & & & \\
& w\text{-preferred} \arrow[dr, Rightarrow] & & w\text{-grounded} \arrow[dl, Rightarrow] \\
& & w\text{-complete} \arrow[d, Rightarrow] & \\
& & w\text{-admissible} \arrow[d, Rightarrow] & \\
& & w\text{-conflict-free} &
\end{tikzcd}$$

Figure 3: Implications among w-semantics.

i) the set of w-conflict-free extensions in WF is equal to the set of conflict-free extensions in F.

ii) the set of w-admissible extensions in WF is a subset of the set of admissible extensions in F.

iii) the set of w-stable extensions in WF is a subset of the set of stable extensions in F.

Proof. Concerning *i*, $\top$ represents a no-attack relation, so $\top$-conflict-free extensions do not include any attack. *ii* and *iii* hold because the notion of w-defence implies the classical notion of defence, but not vice versa (Prop. 1). □

Theorem 2. *Given $F = \langle \mathcal{A}_{rgs}, R\rangle$, and $WF = \langle \mathcal{A}_{rgs}, R, W, \mathbb{S}\rangle$, with $\mathbb{S}$ as desired, then*

i) for each w-complete extension B_{WF} in WF, there exists a complete extension B_F in F, s.t., $B_{WF} \subseteq B_F$.

ii) if B_{WF} is the grounded extension of WF, and if $B_A \neq \emptyset$ is the grounded extension of F, then $B_{WF} \subseteq B_A$.

iii) for each w-preferred extension B_{WF} in WF, there exists a preferred extension B_F in F, s.t. $B_{WF} \subseteq B_F$.

Proof. *iii* follows from Def. 9, and Def. 17: the maximal w-admissible extensions w.r.t. set inclusion are computed over a subset of the admissible ones ($\mathbb{D}_w \Rightarrow \mathbb{D}_0$ in Prop. 1). Therefore, each of them is a subset of at least one preferred extension in the corresponding unweighted framework. Same considerations hold for *i*, since $\mathbb{D}_w \Rightarrow \mathbb{D}_0$ less arguments need to be taken in order to have a valid w-complete extension. *ii* follows from *i*: the least w-complete extension is computed over subsets of Dung's complete extensions, thus their intersection may result in a smaller set. However, if the grounded extension B_F is $\emptyset$, having smaller w-complete extensions may lead to $B_F \subseteq B\,WF$. □

Moreover, Prop. 4 restates the classical implication chain between semantics (Dung 1995), which still holds for w-semantics as well. This is visually represented also by Fig. 3.

Proposition 4. *The following implications hold between w-semantics: w-stable $\Rightarrow$ w-preferred $\Rightarrow$ w-complete $\Rightarrow$ w-admissible $\Rightarrow$ w-conflict-free, and w-grounded $\Rightarrow$ w-complete.*

Proof. All the implications are proved by definition, from Def. 14 to Def. 18. □

5 A Comparison with Related Work

Two of the most related definitions of weighted defence (i.e., Def. 13) are (Coste-Marquis et al. 2012) and (Martínez, García, and Simari 2008). In the following we condense their main features and we show how our approach differs.

In (Martínez, García, and Simari 2008) attacks are relatively ordered by their force, i.e., $R(a,b) \gg R(b,a)$ means that the former attack is stronger than the latter (vice-versa, a weaker attack). Equivalent and incomparable classes are considered as well, i.e., respectively $R(a,b) \approx R(b,a)$ and $R(a,b)?R(b,a)$. This is accordingly reflected by the defence definition, where considering $R(a,b)$ and $R(c,a)$ we can have that c is a *strong*, *weak*, *normal*, or *unqualified* defender of b. Therefore, an argument b is defended by $\mathcal{B}$ if, and only if, for any argument a such that $R(a,b)$, there is an argument $c \in \mathcal{B}$ such that $R(c,a)$, and according to the desired defence strength, $R(c,a) \gg R(a,b)$, $R(c,a) \ll R(a,b)$, $R(c,a) \approx R(a,b)$, and $R(c,a)?R(a,b)$. For instance, when requiring a level $[\gg, \approx]$, for each attacker a of b there must must be either a strong or a normal defender $c \in \mathcal{B}$. In Def. 19 we exactly rephrase such defence by modelling the total order defined by $[\gg, \approx]$ with a c-semiring $\mathbb{S}$:

Definition 19 ($\mathbb{D}_1$)**.** *Given* $WF = \langle \mathcal{A}_{rgs}, R, W, \mathbb{S}\rangle$, $a, b, c \in \mathcal{A}_{rgs}$, $\mathcal{B} \subseteq \mathcal{A}_{rgs}$, *then* b *is defended by* $\mathcal{B}$ *if for any* $R(a,b), \exists c \in \mathcal{B}$ *s.t.* $W(a,b) \geq_{\mathbb{S}} W(c,a)$.

In (Coste-Marquis et al. 2012) the authors define $\sigma^{\oplus}$-extensions, where σ is one of the given semantics (e.g., admissible), and $\oplus$ is an *aggregation function* ($\times$ in a c-semiring). $\oplus$ needs to satisfy non-decreasingness, minimality, and identity:[2] two examples are the arithmetic sum and $\max$. Even the notion of defence is refined: in Def. 20 we cast it in the same semiring-based framework.

Definition 20 ($\mathbb{D}_2$)**.** *Given* $WF = \langle \mathcal{A}_{rgs}, R, W, \mathbb{S}\rangle$, *an argument* b *is defended by a subset of arguments* $\mathcal{B}$ *if* $\forall a \in \mathcal{A}_{rgs}$ *s.t.* $R(a,b)$, *we have that* $W(a,b) \geq_{\mathbb{S}} W(\mathcal{B},a)$.

Thus, an argument b is $\oplus$-acceptable if for each attack from a against b, the aggregated weight of the collective defence of b is greater than $W(a,b)$. Such phrasing of defence is also equivalent to (Bistarelli and Santini 2010).

By using the same semiring-based framework, it is now possible to relate such notions of defence together (we remind that $\mathbb{D}_w$ stands for w-defence).

Theorem 3. $\mathbb{D}_w \Rightarrow \mathbb{D}_2$.

Proof. If $W(a, \mathcal{B} \cup \{b\}) \geq_{\mathbb{S}} W(\mathcal{B},a)$ (i.e., Def. 13 holds) then $W(a,b) \geq_{\mathbb{S}} W(\mathcal{B},a)$ (also (Coste-Marquis et al. 2012) holds), due to $W(a,b) \geq_{\mathbb{S}} W(a, \mathcal{B} \cup \{b\})$ (monotonicity of $\times$ operator, see Sec. 2.1). □

Moreover, we can link $\mathbb{D}_1$ and $\mathbb{D}_2$ as well:

Theorem 4. $\mathbb{D}_1 \Rightarrow \mathbb{D}_2$.

Proof. This is equivalent to prove that $\forall R(a,b), \exists c \in \mathcal{B}$ s.t. $W(a,b) \geq_{\mathbb{S}} W(c,a) \Rightarrow W(a,b) \geq_{\mathbb{S}} W(\mathcal{B},a)$. If such c exists, we also know that $W(c,a) \geq_{\mathbb{S}} W(\mathcal{B},a)$ holds, given Def. 12 and monotonicity of $\times$; transitivity leads to $W(a,b) \geq_{\mathbb{S}} W(\mathcal{B},a)$, proving $\Rightarrow$. □

[2] Such properties are satisfied by a c-semiring (see Sec. 2.1), *e.g.*, minimality in (Coste-Marquis et al. 2012) corresponds to the absorptivity of $\times$ w.r.t. $\perp$.

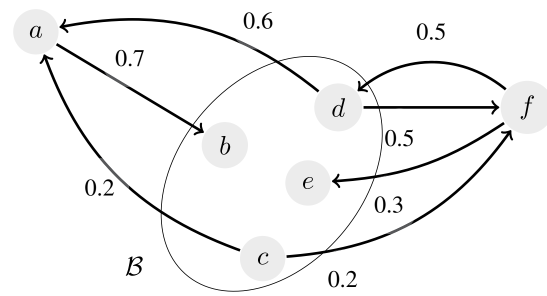

Figure 4: An example of WAAF where $\{b,d,e\}$ are defended by $\mathcal{B}$ according to $\mathbb{D}_2$ (using the weighted semiring).

In case the c-semiring we use is the fuzzy one, i.e., $\langle [0,1], \max, \min, 0, 1\rangle$, then $\mathbb{D}_2$ collapses into $\mathbb{D}_1$, as Th. 5 states.

Theorem 5. *If* $\mathbb{S} = \langle [0,1], \max, \min, 0, 1\rangle$, *then* $\mathbb{D}_1 \Leftrightarrow \mathbb{D}_2$.

Proof. This is equivalent to prove that $\forall R(a,b), \exists c \in \mathcal{B}$ s.t. $W(a,b) \geq_{\mathbb{S}} W(c,a) \Leftrightarrow W(a,b) \geq_{\mathbb{S}} W(\mathcal{B},a)$. $\Rightarrow$ can be proven for any semiring $\mathbb{S}$ (see Th. 4). In order to prove $\Leftarrow$, $W(\mathcal{B},a)$ is computed by using Def. 12 and $\min$, hence there exists at least one $c \in \mathcal{B}$ s.t. $W(a,b) \geq_{\mathbb{S}} W(c,a)$. □

This results permits to relate $\mathbb{D}_w$ and $\mathbb{D}_1$ when using the fuzzy c-semiring:

Corollary 6. *If* $\mathbb{S} = \langle [0,1], \max, \min, 0, 1\rangle$, *then* $\mathbb{D}_w \Rightarrow \mathbb{D}_1$.

Proof. This directly follows from Th. 3 when using a fuzzy c-semiring, i.e., Th. 5. □

To conclude, we show that all the three $\mathbb{D}_w$, $\mathbb{D}_1$, and $\mathbb{D}_2$ collapse to the classical defence $\mathbb{D}_0$ (Dung 1995) when considering the framework without weights.

Theorem 7. *If* $\mathbb{S} = \langle \{true, false\}, \vee, \wedge, false, true\rangle$, *then* $\mathbb{D}_w \Leftrightarrow \mathbb{D}_0 \Leftrightarrow \mathbb{D}_1 \Leftrightarrow \mathbb{D}_2$.

Proof. $\mathbb{D}_w \Leftrightarrow \mathbb{D}_0$ is proved in Prop. 2. To show $\mathbb{D}_w \Leftrightarrow \mathbb{D}_2$ we only need $\mathbb{D}_2 \Rightarrow \mathbb{D}_w$ ($\Leftarrow$ holds from Th. 3): this holds because if $W(a,b) \geq_{\mathbb{S}} W(\mathcal{B},a)$ it means that if $W(a,b)$ is *false* then $W(\mathcal{B},a)$ is *false* (due to $\geq_{\mathbb{S}}$); hence, $W(a, \mathcal{B} \cup \{b\}) \geq_{\mathbb{S}} W(\mathcal{B},a)$, i.e., $\mathbb{D}_w$, holds as well. To show $\mathbb{D}_1 \Leftrightarrow \mathbb{D}_2$ we only need $\mathbb{D}_2 \Rightarrow \mathbb{D}_1$ ($\Leftarrow$ holds from Th. 4): similarly, if $W(\mathcal{B},a)$ is *false*, than $\exists c \in \mathcal{B}$ s.t. $W(a,b) \geq_{\mathbb{S}} W(c,a)$, since $\exists c \in \mathcal{B}$ s.t. $W(c,a) = false$ (i.e., c attacks a). □

An example on how $\mathbb{D}_w$, $\mathbb{D}_1$, and $\mathbb{D}_2$ differently work is provided in Fig. 4. We read this example by considering the weighted c-semiring, i.e.,. $\mathbb{S} = \langle \mathbb{R}^+ \cup \{+\infty\}, min, \hat{+}, +\infty, 0\rangle$. Argument b is defended by $\mathcal{B} = \{b,c,d,e\}$ according to $\mathbb{D}_w$ and (consequently) $\mathbb{D}_2$, since $W(a, \mathcal{B} \cup \{b\}) \geq_{\mathbb{S}} W(\mathcal{B},a)$ $(0.7 \leq 0.8)$. It is not defended according to $\mathbb{D}_1$, since $W(d,a) \geq_{\mathbb{S}} W(a,b)$ $(0.6 \leq 0.7)$

and $W(c,a) \geq_{\mathbb{S}} W(a,b)$ ($0.2 \leq 0.7$). On the other hand, considering the attacks from f instead, $\mathbb{D}_w$ does not hold: $W(f,d) \times W(f,e) \not\geq_{\mathbb{S}} W(d,f) \times W(c,f)$ (i.e., $0.8 \not\leq 0.7$); however, $\mathbb{D}_2$ holds because $W(f,d) \geq_{\mathbb{S}} W(\{d,c\},f)$ (i.e., $0.5 \leq 0.7$) and $W(f,e) \geq_{\mathbb{S}} W(\{d,c\},f)$ (i.e., $0.3 \leq 0.7$). With respect to the attacks from f, even $\mathbb{D}_2$ holds, since $W(f,d) \leq W(d,f)$ and $W(f,e) \leq W(d,f)$. Therefore, over the whole WAAF in Fig. 4, only $\mathbb{D}_2$ holds. Reading the same example in Fig. 4 with $\mathbb{S} = \langle [0,1], \max, \min, 0, 1\rangle$ instead, $\mathbb{D}_2$ collapses to $\mathbb{D}_1$ (Th. 5) and $\mathbb{D}_1$ does not hold due to $R(a,b)$. According to Th. 6, since $\mathbb{D}_1$ is not valid then $\mathbb{D}_w$ cannot hold as well.

In the concluding part of this section we study how admissible semantics are related considering $\mathbb{D}_w$, $\mathbb{D}_1$, and $\mathbb{D}_2$ in the fuzzy c-semiring. We focus on this semantics because it is at the core of the other ones proposed in (Dung 1995) (see Sec. 2.2), explicitly (i.e., complete, preferred, grounded), or implicitly (i.e., stable). We respectively call adm_1 and adm_2 the set of admissible extensions using $\mathbb{D}_1$ and $\mathbb{D}_2$, adm_w is our proposal (using $\mathbb{D}_w$ in Def. 15), and adm_0 adopt the classical definition of defence $\mathbb{D}_0$ (Dung 1995).

Theorem 8. *Given* $WF = \langle \mathcal{A}_{rgs}, R, W, \mathbb{S}\rangle$ *where* $\mathbb{S} = \langle [0,1], \max, \min, 0, 1\rangle$*, then* $adm_w(WF) = adm_1(WF) = adm_2(WF) \subseteq adm_0(WF)$.

Proof. $adm_1(WF) = adm_2(WF)$ directly derives from $\mathbb{D}_1 \Leftrightarrow \mathbb{D}_2$ (see Th. 5). In order to prove $adm_w(WF) = adm_1(WF) = adm_2(WF)$, since we have already proved Th. 6, we only need to show that $\mathbb{D}_1 \Rightarrow \mathbb{D}_w$, i.e., $\forall R(a,b), \exists c \in \mathcal{B}$ s.t. $W(a,b) \geq_{\mathbb{S}} W(c,a) \Rightarrow W(a, \mathcal{B} \cup \{b\}) \geq_{\mathbb{S}} W(\mathcal{B}, a)$. Since we need to prove that $\mathcal{B}$ is w-admissible, $b \in \mathcal{B}$. Therefore, $\exists c \in \mathcal{B}$ s.t. $W(a,b) \geq_{\mathbb{S}} W(a,\mathcal{B}) \geq_{\mathbb{S}} W(c,a)$, since $W(a,\mathcal{B})$ is the worst $(\min)$ of the attacks from a to $\mathcal{B}$. Given $W(c,a) \geq_{\mathbb{S}} W(\mathcal{B},a)$ (see Th. 6), consequently we have $W(a,\mathcal{B}) \geq_{\mathbb{S}} W(\mathcal{B},a)$. Finally, due to Prop. 1, we have the last inclusion of the theorem, i.e., $adm_w(WF) \subseteq adm_0(WF)$. □

6 Conclusion and Future Work

In this work we have defined a new notion of defence for WAAFs. Since defence is collective in the literature (i.e., it considers all the counter-attacks from $\mathcal{B}$ as a whole), our main motivation is to provide a similar view also for all the attacks from a to $\mathcal{B}$, here considered by summing all the attacks weights together. In addition, by casting similar proposals (Coste-Marquis et al. 2012; Martínez, García, and Simari 2008) in the same parametric algebraic-framework, it is possible to show all their relations in detail.

In the future we plan to compute a relaxation of w-defence in order to let it exactly match again to the classical definition of defence given by Dung. Our intent is also to have a computational framework where it is possible to relax *i)* w-defence and *ii)* the internal conflict of extensions, by allowing also the possibility to tolerate small amount of attack strength-weights in the conflict-free semantics (and, consequently in all the semantics based on it). Either by relaxing *i* or *ii* the number of extensions (for a given semantics) increases; therefore, we want to play on such two effects singly and simultaneously on randomly-generated and real-world frameworks to study the outcome and formalise further properties on WAAFs. In addition we plan to extend w-defence to coalitions as well (Bistarelli and Santini 2013).

References

Amgoud, L., and Cayrol, C. 1998. On the acceptability of arguments in preference-based argumentation. In *UAI '98: Proceedings of the Fourteenth Conference on Uncertainty in Artificial Intelligence*, 1–7. Morgan Kaufmann.

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 Santini, F. 2010. A common computational framework for semiring-based argumentation systems. In Coelho, H.; Studer, R.; and Wooldridge, M., eds., *ECAI*, volume 215 of *Frontiers in Artificial Intelligence and Applications*, 131–136. IOS Press.

Bistarelli, S., and Santini, F. 2013. Coalitions of arguments: An approach with constraint programming. *Fundam. Inform.* 124(4):383–401.

Bistarelli, S.; Montanari, U.; and Rossi, F. 1997. Semiring-based constraint satisfaction and optimization. *J. ACM* 44(2):201–236.

Coste-Marquis, S.; Konieczny, S.; Marquis, P.; and Ouali, M. A. 2012. Weighted attacks in argumentation frameworks. In *Principles of Knowledge Representation and Reasoning: Proceedings of the Thirteenth International Conference, KR 2012, Rome, Italy, June 10-14, 2012*. 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):321–357.

Dunne, P. E.; Hunter, A.; McBurney, P.; Parsons, S.; and Wooldridge, M. 2009. Inconsistency tolerance in weighted argument systems. In *Conf. on Autonomous Agents and Multiagent Systems*, 851–858. IFAAMS.

Dunne, P. E.; Hunter, A.; McBurney, P.; Parsons, S.; and Wooldridge, M. 2011. Weighted argument systems: Basic definitions, algorithms, and complexity results. *Artif. Intell.* 175(2):457–486.

Janssen, J.; Cock, M. D.; and Vermeir, D. 2008. Fuzzy argumentation frameworks. In *Information Processing and Management of Uncertainty in Knowledge-based Systems*, 513–520.

Li, H.; Oren, N.; and Norman, T. J. 2011. Probabilistic argumentation frameworks. In *Theories and Applications of Formal Argumentation - First International Workshop, Revised Selected Papers*, volume 7132 of *Lecture Notes in Computer Science*, 1–16. Springer.

Martínez, D. C.; García, A. J.; and Simari, G. R. 2008. An abstract argumentation framework with varied-strength attacks. In *Principles of Knowledge Representation and Reasoning: Proceedings of the Eleventh International Conference*, 135–144. AAAI Press.

Modgil, S. 2009. Reasoning about preferences in argumentation frameworks. *Artif. Intell.* 173(9-10):901–934.