International Journal of Approximate Reasoning ••• (••••) •••–•••

Contents lists available at ScienceDirect

International Journal of Approximate Reasoning

www.elsevier.com/locate/ijar

A novel weighted defence and its relaxation in abstract argumentation ☆

Stefano Bistarelli, Fabio Rossi, Francesco Santini *

Department of Mathematics and Computer Science, University of Perugia, Italy

A R T I C L E   I N F O

Article history:
Received 15 October 2016
Received in revised form 1 August 2017
Accepted 2 October 2017
Available online xxxx

Keywords:
Abstract Argumentation Frameworks
Weighted attacks
Inconsistency tolerance
Relaxation
Semirings

A B S T R A C T

When dealing with Weighted Abstract Argumentation, having weights on attacks clearly brings more information. The advantage, for instance, is the possibility to define a different notion of defence, checking also if the weight associated with it is stronger than the attack weight. In this work we study two different relaxations, one related to the new weighted defence we propose, by checking the difference between the composition of inward and outward attack-weights. The second one is related to how much inconsistency we are willing to tolerate inside an extension; such amount is computed by aggregating the costs of the attacks between any two arguments both inside an extension. These two relaxations are strictly linked: allowing a small conflict may lead to have more arguments into an extension, and consequently result in a stronger or weaker defence. Weights are represented by a semiring structure, which can be instantiated to different metrics used in the literature (*e.g.*, costs, probabilities, fuzzy levels).

1. Introduction

An *Abstract Argumentation Framework* (*AAF*) [1] is represented by 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 collectively 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 prevails. However, having weights on attacks results in such additional information, which can be fruitfully exploited in this direction. For instance, in case the attack $R(a, b)$ is stronger than (or preferred to) $R(b, a)$, taking the position defined by a may result in a better choice for an intelligent agent, since it can be defended better.

The aim of this work is to first *i)* propose a new notion of weighted defence and, from this, relax classically exact and sharp concepts in *Weighted Abstract Argumentation Frameworks* (*WAAFs*, see Section 11). This is accomplished by allowing *ii)* an internal conflict *inside* extensions satisfying a given extension-based semantics, and *iii)* by relaxing the defence of arguments w.r.t. the attacks coming from *outside* an extension. These are the three main results of the work.

☆ This paper is part of the Virtual special issue on Uncertainty Reasoning, Edited by Robert E. Mercer and Salem Benferhat.

* Corresponding author.
E-mail addresses: bista@dmi.unipg.it (S. Bistarelli), rossi@dmi.unipg.it (F. Rossi), francesco.santini@dmi.unipg.it (F. Santini).

https://doi.org/10.1016/j.ijar.2017.10.006
0888-613X/

The first goal is to provide a new definition of defence for WAAFs, here called *w-defence*, which encompasses weights in the style of similar works, as [2] and [3]. 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 composition (a parametric $\otimes$ operation from a *c-semiring* structure [4]) of all the attack weights from $\mathcal{B}$ to a is stronger than the composition of *all* the attacks from a to $\mathcal{B}$. Differently from [2], 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 $\mathcal{B}$. Therefore, both our proposal and the one given by [2] suggest a collective defence from $\mathcal{B}$ to a, but, differently, in this paper we consider the group of attacks from a to $\mathcal{B}$ as a single entity, i.e., with a single global weight. We believe such a 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 admissibility), but each attack from a to $\mathcal{B}$ is treated separately (however, in case of fuzzy aggregation of weights, the two approaches are equivalent). Our intent is to normalise such dis-homogeneity.

Once having defined w-defence, we then proceed with the relaxation of the framework. Such two issues represent orthogonal relaxations, concerning either only internal arguments, or the relationship between the set of acceptable and not-acceptable arguments (interval versus external arguments). In this way, an autonomous reasoning-agent has more instruments to understand, for instance, whether tolerating a small conflict among its arguments considerably changes its point of view. As a possible scenario, a debate can be permeated by arguments advanced by *trolls* [5], which can accordingly generate noise in an abstract framework. Mitigating the inconsistency produced by them may let other arguments grow in strength (see Section 9). Internal inconsistency arises in many areas of AI and computing: merging information from heterogeneous sources, negotiation in multi-agent systems, or understanding natural language dialogues [6]. On the other hand, an agent could also be interested in defending its arguments with a higher or lower level of strength, in order to respectively strengthen the defence or increase the number of satisfied extensions. For instance, increasing both α and γ, something close to a stable extension can appear (there are AAF where the stable semantics is not satisfied).

In the end, we design α^γ-semantics (Section 6), where α is the amount of weight tolerated inside an extension satisfying it, and γ is the weight difference between a "full" defence (i.e., w-defence) and the "weaker" defence implemented by an α^γ-extension. We obtain them from classical formulations [1], hence we call them α^γ-conflict-free and α^γ-admissible sets, α^γ-complete, α^γ-preferred, and α^γ-stable semantics. These two parameters strictly influence each other: by relaxing α one can allow one or more new arguments be accepted into an extension, but, at the same time, their attacks contribute to the defence strength, i.e., to the computation of γ: if a is accepted in $\mathcal{B}$ because α is increased, $\mathcal{B} \cup \{a\}$ could be not γ-defensible anymore, or, on the other hand, it could become defensible even with a stricter γ. An agent can play with these two parameters with the purpose to "explore the neighbourhoods" of classical extensions, and take different decisions.

To have a general and formal representation of weights and operations on them (i.e., aggregation and preference), in this work we instead adopt a parametric algebraic framework based on c-semirings [4]. Hence, it is possible to consider different metrics within the same computational framework, as fuzzy or probabilistic scores, and model different kinds of AAFs in the literature (see Section 10). This represents a further result of the paper.

This paper elaborates on [7] and extends the works in [8] and [9]. All these papers are here integrated to offer an thorough view on the topic, by providing proofs, extended examples, and tests, which are not present in the single contributions; for instance, the case-study in Section 9.2 is new. The paper is structured as follows: Section 2.1 recollects the basic definitions of AAF given by [1], while in Section 2.2 we introduce c-semirings. Section 3 presents WAAFs, w-defence, and Section 4 reports a comparison with related notions in the literature [2,3]. Section 5 relaxes w-defence by proposing γ-defence, where γ is the amount by which defence is weakened. In Section 6 we propose α^γ-semantics (*e.g.*, α^γ-stable), which extend classical ones by considering α and γ at the same time. Section 7 collects some formal results concerning such new semantics, *e.g.*, inclusion relations. In Section 8 we briefly describe an implementation of the proposed framework, together with some tests on random WAAFs. Section 9 presents two different case-studies to better motivate the formal results in the paper: one example is based on a real-world case-study, where the considered WAAF is directly extracted from *Amazon.com* reviews on a chosen product (a ballet tutu for kids). In Section 10 we describe related work, and, finally, Section 11 wraps up the paper by drawing final conclusions and suggesting future work.

2. Background

In the following of this section we first recollect the main definitions at the basis of AAFs [1] (Section 2.1), and then introduce c-semirings (Section 2.2). C-semirings represent a parametric framework where to deal with attack-weights. By changing the underlying c-semiring instantiation, it is possible to capture different metrics (*e.g.*, fuzzy or probabilistic ones).

2.1. Abstract Argumentation Frameworks

In his pioneering work [1], 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 1. 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 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).

Please cite this article in press as: S. Bistarelli et al., A novel weighted defence and its relaxation in abstract argumentation, Int. J. Approx. Reason. (2017), https://doi.org/10.1016/j.ijar.2017.10.006