Gradual Semantics for Probabilistic Argumentation Frameworks

A 30 ECTS thesis submitted to the master Artificial Intelligence program

Jeroen Paul Spaans

Student ID: 6612385

Supervisor:

Dr. Dragan Doder

Second Examiner:

Dr. Annemarie Borg

Utrecht University

27th July 2023

Abstract. Gradual semantics are methods that evaluate overall strengths of individual arguments in graphs. In this thesis, we investigate gradual semantics for extended frameworks in which probabilities are used to quantify the uncertainty about arguments and attacks belonging to the graph. We define the likelihoods of an argument's possible strengths when facing uncertainty about the topology of the argumentation framework. We also define an approach to compare the strengths of arguments in this probabilistic setting. Finally, we propose a method to calculate the overall strength of each argument in the framework, and we evaluate this method against a set of principles.

Contents

1. Introduction

Within the last couple of decades, argumentation has emerged as a popular field in Artificial Intelligence [9, 15]. It has been shown to be useful in several domains, such as decision making [46], reasoning under inconsistency [17] and non-monotonic reasoning [34] and is applicable in the domains of law and medicine [9]. The underlying structure of formal models of abstract argumentation takes the form of directed graphs, whose nodes represent arguments and whose directed edges indicate attacks between attacks.

Two main classes of semantics were proposed to reason about such structures and to evaluate arguments in the graphs. *Extension-based* semantics are proposed with the goal of identifying jointly acceptable sets of arguments (extensions) based on specific properties within the graph [23]. The acceptability status of an argument is then derived from these extensions. The argument is *sceptically accepted* if it belongs to all extensions, *credulously accepted* if it belongs to some of the extensions, and *rejected* otherwise. On the other hand, *gradual semantics* [21] focus on individual arguments and quantify their strengths in graphs, using a richer scale (usually the unit interval of reals $[0, 1]$). They typically define strength of an argument to depend on strengths of its direct attackers. One intuitive difference between gradual and extension-based semantics is that, in the latter approach, the attack relation is used to destroy its target (two conflicting arguments cannot be in the same extension), while in ranking based semantics it is often used to only weaken its target. Examples of gradual semantics are h-Categoriser [17], Simple product semantics [32], Trust-based semantics [38], Iterative Schema [25], Max-based and Cardinality-based semantics [8]. Some of the approaches were adapted to frameworks where arguments and/or attacks have a base weight [5, 6, 7, 8], bipolar frameworks in which both support and attack relations are present in a graph [4, 40, 41], and weighted bipolar SETAFs where a set of arguments can attack or support a target together [45]. Gradual semantics are similar in spirit to *ranking semantics* [2, 20], which focus on the strength of arguments relative to that of other arguments, and return a preorder on arguments, thus ranking them from the strongest to the weakest ones. Obviously, each gradual semantics can be used to generate a ranking semantics (but not vice versa).

For many applications in which there is uncertainty about topology of the argumentation graph, simple attack frameworks appear too simple for convenient modelling of those aspects of an argumentation problem. There are different scenarios in which uncertainty about the presence of arguments and attacks in a graph arises naturally: at times, ambiguities in the language used for presenting an argument, or the presentation of arguments with incomplete premises or claims may lead to uncertainty about the correct interpretation of attacks between arguments [28]; other times, arguments are presented with explicit uncertainty in their claims [28]; an audience to some argumentation is often unsure of the exact set of arguments being put forward, and a participant in some argumentation may be unsure which arguments the audience has in mind [30]. In order to handle these uncertainties, Li, Oren and Norman [33] proposed Probabilistic Argumentation Frameworks which augment argumentation graphs with probabilities. In this approach, named the *constellations approach* by Hunter[1][27], probability values are added to the arguments and attacks, allowing for the modelling of uncertainty in which elements should be present in the argumentation framework. In the constellations approach, a considered extension semantics is used to determine the probability of an argument being (credulously/sceptically) accepted. This approach to probabilistic argumentation has been extensively investigated from an extension-based semantic point of view [26, 27, 28, 30, 39, 18, 35, 29, 37, 19], but never from the perspective of gradual semantics. Therefore, the current results on the constellations approach are well suited to answering the question of probability of (joint) acceptance of (sets of) arguments, but not the questions about probability of strength of an individual argument, or the question which argument is stronger in a probabilistic setting. Such questions are instead a particularly good match for applications of gradual semantics to graphs augmented with probabilities.

The aim of this thesis is to study gradual semantics in probabilistic argumentation frameworks, following the constellations approach. Working towards this goal we set ourselves the following objectives: Our first goal is to define semantics that return probabilities of an argument's acceptance with respect to any strength threshold, providing a richer scale of acceptability statuses than would be possible using Dung semantics. The second goal is to employ rankings to determine the probability that one argument is stronger than another and to extend this ranking-based approach to enable us to find the probability of certain, possibly more complex, *ranking queries* of interest being satisfied. This, for example, would allow us to calculate the probability that the argument a is stronger then either b or c. For both of these objectives, our intent is to investigate the formal properties of and connections between the approaches adopted in pursuing them. Our third goal is to investigate the challenging problem of defining semantics that assign unique overall strength to each argument in a probabilistic argumentation framework. In this investigation we aim to propose desirable principles for such semantics, following the constellations approach, inspired by existing principles for gradual semantics from the literature

[1]In the same paper, the author introduces the notion of the *epistemic approach* to probabilistic argumentation, which can be used to represent the degree to which an argument is believed [27, 31, 42].

[5]. We intend to investigate properties of those principles and show their compatibility. Finally, we aim to propose the first family of such semantics, by providing a method for generalising gradual semantics for argumentation graphs to our probabilistic framework. We mean to show that if the considered underlying gradual semantics satisfies existing principles [5], then its generalisation satisfies our novel principles.

The remainder of this thesis is organised as follows: Section 2 provides the formal background to the sections that follow it. In section 3, we pursue our first goal and investigate the use of gradual semantics with PrAFs to determine the probability of acceptability of arguments. Section 4 is concerned with the comparison of arguments in a PrAF in pursuit of our second goal. In it, we employ gradual semantics to determine argument rankings and investigate the probability that a demand on such rankings is met in a PrAF under a gradual semantics. In Section 5, in pursuit of our third and final goal, means of assigning each argument in a PrAF an overall strength using gradual semantics are investigated in keeping with the constellations approach to probabilistic argumentation. Finally, in Section 6, we discuss work related to this thesis, we discuss work that may follow this thesis in the future, and conclude on the work presented in this thesis.

2. Background

2.1. **Dung's Argumentation Framework.** The core concept that all topics we will introduce hereafter relate to is that of *argumentation graphs*, introduced by Dung in his seminal 1995 paper [23] under the name argumentation frameworks; a formalism based on the notion that arguments are defeasible and may attack each other and that deciding which arguments to accept requires evaluation of the attacks between arguments. Given that conflict between arguments is possible, whether an argument can reasonably be accepted depends on the existence of counterarguments, which may in turn be subject to their own counterarguments, and so on. The argumentation framework allows for the evaluation of a set of arguments by placing them in a directed graph, where arguments make up the nodes and attacks between arguments are modelled as the edges between them. Formally we have:

Definition 2.1 (Argumentation Graph)**.** *An argumentation graph, or AG, is an ordered pair $\mathbf{G} = \langle \mathcal{A}, \mathcal{R} \rangle$, where $\mathcal{A}$ is a non-empty finite set of arguments and $\mathcal{R} \subseteq \mathcal{A} \times \mathcal{A}$ is an attack relation between arguments. Let* $\mathtt{AG}$ *denote the set of all argumentation graphs.*

Considering argumentation graphs, we have the following:

- We say an argument $a \in \mathcal{A}$ attacks an argument $b \in \mathcal{A}$, $a\mathcal{R}b$ for short, iff $(a, b) \in \mathcal{R}$.
- If an argument $a \in \mathcal{A}$ attacks an argument $b \in \mathcal{S} \subseteq \mathcal{A}$, we say a attacks $\mathcal{S}$.
- Similarly we say $\mathcal{S} \subseteq \mathcal{A}$ attacks $a \in \mathcal{A}$ iff there exists some $b \in \mathcal{S}$ such that b attacks a.
- An argument $a \in \mathcal{A}$ or set of arguments $\mathcal{S} \subseteq \mathcal{A}$ defends an argument $b \in \mathcal{A}$ if it attacks all arguments $c \in \mathcal{A}$ that attack b.
- Likewise, an argument $a \in \mathcal{A}$ or set of arguments $\mathcal{S} \subseteq \mathcal{A}$ defends a set of arguments $\mathcal{S}' \subseteq \mathcal{A}$ if it defends all members of $\mathcal{S}'$.
- The function $F_{\mathbf{G}} : 2^{\mathcal{A}} \to 2^{\mathcal{A}}$ such that for $\mathcal{S} \subseteq \mathcal{A}$ $F_{\mathbf{G}}(\mathcal{S}) = \{a \mid \mathcal{S} \text{ defends } a\}$ is called the characteristic function of $\mathbf{G}$.

The argumentation graph formalism considers arguments and attacks as purely abstract entities, doing away entirely with features such as the structure and origin of arguments and the nature of attacks. This simplification makes it so the framework can be regarded more as a calculus of conflict, applicable to a wide variety of areas only loosely related to argumentation such as nonmonotonic reasoning and game theory. This broad applicability helps explain the academic interest the formalism has enjoyed, even as a subject upon itself rather than a tool to be used in another setting.

We see a very simple example of an argumentation graph in Figure 1.

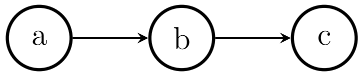

FIGURE 1. A simple argumentation graph with arguments a, b, and c and attacks (a, b) and (b, c).

With argumentation frameworks encoding sets of arguments and the conflict between them, Dung seeks to identify sets of arguments in an argumentation graph, called extensions, that together represent a reasonable position one might take. The first constraint for such an extensions to be considered reasonable is that they should be internally consistent.

Definition 2.2 (Conflict-freeness)**.** *Let* $\mathbf{G} = \langle \mathcal{A}, \mathcal{R} \rangle$ *be an argumentation graph. A set of arguments* $\mathcal{S} \subseteq \mathcal{A}$ *is conflict-free iff there are no* $a, b \in \mathcal{S}$ *s.t.* $(a, b) \in \mathcal{R}$.

As a minimum, one more constraint is placed on reasonable extensions. We require that such an extension defends all of its members from outside attacks.

Definition 2.3 (Admissibility)**.** *Let* $\mathbf{G} = \langle \mathcal{A}, \mathcal{R} \rangle$ *be an argumentation graph. A set of arguments* $\mathcal{S} \subseteq \mathcal{A}$ *is admissible iff it is conflict-free and it defends all its elements.*

Further constraints on these extensions produce special classes of admissible sets of arguments. As a mechanism for identifying such extensions, Dung introduces semantics—commonly collectively referred to as Dung semantics or extension semantics—that map an AG to its extensions belonging to a specific class.

Definition 2.4 (Extension Semantics)**.** *An extension-based semantics is a function* $\mathbf{S}$ *mapping any argumentation framework* $\mathbf{G} = \langle \mathcal{A}, \mathcal{R} \rangle$ *to a subset of* $2^{\mathcal{A}}$*, denoted as* $\mathtt{E}_{\mathbf{S}}(\mathbf{G})$.

- *Complete semantics* $\mathtt{Co}$ *maps any* $\mathbf{G} = \langle \mathcal{A}, \mathcal{R} \rangle$ *to its complete extensions: those* $\mathcal{S} \subseteq \mathcal{A}$ *that are conflict-free and for which it holds that* $\mathcal{S} = F_{\mathbf{G}}(\mathcal{S})$.
- *Grounded semantics* $\mathtt{Gr}$ *maps any* $\mathbf{G} = \langle \mathcal{A}, \mathcal{R} \rangle$ *to its grounded extension: its minimal (w.r.t. set inclusion) complete extension.*
- *Preferred semantics* $\mathtt{Pr}$ *maps any* $\mathbf{G} = \langle \mathcal{A}, \mathcal{R} \rangle$ *to its preferred extensions: its maximal (w.r.t. set inclusion) complete extensions.*
- *Stable semantics* $\mathtt{St}$ *maps any* $\mathbf{G} = \langle \mathcal{A}, \mathcal{R} \rangle$ *to its stable extensions: those of its preferred extensions* $\mathbf{S}$ *that attack all arguments in* $\mathcal{A} \setminus \mathbf{S}$.

We say an argument is sceptically accepted under an extension semantics if it belongs to all of its extensions, credulously accepted if it belongs to at least one of its extensions, and rejected otherwise.

Example 1 (A running example)**.** The argumentation graph presented in Figure 1 consists of arguments a, b, and c with $a\mathcal{R}b$ and $b\mathcal{R}c$. The extension of this graph characterised by grounded semantics, is $\{a, c\}$. As this is the only extension produced, these arguments will both be sceptically accepted. Informally we have the following: a is part of the extension because it is not attacked. Including a requires us to exclude b as a attacks b. This attack that rules out b then defends c from the attack it receives, which allows us to include it.

2.2. **Gradual Semantics for Argumentation Graphs.** As we explained in the introduction, gradual semantics do not identify sets of arguments that are acceptable together, as extension semantics do, but rather assign to each argument a unique overall strength value or acceptability degree considering the strengths of their attackers. This richer evaluative scale, where arguments with a higher acceptability degree are considered more acceptable, enables us to make comparisons between alternative arguments, even when both would be assigned the same acceptability status by a Dung semantics.

Following an approach already accepted by some authors from the field [8], for simplicity we use the unit interval of reals as the evaluative scale.

Definition 2.5 (Weighting)**.** [8] *A weighting on a set $\mathcal{X}$ is a function from $\mathcal{X}$ to the interval* $[0, 1]$.

Now we can define gradual semantics in a formal way.

Definition 2.6 (Gradual Semantics)**.** [8] *A semantics is a function* $\mathbf{S}$ *transforming any argumentation graph* $\mathbf{G} = \langle \mathcal{A}, \mathcal{R} \rangle \in \mathtt{AG}$ *into a weighting* $\mathtt{Deg}^{\mathbf{S}}_{\mathbf{G}}$ *on* $\mathcal{A}$ *(i.e.,* $\mathtt{Deg}^{\mathbf{S}}_{\mathbf{G}} : \mathcal{A} \to [0, 1]$*). For any* $a \in \mathcal{A}$*,* $\mathtt{Deg}^{\mathbf{S}}_{\mathbf{G}}(a)$ *represents the strength of* a.

A well-studied gradual semantics that will be used as an example in this work is the h-categoriser, proposed by Besnard and Hunter [17]:

Definition 2.7. *The h-categoriser is a gradual semantics* $\mathtt{Hbs}$ *s.t.* $\forall \mathbf{G} = \langle \mathcal{A}, \mathcal{R} \rangle \in \mathtt{AG}$, $\forall a \in \mathcal{A}$,

$$\mathtt{Deg}^{\mathtt{Hbs}}_{\mathbf{G}}(a) = \frac{1}{1 + \sum_{b_i \in \mathtt{Att}_{\mathbf{G}}(a)} \mathtt{Deg}^{\mathtt{Hbs}}_{\mathbf{G}}(b_i)}.$$

Example 2 (Running example, continued)**.** Applying $\mathtt{Hbs}$ to the argumentation graph presented in Figure 1, we get that $\mathtt{Deg}^{\mathtt{Hbs}}_{\mathbf{G}}(a) = 1$, $\mathtt{Deg}^{\mathtt{Hbs}}_{\mathbf{G}}(b) = 1/2$, and $\mathtt{Deg}^{\mathtt{Hbs}}_{\mathbf{G}}(c) = 2/3$. Comparing this to Example 1 we note that under grounded semantics arguments a and b received the same status, "accepted", while under the h-categoriser argument a is stronger than argument c and we note that while argument b was not accepted under grounded semantics (the attack from a ruled out its acceptance) it receives a positive but lowered strength under the h-categoriser.

In the literature on gradual semantics, several works have been devoted to development of *principles* that represent desirable formal properties of semantics, with the purpose to serve as a tool for analysis and comparison of gradual semantics [3, 5, 8, 10, 12].

In what follows, we recall some principles from [5], adjusted to non-weighted graphs.

Before the first principle, we recall what it means for an isomorphism between argumentation graphs to exist.

Definition 2.8 (Isomorphism)**.** *Let* $\mathbf{G} = \langle \mathcal{A}, \mathcal{R} \rangle$ *and* $\mathbf{G}' = \langle \mathcal{A}', \mathcal{R}' \rangle$ *be two AGs. An isomorphism from* $\mathbf{G}$ *to* $\mathbf{G}'$ *is a bijective function* f *from* $\mathcal{A}$ *to* $\mathcal{A}'$ *such that* $\forall a, b \in \mathcal{A}, a\mathcal{R}b$ *iff* $f(a)\mathcal{R}'f(b)$.

The first principle, *Anonymity*, requires that an argument's strength depends only on the topology of the argumentation graph. I.e. the argument's name and contents are irrelevant to its strength.

Principle 2.9 (Anonymity)**.** *A gradual semantics* $\mathbf{S}$ *satisfies Anonymity iff for any AGs* $\mathbf{G} = \langle \mathcal{A}, \mathcal{R} \rangle$ *and* $\mathbf{G}' = \langle \mathcal{A}', \mathcal{R}' \rangle$*, for any isomorphism* f *from* $\mathbf{G}$ *to* $\mathbf{G}'$*, the following holds:* $\forall\, a \in \mathcal{A}, \mathtt{Deg}^{\mathbf{S}}_{\mathbf{G}}(a) = \mathtt{Deg}^{\mathbf{S}}_{\mathbf{G}'}(f(a))$.

Independence states that an argument's strength should not depend on arguments that are not connected to it. Here, for any $\mathbf{G} = \langle \mathcal{A}, \mathcal{R} \rangle, \mathbf{G}' = \langle \mathcal{A}', \mathcal{R}' \rangle \in \mathtt{AG}$ s.t. $\mathcal{A} \cap \mathcal{A}' = \emptyset$, $\mathbf{G} \oplus \mathbf{G}'$ is the AG $\langle \mathcal{A} \cup \mathcal{A}', \mathcal{R} \cup \mathcal{R}' \rangle$.

Principle 2.10 (Independence)**.** *A gradual semantics* $\mathbf{S}$ *satisfies Independence iff for any AGs* $\mathbf{G} = \langle \mathcal{A}, \mathcal{R} \rangle$ *and* $\mathbf{G}' = \langle \mathcal{A}', \mathcal{R}' \rangle$ *s.t* $\mathcal{A} \cap \mathcal{A}' = \emptyset$*, the following holds:* $\forall\, a \in \mathcal{A}$, $\mathtt{Deg}^{\mathbf{S}}_{\mathbf{G}}(a) = \mathtt{Deg}^{\mathbf{S}}_{\mathbf{G} \oplus \mathbf{G}'}(a)$ *where* $\mathbf{G} \oplus \mathbf{G}' = \langle \mathcal{A} \cup \mathcal{A}', \mathcal{R} \cup \mathcal{R}' \rangle$.

The *Directionality* principle expresses that an argument's strength may only depend on arguments from which a path to it exists. I.e. for any $\mathbf{G} = \langle \mathcal{A}, \mathcal{R} \rangle \in \mathtt{AG}$, the strength of an argument $a \in \mathcal{A}$ may only depend on an argument $b \in \mathcal{A}$ if there exists a finite non-empty sequence $\langle x_1, \ldots, x_n \rangle$ of arguments $x_i \in \mathcal{A}$ s.t. $x_1 = b, x_n = a$ and $\forall i < n, x_i \mathcal{R} x_{i+1}$.

Principle 2.11 (Directionality)**.** *A gradual semantics* $\mathbf{S}$ *satisfies Directionality iff for any AGs* $\mathbf{G} = \langle \mathcal{A}, \mathcal{R} \rangle$ *and* $\mathbf{G}' = \langle \mathcal{A}, \mathcal{R}' \rangle$ *s.t.* $\mathcal{R}' = \mathcal{R} \cup \{(a, b)\}$ *and* $\forall r \in \mathcal{R}, P'_{\mathcal{R}}(r) = P_{\mathcal{R}}(r)$*, it holds that:* $\forall x \in \mathcal{A}$*, if there is no path from* b *to* x*,* $\mathtt{Deg}^{\mathbf{S}}_{\mathbf{G}}(a) = \mathtt{Deg}^{\mathbf{S}}_{\mathbf{G}'}(f(a))$.

Equivalence ensures that an argument's overall strength depends only on the strengths of its attackers.

Principle 2.12 (Equivalence)**.** *A gradual semantics* $\mathbf{S}$ *satisfies Equivalence iff for any AG* $\mathbf{G} = \langle \mathcal{A}, \mathcal{R} \rangle$*,* $\forall a, b \in \mathcal{A}$*, the following holds: if there exists a bijective function* f *from* $\mathtt{Att}_{\mathbf{G}}(a)$ *to* $\mathtt{Att}_{\mathbf{G}}(b)$ *s.t.* $\forall x \in \mathtt{Att}_{\mathbf{G}}(a)$*,* $\mathtt{Deg}^{\mathbf{S}}_{\mathbf{G}}(x) = \mathtt{Deg}^{\mathbf{S}}_{\mathbf{G}}(f(x))$*, then* $\mathtt{Deg}^{\mathbf{S}}_{\mathbf{G}}(a) = \mathtt{Deg}^{\mathbf{S}}_{\mathbf{G}}(b)$.

Maximality states that a non-attacked argument will have maximal strength (i.e., the strength 1).

Principle 2.13 (Maximality)**.** *A gradual semantics* $\mathbf{S}$ *satisfies Maximality iff for any AG* $\mathbf{G} = \langle \mathcal{A}, \mathcal{R} \rangle$*,* $\forall a \in \mathcal{A}$*, it holds that: if* $\mathtt{Att}_{\mathbf{G}}(a) = \emptyset$*, then* $\mathtt{Deg}^{\mathbf{S}}_{\mathbf{G}}(a) = 1$*.*

The *Neutrality* principle states that an argument with strength zero has no impact on the strength of any argument it attacks.

Principle 2.14 (Neutrality)**.** *A gradual semantics* $\mathbf{S}$ *satisfies Neutrality iff for any AG* $\mathbf{G} = \langle \mathcal{A}, \mathcal{R} \rangle$*,* $\forall a, b \in \mathcal{A}$*, it holds that: if* $\mathtt{Att}_{\mathbf{G}}(b) = \mathtt{Att}_{\mathbf{G}}(a) \cup \{x\}$ *s.t.* $x \in \mathcal{A} \setminus \mathtt{Att}_{\mathbf{G}}(a)$ *and* $\mathtt{Deg}^{\mathbf{S}}_{\mathbf{G}}(x) = 0$*, then* $\mathtt{Deg}^{\mathbf{S}}_{\mathbf{G}}(a) = \mathtt{Deg}^{\mathbf{S}}_{\mathbf{G}}(b)$*.*

Weakening states that if an argument has an attacker with positive strength, then its own strength cannot be maximal.

Principle 2.15 (Weakening)**.** *A gradual semantics* $\mathbf{S}$ *satisfies Weakening iff for any AG* $\mathbf{G} = \langle \mathcal{A}, \mathcal{R} \rangle$*,* $\forall a \in \mathcal{A}$*, it holds that if* $\mathtt{Att}_{\mathbf{G}}(a) \neq \emptyset$*, then* $\mathtt{Deg}^{\mathbf{S}}_{\mathbf{G}}(a) < 1$*.*

Resilience states that in any graph, every argument will have strictly positive strength.

Principle 2.16 (Resilience)**.** $\mathbf{S}$ *satisfies Resilience iff for any AG* $\mathbf{G} = \langle \mathcal{A}, \mathcal{R} \rangle$*,* $\forall a \in \mathcal{A}$*,* $\mathtt{Deg}^{\mathbf{S}}_{\mathbf{G}}(a) > 0$*.*

The *Reinforcement* principle states that the intensity of an attack is proportionate to the strength of its source.

Principle 2.17 (Reinforcement)**.** *A gradual semantics* $\mathbf{S}$ *satisfies Reinforcement iff for any AG* $\mathbf{G} = \langle \mathcal{A}, \mathcal{R} \rangle$*,* $\forall a, b \in \mathcal{A}$*, the following holds: if* $\mathtt{Deg}^{\mathbf{S}}_{\mathbf{G}}(a) > 0$*,* $\mathtt{Att}_{\mathbf{G}}(a) \setminus \mathtt{Att}_{\mathbf{G}}(b) = \{x\}$*,* $\mathtt{Att}_{\mathbf{G}}(b) \setminus \mathtt{Att}_{\mathbf{G}}(a) = \{y\}$*, and* $\mathtt{Deg}^{\mathbf{S}}_{\mathbf{G}}(y) > \mathtt{Deg}^{\mathbf{S}}_{\mathbf{G}}(x)$*, then* $\mathtt{Deg}^{\mathbf{S}}_{\mathbf{G}}(a) > \mathtt{Deg}^{\mathbf{S}}_{\mathbf{G}}(b)$*.*

Where Weakening leads to a loss of strength whenever an argument is attacked by at least one argument with positive strength, *Counting* requires that each alive attacker negatively affects the overall strength of an argument.

Principle 2.18 (Counting)**.** *A gradual semantics* $\mathbf{S}$ *satisfies Counting iff for any AG* $\mathbf{G} = \langle \mathcal{A}, \mathcal{R} \rangle$*,* $\forall a, b \in \mathcal{A}$*, it holds that: if* $\mathtt{Att}_{\mathbf{G}}(b) = \mathtt{Att}_{\mathbf{G}}(a) \cup \{x\}$ *with* $x \notin \mathtt{Att}_{\mathbf{G}}(a)$*,* $\mathtt{Deg}^{\mathbf{S}}_{\mathbf{G}}(x) > 0$*, and* $\mathtt{Deg}^{\mathbf{S}}_{\mathbf{G}}(a) > 0$*, then* $\mathtt{Deg}^{\mathbf{S}}_{\mathbf{G}}(a) > \mathtt{Deg}^{\mathbf{S}}_{\mathbf{G}}(b)$*.*

The *Weakening Soundness* principle, going beyond Weakening, states that attacks are the only reason for a decrease in argument strength.

Principle 2.19 (Weakening Soundness)**.** *A gradual semantics* $\mathbf{S}$ *satisfies Weakening Soundness iff for any AG* $\mathbf{G} = \langle \mathcal{A}, \mathcal{R} \rangle$*,* $\forall a \in \mathcal{A}$*, it holds that: if* $\mathtt{Deg}^{\mathbf{S}}_{\mathbf{G}}(a) < 1$*, then* $\exists b \in \mathtt{Att}_{\mathbf{G}}(a)$ *s.t.* $\mathtt{Deg}^{\mathbf{S}}_{\mathbf{G}}(b) > 0$*.*

2.3. **Probabilistic Argumentation Frameworks.** When using argumentation graphs and semantics to evaluate arguments, we assume that all relevant arguments and attacks are considered and no arguments or attacks in the graph are irrelevant. In reality, however, it is rarely clear which arguments and attacks apply and should thus be placed in the graph. We might, for example, run into natural language being imprecise such that it is unclear whether claims are contradictory or there might be multiple ways to model the premises and claims often left implicit in conversation. We might also be in a situation where we are unsure what arguments a dialogue partner or an audience has in their mind. At other times we may be dealing with explicit uncertainty, such as with utterances like 'I am 99% sure' or 'I may have seen Robin yesterday'.

Example 3. [28] Suppose we are part of the audience to a discussion between a proponent and an opponent of expanding Heathrow airport. In the debate the proponent offers up the first and third of the following arguments and the opponent presents the second:

- x = We should build a third runway at Heathrow because everyone will benefit from the increased capacity.
- y = It is not true that everyone will benefit in the community.
- z = Local residents won't have problems with traffic because we will increase public transport to the airport.

The first and third arguments make both their premises and claim explicit. The second argument, however, leaves its premises implicit; it is an enthymeme.

Now suppose that with the knowledge we have available, we deem that the second argument can be made explicit in one of two ways:

- y' = It is not true that everyone will benefit in the community. There are local residents who will suffer from increased noise from the increased number of aircraft.
- y'' = It is not true that everyone will benefit in the community. There are local residents who will have problems from increased traffic on the roads to the airport.

We now see that both interpretations of y attack x, but only the second is attacked by z. It is thus uncertain whether we should include the attack from z to y in our graph.

Uncertainty such as that described above is captured by probabilistic argumentation frameworks: an extension of argumentation frameworks, originally proposed by Li, Oren, and Norman [33], where each argument and attack is given a certain likelihood of appearing in a graph.

Definition 2.20 (PrAF)**.** *A probabilistic argumentation framework (PrAF) is a quadruple* $\mathbf{F} = \langle \mathcal{A}, P_{\mathcal{A}}, \mathcal{R}, P_{\mathcal{R}} \rangle$*, where* $\langle \mathcal{A}, \mathcal{R} \rangle$ *is an argumentation graph and* $P_{\mathcal{A}} : \mathcal{A} \to (0, 1]$ *and* $P_{\mathcal{R}} : \mathcal{R} \to (0, 1]$ *associate* *likelihood values with arguments and attacks respectively.* `PrAF` *denotes the set of all probabilistic argumentation frameworks.*

In a PrAF $\mathbf{F} = \langle \mathcal{A}, P_{\mathcal{A}}, \mathcal{R}, P_{\mathcal{R}} \rangle$, the argumentation graph $\mathbf{G} = \langle \mathcal{A}, \mathcal{R} \rangle$ represents the set of all arguments and attacks that may potentially appear. An instantiated graph that may arise under the uncertainty we face thus contains a subset of the arguments and attacks in $\mathbf{G}$. We call the process of deriving such a graph from a PrAF induction, and name the graphs that result induced graphs (of the PrAF). The functions $P_{\mathcal{A}}$ and $P_{\mathcal{R}}$ represent the uncertainty in the arguments and attacks in $\mathbf{G}$. $P_{\mathcal{A}}$ gives the probability that an argument appears in a graph induced from $\mathbf{G}$ and $P_{\mathcal{R}}$ the conditional probability that an attack appears in an induced graph given that both arguments it relates appear in the graph. The ranges of both functions deliberately exclude 0, as any argument or attack with a zero probability is known never to appear and is thus redundant. The maximum value 1 either function may assign represents certainty that an argument appears in an induced graph or that an attack appears given that its origin and target do as well.

We say that an argument is *perfect* in a PrAF, if it is both non-attacked and its probability is 1.

Definition 2.21 (Induced Graph)**.** *An argumentation graph* $\mathbf{G}' = \langle \mathcal{A}', \mathcal{R}' \rangle$ *is induced from a probabilistic argumentation framework* $\mathbf{F} = \langle \mathcal{A}, P_{\mathcal{A}}, \mathcal{R}, P_{\mathcal{R}} \rangle$ *iff all of the following hold:*

- $\mathcal{A}' \subseteq \mathcal{A}$
- $\mathcal{R}' \subseteq \mathcal{R} \cap (\mathcal{A}' \times \mathcal{A}')$
- $\forall a \in \mathcal{A}$ *such that* $P_{\mathcal{A}}(a) = 1$*,* $a \in \mathcal{A}'$
- $\forall (a, b) \in \mathcal{R}$ *such that* $P_{\mathcal{R}}((a, b)) = 1$ *and* $a, b \in \mathcal{A}'$*,* $(a, b) \in \mathcal{R}'$

$\mathtt{I}(\mathbf{F})$ *denotes the set of all argumentation graphs that may be induced from a probabilistic argumentation framework* $\mathbf{F}$*.*

Example 4 (Running example, continued)**.** Looking back to the previous example, we see that the argumentation graph we have so far been using for our running example may represent the dialogue under interpretation y''. Argument a in our graph corresponds to argument z in the dialogue, b corresponds to y and c to z. The other interpretation would produce a subgraph of our example graph that excludes the attack from c to b. This whole scenario may then be modelled with the PrAF presented in Figure 2. Here we would have $0 < P_{\mathcal{R}}((a, b)) < 1$. We would have to base the exact value of this probability on our understanding of the likelihood of either interpretation of y being intended. All other probabilities are equal to one.

Our definition of an induced graph differs from the original, proposed in [33], in the fourth bullet; in addition to the attack (a, b) being certain, the original definition requires that arguments a and b both have probability 1 instead of only requiring them to be present in

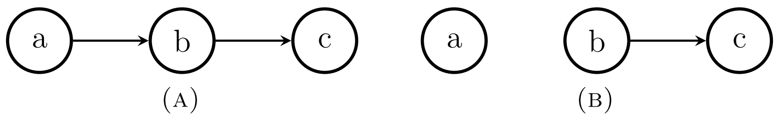

FIGURE 2. A probabilistic argumentation framework with a single uncertain attack (a, b): (a) shows the entire PrAF and its induced graph where (a, b) is present; (b) shows the induced graph where (a, b) is not present.

the graph for (a, b) to be certainly present in the graph. This change eliminates any induced graphs that will receive probability 0 under the following definition 2.22, where an attack with probability 1 is not present even though the arguments it connects are.

Since independence between arguments and attacks is assumed, the probability of some induced graph $\mathbf{G}$ being induced from a PrAF $\mathbf{F}$ can be computed using the joint probabilities of independent variables:

Definition 2.22 (Probability of an Induced Graph)**.** *Given a probabilistic argumentation framework* $\mathbf{F} = \langle \mathcal{A}, P_{\mathcal{A}}, \mathcal{R}, P_{\mathcal{R}} \rangle$*, the probability of some graph* $\mathbf{G}' = \langle \mathcal{A}', \mathcal{R}' \rangle$ *being induced from* $\mathbf{F}$ *is:*

$$P^I_{\mathbf{F}}(\mathbf{G}') = \prod_{a \in \mathcal{A}'} P_{\mathcal{A}}(a) \prod_{a \in \mathcal{A} \setminus \mathcal{A}'} (1 - P_{\mathcal{A}}(a)) \prod_{r \in \mathcal{R}'} P_{\mathcal{R}}(r) \prod_{r \in \mathcal{R}\downarrow_{\mathcal{A}'} \setminus \mathcal{R}'} (1 - P_{\mathcal{R}}(r))$$

where $\mathcal{R} \downarrow_{\mathcal{A}'} = \{(a, b) | a, b \in \mathcal{A}' \text{ and } (a, b) \in \mathcal{R}\}$.

Through definitions 2.20, 2.21 and 2.22 it is easily shown that

$$\forall \mathbf{F} \in \texttt{PrAF},\ \forall \mathbf{G} \in \texttt{I}(\mathbf{F}),\ P^I_{\mathbf{F}}(\mathbf{G}) > 0 \tag{1}$$

The following proposition claims that $P^I_{\mathbf{F}}$ is a probability distribution over the induced graphs of a PrAF.

Proposition 2.23. *The sum of probabilities of all argumentation graphs that may be induced from an arbitrary PrAF* $\mathbf{F}$ *is* 1*. i.e.*

$$\sum_{\mathbf{G} \in \texttt{I}(\mathbf{F})} P^I_{\mathbf{F}}(\mathbf{G}) = 1$$

Proof. Note that this proposition is essentially the same as the analogous statement for PrAFs by [33]. Since we slightly modified the definition of an induced graph, we present this statement for completeness without claiming that the result is original. We prove this proposition by induction. For the base case, consider a PrAF $\mathbf{F} = \langle \mathcal{A}, P_{\mathcal{A}}, \mathcal{R}, P_{\mathcal{R}} \rangle$ with $\mathcal{A} = \{a\}$. Either $P_{\mathcal{A}}(a) = 1$ or $0 < P_{\mathcal{A}}(a) < 1$. In the former case, $\texttt{I}(\mathbf{F}) = \{\langle \mathcal{A}, \mathcal{R} \rangle\}$ with $P^I_{\mathbf{F}}(\langle \mathcal{A}, \mathcal{R} \rangle) = 1$. In the latter case, $\texttt{I}(\mathbf{F}) = \{\langle \mathcal{A}, \mathcal{R} \rangle, \langle \emptyset, \mathcal{R} \rangle\}$ with $P^I_{\mathbf{F}}(\langle \mathcal{A}, \mathcal{R} \rangle) = P_{\mathcal{A}}(a)$ and

$P^I_{\mathbf{F}}(\langle\emptyset, \mathcal{R}\rangle) = 1 - P_{\mathcal{A}}(a)$. Since $0 < P_{\mathcal{A}}(a) < 1$, $P_{\mathcal{A}}(a) + (1 - P_{\mathcal{A}}(a)) = 1$. In either case, $\sum_{\mathbf{G} \in \mathtt{I}(\mathbf{F})} P^I_{\mathbf{F}}(\mathbf{G}) = 1$.

In general, there are four ways to expand a PrAF $\mathbf{F} = \langle \mathcal{A}, P_{\mathcal{A}}, \mathcal{R}, P_{\mathcal{R}} \rangle$ to make PrAF $\mathbf{F}' = \langle \mathcal{A}', P'_{\mathcal{A}}, \mathcal{R}', P'_{\mathcal{R}} \rangle$: we may add a certain argument, we may add a certain attack, we may add an uncertain argument, or we may add an uncertain attack. We consider these as four induction steps. For all, assume that for $\mathbf{F}$ it holds that $\sum_{\mathbf{G} \in \mathtt{I}(\mathbf{F})} P^I_{\mathbf{F}}(\mathbf{G}) = 1$.

(1) If we add a certain argument a, we have a bijective function f from $\mathtt{I}(\mathbf{F})$ to $\mathtt{I}(\mathbf{F}')$ s.t. $\forall \mathbf{G} \in \mathtt{I}(\mathbf{F})$, $P^I_{\mathbf{F}}(\mathbf{G}) = P^I_{\mathbf{F}'}(f(\mathbf{G}))$, so $\sum_{\mathbf{G} \in \mathtt{I}(\mathbf{F})} P^I_{\mathbf{F}}(\mathbf{G}) = \sum_{\mathbf{G}' \in \mathtt{I}(\mathbf{F}')} P^I_{\mathbf{F}'}(\mathbf{G}') = 1$. Here $f(\mathbf{G})$ is $\mathbf{G}$ with the addition of the new argument a.
(2) If we add a certain attack $r = (a, b)$, we also have a bijective function f from $\mathtt{I}(\mathbf{F})$ to $\mathtt{I}(\mathbf{F}')$ s.t. $\forall \mathbf{G} \in \mathtt{I}(\mathbf{F})$, $P^I_{\mathbf{F}}(\mathbf{G}) = P^I_{\mathbf{F}'}(f(\mathbf{G}))$, so $\sum_{\mathbf{G} \in \mathtt{I}(\mathbf{F})} P^I_{\mathbf{F}}(\mathbf{G}) = \sum_{\mathbf{G}' \in \mathtt{I}(\mathbf{F}')} P^I_{\mathbf{F}'}(\mathbf{G}') = 1$. Here $f(\mathbf{G})$ is $\mathbf{G}$ with the addition of r if a and b are in $\mathbf{G}$ and simply $\mathbf{G}$ otherwise.
(3) If we add an argument a with $0 < P_{\mathcal{A}}(a) < 1$, we have a surjection f from $\mathtt{I}(\mathbf{F}')$ to $\mathtt{I}(\mathbf{F})$ s.t. for each $\mathbf{G} \in \mathtt{I}(\mathbf{F})$ there are exactly two graphs $\mathbf{G}', \mathbf{G}'' \in \mathtt{I}(\mathbf{F}')$ with $f(\mathbf{G}') = f(\mathbf{G}'') = \mathbf{G}$ where $\mathbf{G}' = \mathbf{G}$ and $\mathbf{G}''$ is $\mathbf{G}$ with the addition of a. We have $P^I_{\mathbf{F}'}(\mathbf{G}') = P^I_{\mathbf{F}}(\mathbf{G}) \cdot (1 - P_{\mathcal{A}}(a))$ and $P^I_{\mathbf{F}'}(\mathbf{G}'') = P^I_{\mathbf{F}}(\mathbf{G}) \cdot P_{\mathcal{A}}(a)$. Since $0 < P_{\mathcal{A}}(a) < 1$, $P^I_{\mathbf{F}'}(\mathbf{G}') + P^I_{\mathbf{F}'}(\mathbf{G}'') = P^I_{\mathbf{F}}(\mathbf{G})$, so $\sum_{\mathbf{G} \in \mathtt{I}(\mathbf{F})} P^I_{\mathbf{F}}(\mathbf{G}) = \sum_{\mathbf{G}' \in \mathtt{I}(\mathbf{F}')} P^I_{\mathbf{F}'}(\mathbf{G}')$.
(4) If we add an attack $r = (a, b)$ with $0 < P_{\mathcal{R}}(r) < 1$, we get that each $\mathbf{G} \in \mathtt{I}(\mathbf{F})$ that does not contain arguments a and b corresponds to a single identical $\mathbf{G}' \in \mathtt{I}(\mathbf{F}')$ with $P^I_{\mathbf{F}}(\mathbf{G}) = P^I_{\mathbf{F}'}(\mathbf{G}')$ and that each $\mathbf{G} \in \mathtt{I}(\mathbf{F})$ that does contain arguments a and b corresponds to two graphs $\mathbf{G}', \mathbf{G}'' \in \mathtt{I}(\mathbf{F}')$ s.t. $\mathbf{G}' = \mathbf{G}$, $\mathbf{G}''$ is $\mathbf{G}$ with the addition of r, $P^I_{\mathbf{F}'}(\mathbf{G}') = P^I_{\mathbf{F}}(\mathbf{G}) \cdot (1 - P_{\mathcal{R}}(r))$ and $P^I_{\mathbf{F}'}(\mathbf{G}'') = P^I_{\mathbf{F}}(\mathbf{G}) \cdot P_{\mathcal{R}}(r)$. Since $0 < P_{\mathcal{R}}(r) < 1$, $P^I_{\mathbf{F}'}(\mathbf{G}') + P^I_{\mathbf{F}'}(\mathbf{G}'') = P^I_{\mathbf{F}}(\mathbf{G})$, so $\sum_{\mathbf{G} \in \mathtt{I}(\mathbf{F})} P^I_{\mathbf{F}}(\mathbf{G}) = \sum_{\mathbf{G}' \in \mathtt{I}(\mathbf{F}')} P^I_{\mathbf{F}'}(\mathbf{G}')$.

□

The probability distribution $P^I_{\mathbf{F}}$ is used in [33] to define the probability of a set of arguments $\mathcal{X}$ being accepted in a PrAF as the sum of probabilities of those induced graphs in which all arguments are accepted.

Definition 2.24 (Probability of Extension-based Acceptability)**.** *Given a probabilistic argumentation framework* $\mathbf{F} = \langle \mathcal{A}, P_{\mathcal{A}}, \mathcal{R}, P_{\mathcal{R}} \rangle$*, an inference mode* $i \in \{$*sceptical*, *credulous*$\}$*, and an extension semantics* $\mathbf{S}$*, the probability that some set of arguments* $\mathcal{X} \subseteq \mathcal{A}$ *is acceptable in* $\mathbf{F}$ *w.r.t.* $\mathbf{S}$ *and* i *is:*

$$P^{\mathbf{S},i}_{\mathbf{F}}(\mathcal{X}) = \sum_{\mathbf{G} \in \mathtt{I}(\mathbf{F}), \forall a \in \mathcal{X} \mathbf{G} \mid\!\sim^{i}_{\mathbf{S}} a} P^I_{\mathbf{F}}(\mathbf{G})$$

Where $\mathbf{G} \vdash^{sceptical}_{\mathbf{S}} a$ *iff* $\forall \mathcal{S} \in \mathtt{E}_{\mathbf{S}}(\mathbf{G})$, $a \in \mathcal{S}$ *and* $\mathbf{G} \vdash^{credulous}_{\mathbf{S}} a$ *iff* $\exists \mathcal{S} \in \mathtt{E}_{\mathbf{S}}(\mathbf{G})$, *s.t.* $a \in \mathcal{S}$.

What follows is a collection of special notations used in the thesis.

Notation. *Let* $\mathbf{F} = \langle \mathcal{A}, P_{\mathcal{A}}, \mathcal{R}, P_{\mathcal{R}} \rangle$ *be a PrAF,* $\mathbf{G} = \langle \mathcal{A}, \mathcal{R} \rangle$ *be an AG, and* $a \in \mathcal{A}$.

- *We write* $a\mathcal{R}b$ *iff* $(a, b) \in \mathcal{R}$.
- $\mathtt{Att}_{\mathbf{F}}(a)$ *and* $\mathtt{Att}_{\mathbf{G}}(a)$ *denote the set of all attackers of* a *in* $\mathbf{F}$ *and* $\mathbf{G}$ *respectively (i.e.,* $\mathtt{Att}_{\mathbf{F}}(a) = \mathtt{Att}_{\mathbf{G}}(a) = \{b \in \mathcal{A} | b\mathcal{R}a$*).*
- *For* $a, b \in \mathcal{A}$*, we say there is a path from* b *to* a *if there exists a finite non-empty sequence* $\langle x_1, \ldots, x_n \rangle$ *of arguments* $x_i \in \mathcal{A}$ *s.t.* $x_1 = b, x_n = a$ *and* $\forall i < n, x_i \mathcal{R} x_{i+1}$.
- *For any* $\mathbf{F} = \langle \mathcal{A}, P_{\mathcal{A}}, \mathcal{R}, P_{\mathcal{R}} \rangle, \mathbf{F}' = \langle \mathcal{A}', P'_{\mathcal{A}}, \mathcal{R}', P'_{\mathcal{R}} \rangle \in \mathtt{PrAF}$ *s.t.* $\mathcal{A} \cap \mathcal{A}' = \emptyset$, $\mathbf{F} \oplus \mathbf{F}'$ *is the PrAF* $\langle \mathcal{A} \cup \mathcal{A}', P''_{\mathcal{A}}, \mathcal{R} \cup \mathcal{R}', P''_{\mathcal{R}} \rangle$ *where for any* $a \in \mathcal{A}$ *(respectively* $a \in \mathcal{A}'$*)* $P''_{\mathcal{A}}(a) = P_{\mathcal{A}}(a)$ *(respectively* $P''_{\mathcal{A}}(a) = P'_{\mathcal{A}}(a)$*) and for any* $r \in \mathcal{R}$ *(respectively* $r \in \mathcal{R}'$*)* $P''_{\mathcal{R}}(r) = P_{\mathcal{R}}(r)$ *(respectively* $P''_{\mathcal{R}}(r) = P'_{\mathcal{R}}(r)$*).*
- *For any* $\mathbf{G} = \langle \mathcal{A}, \mathcal{R} \rangle, \mathbf{G}' = \langle \mathcal{A}', \mathcal{R}' \rangle \in \mathtt{AG}$ *s.t.* $\mathcal{A} \cap \mathcal{A}' = \emptyset$, $\mathbf{G} \oplus \mathbf{G}'$ *is the AG* $\langle \mathcal{A} \cup \mathcal{A}', \mathcal{R} \cup \mathcal{R}' \rangle$.
- *For any* $\mathbf{F} = \langle \mathcal{A}, P_{\mathcal{A}}, \mathcal{R}, P_{\mathcal{R}} \rangle$ *and* $S \subseteq \mathcal{A}$, $\mathbf{F}_{|S} = \langle S, (P_{\mathcal{A}})_{|S}, \mathcal{R}_{|S \times S}, (P_{\mathcal{R}})_{|(\mathcal{R}_{|S \times S})} \rangle$.

3. Probabilistic Acceptability of Arguments

In the work of Li, Oren, and Norman [33], Definition 2.22 is used to determine the probability that some set of arguments is acceptable under a given Dung semantics by adding together the probabilities of those induced graphs where the set is acceptable. This section explores the possibility of determining that probability under a gradual semantics instead.

Where Dung semantics determine the acceptability of an argument directly, without explicitly considering the overall strength of the argument, gradual semantics may be used to determine acceptability indirectly through the strengths they assign arguments. An approach to deriving argument acceptability from the strengths assigned to arguments by a gradual semantics is proposed in [1]; we may simply accept those arguments whose strength meets or exceeds a threshold we choose.

Definition 3.1 (Acceptability Under Gradual Semantics)**.** *Given an argumentation graph* $\mathbf{G} = \langle \mathcal{A}, \mathcal{R} \rangle$ *and a gradual semantics* $\mathbf{S}$*, an argument* $a \in \mathcal{A}$ *is threshold accepted in* $\mathbf{G}$ *with respect to* $\mathbf{S}$ *and some threshold* $t \in [0, 1]$*, denoted* $\mathbf{G} \mathrel{\mid\!\sim^{t}_{\mathbf{S}}} a$*, iff* $\mathtt{Deg}^{\mathbf{S}}_{\mathbf{G}}(a) \geq t$*.*

Now that we can derive the acceptability of an argument from its strength, we may follow the same approach as Li, Oren, and Norman to determine the probability that a set of arguments is acceptable.

Definition 3.2 (Probability of Acceptability)**.** *Given a probabilistic argumentation framework* $\mathbf{F} = \langle \mathcal{A}, P_{\mathcal{A}}, \mathcal{R}, P_{\mathcal{R}} \rangle$*, a gradual semantics* $\mathbf{S}$*, and a threshold* $t \in [0, 1]$*, the probability that some set of arguments* $\mathcal{X} \subseteq \mathcal{A}$ *is acceptable in* $\mathbf{F}$ *w.r.t.* $\mathbf{S}$ *and* t *is:*

$$P^{\mathbf{S},t}_{\mathbf{F}}(\mathcal{X}) = \sum_{\mathbf{G} \in \mathbf{I}(\mathbf{F}), \forall a \in \mathcal{X} \mathbf{G} \mid\!\sim^{t}_{\mathbf{S}} a} P^{I}_{\mathbf{F}}(\mathbf{G})$$

That is, $P^{\mathbf{S},t}_{\mathbf{F}}(\mathcal{X})$ *is the sum of the probabilities of the induced graphs of* $\mathbf{F}$ *where all arguments in* $\mathcal{X}$ *are accepted. For brevity, we write* $P^{\mathbf{S},t}_{\mathbf{F}}(x)$ *instead of* $P^{\mathbf{S},t}_{\mathbf{F}}(\mathcal{X})$ *for singleton sets* $\mathcal{X} = \{x\}$*.*

Example 5 (Running example, continued)**.** In our running example, we are trying to answer the question whether Heathrow airport should be expanded. For this, we may wish to find the probability of argument c being accepted. Suppose we pick an arbitrary acceptance threshold $t = 2/3$. In induced graph $\mathbf{G}_1$, shown in Figure 2a, we have $\mathtt{Deg}^{\mathtt{Hbs}}_{\mathbf{G}_1}(c) = 2/3$. In this induced graph c is accepted. In induced graph $\mathbf{G}_2$, shown in Figure 2b, we have $\mathtt{Deg}^{\mathtt{Hbs}}_{\mathbf{G}_2}(c) = 1/2$. Here c is not accepted. The probability of either induced graph, and by extension the probability of acceptability, depends on $P_{\mathcal{R}}((b, c))$. We see that $P^{\mathtt{Hbs},t}_{\mathbf{F}}(c) = P^{I}_{\mathbf{F}}(\mathbf{G}_1) = P_{\mathcal{R}}((b, c))$.

The following result states that the probability of acceptance of any argument in a framework is bounded from above by its probability of being present in the graph.

Proposition 3.3. *Let* $\mathbf{F} = \langle \mathcal{A}, P_{\mathcal{A}}, \mathcal{R}, P_{\mathcal{R}} \rangle$ *be a probabilistic argumentation framework and let* $\mathbf{S}$ *be a gradual semantics. For every* $t \in [0,1]$ *we have* $P_{\mathbf{F}}^{\mathbf{S},t}(a) \leq P_{\mathcal{A}}(a)$.

Proof. Let $\mathbf{F} = \langle \mathcal{A}, P_{\mathcal{A}}, \mathcal{R}, P_{\mathcal{R}} \rangle$ be a PrAF with $a \in \mathcal{A}$, let $\mathbf{S}$ be a gradual semantics, and let $t \in [0,1]$. It follows from Definition 2.22 that

$$P_{\mathcal{A}}(a) = \sum_{\mathbf{G}=\langle \mathcal{A},\mathcal{R}\rangle \in \mathtt{I}(\mathbf{F}), a \in \mathcal{A}} P_{\mathbf{F}}^{I}(\mathbf{G}) \tag{2}$$

It follows from Definition 3.1 that any for any $\mathbf{G} = \langle \mathcal{A}, \mathcal{R} \rangle \in \mathtt{I}(\mathbf{F})$ where $a \notin \mathcal{A}$, for any threshold t, a is not accepted in $\mathbf{G}$ w.r.t $\mathbf{S}$. From here we see that if for any $\mathbf{G} = \langle \mathcal{A}, \mathcal{R} \rangle \in \mathcal{A}$, $\mathbf{G} \mathrel{\vdash\!\!\sim}^{t}_{\mathbf{S}} a$, then $P_{\mathbf{F}}^{\mathbf{S},t}(a) = P_{\mathcal{A}}(a)$ and if not, then $P_{\mathbf{F}}^{\mathbf{S},t}(a) < P_{\mathcal{A}}(a)$. □

Now we state a form of monotonicity property that compares probabilities of acceptance when different thresholds are considered.

Proposition 3.4. *Let* $\mathbf{F} = \langle \mathcal{A}, P_{\mathcal{A}}, \mathcal{R}, P_{\mathcal{R}} \rangle$ *be a probabilistic argumentation framework and let* $\mathbf{S}$ *be a gradual semantics. If* $t \leq t'$*, then* $P_{\mathbf{F}}^{\mathbf{S},t}(a) \geq P_{\mathbf{F}}^{\mathbf{S},t'}(a)$.

Proof. We prove the result by contraposition. Let $\mathbf{F} = \langle \mathcal{A}, P_{\mathcal{A}}, \mathcal{R}, P_{\mathcal{R}} \rangle$ be a PrAF with $a \in \mathcal{A}$, let $\mathbf{S}$ be a gradual semantics, and let $t, t' \in [0,1]$. Suppose $P_{\mathbf{F}}^{\mathbf{S},t}(a) < P_{\mathbf{F}}^{\mathbf{S},t'}(a)$. This implies the existence of a $\mathbf{G} \in \mathtt{I}(\mathbf{F})$ s.t. $\mathbf{G} \mathrel{\vdash\!\!\sim}^{t'}_{\mathbf{S}} (a)$ and not $\mathbf{G} \mathrel{\vdash\!\!\sim}^{t}_{\mathbf{S}} (a)$. From Definition 3.1 we see that such a $\mathbf{G}$ may only exist if $t > t'$. □

Next we consider properties that depend on behaviour of the chosen gradual semantics. According to the first property, if the semantics satisfies Maximality, then the probability that a non-attacked argument has maximal strength is maximised.

Proposition 3.5. *Let* $\mathbf{F} = \langle \mathcal{A}, P_{\mathcal{A}}, \mathcal{R}, P_{\mathcal{R}} \rangle$ *be a probabilistic argumentation framework and let* $\mathbf{S}$ *be a gradual semantics that satisfies Maximality. If* $a \in \mathcal{A}$ *is not attacked in* $\mathbf{F}$*, then* $P_{\mathbf{F}}^{\mathbf{S},1}(a) = P_{\mathcal{A}}(a)$.

Proof. Let $\mathbf{F} = \langle \mathcal{A}, P_{\mathcal{A}}, \mathcal{R}, P_{\mathcal{R}} \rangle$ be a PrAF with $a \in \mathcal{A}$, and let $\mathbf{S}$ be a gradual semantics that satisfies Maximality. Suppose a is not attacked in $\mathbf{F}$. This means that a is not attacked in any $\mathbf{G} \in \mathtt{I}(\mathbf{F})$. Combined with the assumption that $\mathbf{S}$ satisfies Maximality, this gives $\forall \mathbf{G} = \langle \mathcal{A}, \mathcal{R} \rangle \in \mathtt{I}(\mathbf{F})$ s.t. $a \in \mathcal{A}$, $\mathtt{Deg}_{\mathbf{G}}^{\mathbf{S}}(a) = 1$. From here, Definition 3.1 gives that $\forall \mathbf{G} = \langle \mathcal{A}, \mathcal{R} \rangle \in \mathtt{I}(\mathbf{F})$ s.t. $a \in \mathcal{A}$, $\mathbf{G} \mathrel{\vdash\!\!\sim}^{1}_{\mathbf{S}} a$. Equation 2 now gives us $P_{\mathbf{F}}^{\mathbf{S},1}(a) = P_{\mathcal{A}}(a)$. □

Satisfaction of the Resilience principle implies non-zero probability that an argument has at least some positive strength.

Proposition 3.6. *Let $\mathbf{F} = \langle \mathcal{A}, P_{\mathcal{A}}, \mathcal{R}, P_{\mathcal{R}} \rangle$ be a probabilistic argumentation framework and let $\mathbf{S}$ be a gradual semantics that satisfies Resilience. Then for some $t > 0$, $P^{\mathbf{S},t}_{\mathbf{F}}(a) > 0$.*

Proof. Let $\mathbf{F} = \langle \mathcal{A}, P_{\mathcal{A}}, \mathcal{R}, P_{\mathcal{R}} \rangle$ be a PrAF with $a \in \mathcal{A}$, and let $\mathbf{S}$ be a gradual semantics that satisfies Resilience. Assume it does not hold that for some $t > 0$, $P^{\mathbf{S},t}_{\mathbf{F}}(a) > 0$. Through Equation 1 and Definition 3.2 we see this means that for any $t > 0$ there is no $\mathbf{G} \in \mathtt{I}(\mathbf{F})$ s.t. $\mathbf{G} \mathrel{\substack{t \\ \mid\!\sim_{\mathbf{S}}}} a$. Definition 3.1 then gives, together with $t > 0$, that in all $\mathbf{G} \in \mathtt{I}(\mathbf{F})$ $\mathtt{Deg}^{\mathbf{S}}_{\mathbf{G}}(a) = 0$. $\mathbf{S}$ satisfies resilience, so this is not possible. □

If, in addition, the semantics satisfies Weakening, the probability of acceptance of an attacked argument cannot reach its probability of belonging to the graph.

Proposition 3.7. *Let the semantics satisfy Resilience and Weakening. If a is attacked, then for $t = 1$, $P^{\mathbf{S},t}_{\mathbf{F}}(a) < P_{\mathcal{A}}(a)$.*

Proof. Let $\mathbf{F} = \langle \mathcal{A}, P_{\mathcal{A}}, \mathcal{R}, P_{\mathcal{R}} \rangle$ be a PrAF with $a \in \mathcal{A}$ s.t. $\mathtt{Att}_{\mathbf{F}}(a) \neq \emptyset$, let $\mathbf{S}$ be a gradual semantics that satisfies Resilience and Weakening, and let $t = 1$. $\mathtt{Att}_{\mathbf{F}}(a) \neq \emptyset$ implies that there is a $\mathbf{G} = \langle \mathcal{A}, \mathcal{R} \rangle \in \mathtt{I}(\mathbf{F})$ s.t. $a \in \mathcal{A}$ and $\mathtt{Att}_{\mathbf{G}}(a) \neq \emptyset$. $\mathbf{S}$ satisfying Resilience makes it so that $\forall b \in \mathtt{Att}_{\mathbf{G}}(a)$, $\mathtt{Deg}^{\mathbf{S}}_{\mathbf{G}}(b) > 0$. $\mathbf{S}$ satisfying Weakening makes it so that a being attacked in $\mathbf{G}$ and all attackers of a in $\mathbf{G}$ having a non-zero strength imply $\mathtt{Deg}^{\mathbf{S}}_{\mathbf{G}}(a) < 1$. From here, Definition 3.1 gives that not $\mathbf{G} \mathrel{\substack{t \\ \mid\!\sim_{\mathbf{S}}}} a$. From Equation 2, Definition 3.1, and Definition 3.2 we see that $P^{\mathbf{S},t}_{\mathbf{F}}(a) = P_{\mathcal{A}}(a)$ if a is accepted in all $\mathbf{G} \in \mathtt{I}(\mathbf{F})$ and $P^{\mathbf{S},t}_{\mathbf{F}}(a) < P_{\mathcal{A}}(a)$ otherwise. Since we have a $\mathbf{G} \in \mathtt{I}(\mathbf{F})$ s.t. a is not accepted, we thus have $P^{\mathbf{S},t}_{\mathbf{F}}(a) < P_{\mathcal{A}}(a)$. □

The last result of this section characterises perfect arguments.

Proposition 3.8. *Let $\mathbf{S}$ satisfy Maximality, Weakening and Resilience. Then a is perfect in $\mathbf{F}$ iff $P^{\mathbf{S},1}_{\mathbf{F}}(a) = 1$.*

Proof. Let $\mathbf{F} = \langle \mathcal{A}, P_{\mathcal{A}}, \mathcal{R}, P_{\mathcal{R}} \rangle$ be a PrAF with $a \in \mathcal{A}$, and let $\mathbf{S}$ be a gradual semantics that satisfies Maximality, Resilience and Weakening. We prove that a is perfect in $\mathbf{F}$ iff $P^{\mathbf{S},1}_{\mathbf{F}} = 1$. i.e. we prove that $\mathtt{Att}_{\mathbf{F}}(a) = \emptyset \wedge P_{\mathcal{A}}(a) = 1 \iff P^{\mathbf{S},1}_{\mathbf{F}} = 1$. Left-to-right we have: $P_{\mathcal{A}}(a) = 1$ implies that $\forall \mathbf{G} = \langle \mathcal{A}, \mathcal{R} \rangle \in \mathtt{I}(\mathbf{F})$, $a \in \mathcal{A}$. Combining this with $\mathbf{S}$ satisfying Maximality and $\mathtt{Att}_{\mathbf{F}}(a) = \emptyset$ we get $\forall \mathbf{G} \in \mathtt{I}(\mathbf{F})$, $\mathtt{Deg}^{\mathbf{S}}_{\mathbf{G}}(a) = 1$. Definition 3.1 gives us $\forall \mathbf{G} \in \mathtt{I}(\mathbf{F})$, $\mathbf{G} \mathrel{\substack{1 \\ \mid\!\sim_{\mathbf{S}}}} (a)$, from which we derive $P^{\mathbf{S},1}_{\mathbf{F}} = 1$ through Proposition 2.23 and Definition 3.2. Right-to-left we prove by contraposition: We assume $\mathtt{Att}_{\mathbf{F}}(a) \neq \emptyset \vee P_{\mathcal{A}}(a) \neq 1$. If $\mathtt{Att}_{\mathbf{F}}(a) \neq \emptyset$, Proposition 3.7 gives us that, because $\mathbf{S}$ satisfies Resilience and Weakening, $P_{\mathbf{F}}^{\mathbf{S},1}(a) < P_{\mathcal{A}}(a)$ and thus that $P_{\mathbf{F}}^{\mathbf{S},1} \neq 1$. If $P_{\mathcal{A}}(a) \neq 1$, we know that $P_{\mathcal{A}}(a) < 1$. From here, Proposition 3.3 gives us that $P_{\mathbf{F}}^{\mathbf{S},1}(a) < 1$ and thus that $P_{\mathbf{F}}^{\mathbf{S},1} \neq 1$. $\square$

4. Ranking Arguments in a Probabilistic Setting

Through the introduction, removal, or alteration of terms in a gradual semantics one may vastly alter the exact values it assigns to arguments. Hence, comparing an argument's strength to an exact threshold can only be informative when we are intimately familiar with the semantics used. We see this reflected in the principles used in the literature to study or define semantics [5, 12]; these principles only speak about strength values relative to those of other arguments, or to minimum or maximum values enforced by the framework and never do so in absolute values. Such properties are ultimately what distinguish two semantics and comparing arguments based on their relative strength is thus an appropriate use of gradual semantics.

Let us consider a practical example to better inform our intuitions. Suppose we are a university hiring committee tasked with filling a PhD position and there are two candidates to consider: Alex and Billy. We may model the suitability for hiring of Alex with an argument a and that of Billy with an argument b. Any reason for questioning the suitability of either candidate can now be modelled as an attacker of either argument. For instance, Alex's suitability may be brought into question based on doubts of their mastery of the English language (argument x with $x\mathcal{R}a$) and Billy may be considered a poor fit in the team as they are known to have had some conflict with other members of the research group (argument y with $y\mathcal{R}b$). Uncertainty is introduced into the graph by argument x (Alex has insufficient mastery of the English language) relying on an assumption made because Alex did not provide formal test results proving the contrary and by doubts whether argument y (there is known conflict between Billy and other team members) should constitute a valid attack on Billy's suitability as a candidate.

Say Alex is part of a demographic that is currently underrepresented in the university's staff, while Billy is not. Based on this fact, we prefer to hire Alex whenever they are at least as suitable a candidate as Billy. In order to select a candidate, we may wish to determine which candidate has the highest probability of being preferred. In other words, we want to find the probability that a is at least as strong as b and the probability that it is not and hire Alex if the former is greater.

Note that Definition 3.2 cannot be used to formalise this problem. To find these probabilities, we first need to rank arguments in each individual induced framework. We now define what it means that a is ranked at least as highly as b in $\mathbf{G}$, denoted by $\mathbf{G} \vDash_{\mathbf{S}} a \preceq b$. Note that here we follow the convention from [2], where $a \preceq b$ means "a is at least as strong as b".

Definition 4.1 (Ranking Arguments in Induced Graphs)**.** *Let* $\mathbf{G}' = \langle \mathcal{A}', \mathcal{R}' \rangle$ *be an induced graph of probabilistic argumentation framework*

$\mathbf{F} = \langle \mathcal{A}, P_{\mathcal{A}}, \mathcal{R}, P_{\mathcal{R}} \rangle$, *and let* $\mathbf{S}$ *be a gradual semantics. For* $a, b \in \mathcal{A}$, $\mathbf{G}' \vDash_{\mathbf{S}} a \preceq b$ *iff one of the following holds:*

- $a, b \in \mathcal{A}'$ *and* $\mathtt{Deg}^{\mathbf{S}}_{\mathbf{G}'}(a) \geq \mathtt{Deg}^{\mathbf{S}}_{\mathbf{G}'}(b)$, *or*
- *either* $\mathtt{Deg}^{\mathbf{S}}_{\mathbf{G}'}(b) = 0$ *or* $b \notin \mathcal{A}'$.

The second condition equates the arguments without any strength with those not present in a graph. Note that now we can define probability of $a \preceq b$ as

$$P_{\mathbf{F},\mathbf{S}}(a \preceq b) = \sum_{\mathbf{G} \in \mathtt{I}(\mathbf{F}), \mathbf{G} \vDash_{\mathbf{S}} a \preceq b} P^{I}_{\mathbf{F}}(\mathbf{G}) \tag{3}$$

and that we get the probability of $a \npreceq b$ as the complement of $a \preceq b$.

Now we emphasise that in some situations calculating probability that $a \preceq b$ is still insufficient. Suppose we introduce a third candidate into our example: Charlie, whose suitability is represented by argument c which also receives some uncertain attack. In this case the definition given by equation 3 no longer serves to determine the probability that, say, Alex is at least as strong candidate as the other two—even if we were to somehow combine $P_{\mathbf{F},\mathbf{S}}(a \preceq b)$ and $P_{\mathbf{F},\mathbf{S}}(a \preceq c)$—as induced graphs in which c is stronger than a may count toward $P_{\mathbf{F},\mathbf{S}}(a \preceq b)$ and graphs in which b is stronger than a may count toward $P_{\mathbf{F},\mathbf{S}}(a \preceq c)$. To properly determine this probability we require a formalism that considers multiple argument inequalities at the same time.

Definition 4.2 (Ranking Query)**.** *Given a probabilistic argumentation framework* $\mathbf{F} = \langle \mathcal{A}, P_{\mathcal{A}}, \mathcal{R}, P_{\mathcal{R}} \rangle$, *a ranking query is a Boolean combination of expressions of the form* $a \preceq b$ *with* $a, b \in \mathcal{A}$.

We will denote ranking queries with Greek letters $\alpha, \beta, \gamma, \ldots$ Note that using ranking queries one can also express that one argument is of strictly higher rank than another ($a \preceq b \wedge \neg b \preceq a$) and that two arguments have equal ranks ($a \preceq b \wedge b \preceq a$). By $\top$ we denote a ranking query of the form $\alpha \vee \neg\alpha$.

For a ranking query α and an induced graph $\mathbf{G}$, we define $\mathbf{G} \vDash_{\mathbf{S}} \alpha$ simply by extending Definition 4.1 with the cases of Boolean connectives in the standard way. For example, we define $\mathbf{G} \vDash_{\mathbf{S}} \alpha \vee \beta$ iff $\mathbf{G} \vDash_{\mathbf{S}} \alpha$ or $\mathbf{G} \vDash_{\mathbf{S}} \beta$. We say that two queries α and β are incompatible if there is no graph such that $\mathbf{G} \vDash_{\mathbf{S}} \alpha$ and $\mathbf{G} \vDash_{\mathbf{S}} \beta$. Note that the query $\top$ holds in every induced graph.

Definition 4.3 (Probability of Ranking Queries)**.** *Given a probabilistic argumentation framework* $\mathbf{F} = \langle \mathcal{A}, P_{\mathcal{A}}, \mathcal{R}, P_{\mathcal{R}} \rangle$ *and a gradual semantics* $\mathbf{S}$, *let* α *be a ranking query. The probability that the ranking on arguments indicated by* α *holds is:*

$$P_{\mathbf{F},\mathbf{S}}(\alpha) = \sum_{\mathbf{G} \in \mathtt{I}(\mathbf{F}), \mathbf{G} \vDash_{\mathbf{S}} \alpha} P^{I}_{\mathbf{F}}(\mathbf{G})$$

That is, $P_{\mathbf{F},\mathbf{S}}(x)$ is the sum of the probabilities of the induced graphs of $\mathbf{F}$ *that entail the query α under semantics* $\mathbf{S}$*.*

The additional expressivity offered by definition 4.3 allows us to successfully find the probability that Alex is at least as strong as the other candidates, namely by calculating $P_{\mathbf{F},\mathbf{S}}(a \preceq b \wedge a \preceq c)$.

The first part of the proposition below expresses finite additivity.

Proposition 4.4. *Given a PrAF* $\mathbf{F}$ *and a gradual semantics* $\mathbf{S}$*, let α be a ranking query.*

(1) *If α and β are incompatible, then $P_{\mathbf{F},\mathbf{S}}(\alpha \vee \beta) = P_{\mathbf{F},\mathbf{S}}(\alpha) + P_{\mathbf{F},\mathbf{S}}(\beta)$.*
(2) $P_{\mathbf{F},\mathbf{S}}(\top) = 1$.

Proof. Let $\mathbf{F}$ be a PrAF, $\mathbf{S}$ be a gradual semantics, and let α, β be ranking queries. We prove that (1) if α and β are incompatible, then $P_{\mathbf{F},\mathbf{S}}(\alpha \vee \beta) = P_{\mathbf{F},\mathbf{S}}(\alpha) + P_{\mathbf{F},\mathbf{S}}(\beta)$, and that (2) $P_{\mathbf{F},\mathbf{S}}(\top) = 1$.

(1) α and β being incompatible means there is no $\mathbf{G} \in \mathtt{I}(\mathbf{F})$ s.t. $\mathbf{G} \vDash_{\mathbf{S}} \alpha$ and $\mathbf{G} \vDash_{\mathbf{S}} \beta$. Whenever $\mathbf{G} \vDash_{\mathbf{S}} \alpha$ or $\mathbf{G} \vDash_{\mathbf{S}} \beta$, $\mathbf{G} \vDash_{\mathbf{S}} (\alpha \vee \beta)$. It follows through Definition 4.3 that $P_{\mathbf{F},\mathbf{S}}(\alpha \vee \beta) = P_{\mathbf{F},\mathbf{S}}(\alpha) + P_{\mathbf{F},\mathbf{S}}(\beta)$.
(2) For any $\mathbf{G} \in \mathtt{I}(\mathbf{F})$, $\mathbf{G} \vDash_{\mathbf{S}} \top$. It follows through Definition 4.3 and Proposition 2.23 that $P_{\mathbf{F},\mathbf{S}}(\top) = 1$.

□

The following result links Definition (4.3) and Definition (3.2).

Theorem 1 (Necessitation)**.** *Let $\mathbf{F} = \langle \mathcal{A}, P_{\mathcal{A}}, \mathcal{R}, P_{\mathcal{R}} \rangle$ be a PrAF and* $\mathbf{S}$ *a gradual semantics. For all arguments $a, b \in \mathcal{A}$, if $P_{\mathbf{F},\mathbf{S}}(a \preceq b) = 1$, then for every $t \in (0, 1]$ we have $P_{\mathbf{F}}^{\mathbf{S},t}(a) \geq P_{\mathbf{F}}^{\mathbf{S},t}(b)$.*

Proof. Let $\mathbf{F} = \langle \mathcal{A}, P_{\mathcal{A}}, \mathcal{R}, P_{\mathcal{R}} \rangle$ be a PrAF with $a, b \in \mathcal{A}$, let $\mathbf{S}$ be a gradual semantics, and let $t \in (0, 1]$. We assume $P_{\mathbf{F},\mathbf{S}}(a \preceq b) = 1$. With Definition 4.1 this implies that for any $\mathbf{G}' = \langle \mathcal{A}', \mathcal{R}' \rangle \in \mathtt{I}(\mathbf{F})$, $\mathbf{G}' \vDash_{\mathbf{S}} a \preceq b$ which means that for any for any $\mathbf{G}' = \langle \mathcal{A}', \mathcal{R}' \rangle \in \mathtt{I}(\mathbf{F})$ we have (1) $a, b \in \mathcal{A}'$ and $\mathtt{Deg}_{\mathbf{G}'}^{\mathbf{S}}(a) \geq \mathtt{Deg}_{\mathbf{G}'}^{\mathbf{S}}(b)$ or (2) either $\mathtt{Deg}_{\mathbf{G}'}^{\mathbf{S}}(b) = 0$ or $b \notin \mathcal{A}'$.

In the case where $a, b \in \mathcal{A}'$, we see that whenever $\mathbf{G} \mathrel{\overset{t}{\mid\!\sim}}_{\mathbf{S}} b$ we also have that $\mathbf{G} \mathrel{\overset{t}{\mid\!\sim}}_{\mathbf{S}} a$, since $\mathtt{Deg}_{\mathbf{G}'}^{\mathbf{S}}(a) \geq \mathtt{Deg}_{\mathbf{G}'}^{\mathbf{S}}(b)$. In the other case, we certainly do not have $\mathbf{G} \mathrel{\overset{t}{\mid\!\sim}}_{\mathbf{S}} b$ since either $\mathtt{Deg}_{\mathbf{G}'}^{\mathbf{S}}(b) = 0$ while $t \in (0, 1]$ or $\mathtt{Deg}_{\mathbf{G}'}^{\mathbf{S}}(b)$ is undefined.

From here we see that for any $\mathbf{G} \in \mathtt{I}(\mathbf{F})$, $\mathbf{G} \mathrel{\overset{t}{\mid\!\sim}}_{\mathbf{S}} b$ implies $\mathbf{G} \mathrel{\overset{t}{\mid\!\sim}}_{\mathbf{S}} a$. This gives us $P_{\mathbf{F}}^{\mathbf{S},t}(a) \geq P_{\mathbf{F}}^{\mathbf{S},t}(b)$. □

5. A Constellations Approach to Argument Strength

In the previous sections we saw how we may use the probabilities of a PrAF's induced graphs together with the strength values assigned to each argument in those induced graphs by a gradual semantics to determine the probability that an argument's strength meets some threshold or to determine the probability that some ranking query on arguments is satisfied. Considering each of the PrAF's induced graphs, as we did in answering both of these questions, is characteristic of the constellations approach to probabilistic argumentation. In this section, we explore the possibility of taking a similar approach in assigning each argument a unique strength; we look for a method that, given a choice of gradual semantics **S**, assigns each argument an overall strength that considers both the probabilities of induced graphs and the strengths assigned by **S** in their context. First, we present a generalised notion of this new method for assigning strengths. Then, we discuss some of the properties and assumptions of this new approach and present a set of principles for them inspired by the principles for gradual semantics present in the literature. Finally, we propose a specification of the generalised method for assigning strengths based on a gradual semantics and discuss the properties of this specification.

Just as a gradual semantics is a function transforming an argumentation graph into a weighting on its elements, our generalised method for assigning unique strengths to arguments in a probabilistic argumentation framework—henceforth called a generalised semantics—is a function transforming a PrAF into a weighting on its arguments.

Definition 5.1 (Generalised Semantics). *A generalised semantics is a function* $\mathbb{S}$ *transforming any probabilistic argumentation framework* $\mathbf{F} = \langle \mathcal{A}, P_{\mathcal{A}}, \mathcal{R}, P_{\mathcal{R}} \rangle \in \mathtt{PrAF}$ *into a weighting* $\mathtt{Deg}^{\mathbb{S}}_{\mathbf{F}}$ *on* $\mathcal{A}$ *(i.e.,* $\mathtt{Deg}^{\mathbb{S}}_{\mathbf{F}} : \mathcal{A} \to [0,1]$*). For any* $a \in \mathcal{A}$*,* $\mathtt{Deg}^{\mathbb{S}}_{\mathbf{F}}(a)$ *represents the strength of* a*.*

5.1. Principles for Generalised Semantics.

As the PrAFs to which generalised semantics are applied are an extension of the AGs to which gradual semantics are applied, it is worth investigating which principles used in the study of gradual semantics (or at least the intuitions underlying them) transfer to the constellations approach. We study the generally desirable principles proposed in [5] which are recalled and adjusted to non-weighted graphs in the appendix and find that while many translate naturally to the new setting, three require more extensive alteration to be sensible, and two are not generally desirable. Before proposing the principles resulting from this examination, let us consider these two principles, equivalence and reinforcement, that are not generally desirable in the new setting where we want both the probabilities of induced graphs and the strengths of arguments in them to contribute to the overall strength of an argument.

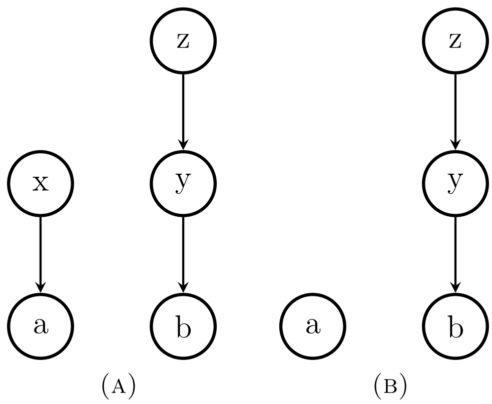

FIGURE 3. A probabilistic argumentation framework illustrating compensation between strength and probability where argument x is uncertain and all other elements are certain: (a) shows the entire PrAF and its induced graph where x is present; (b) shows the induced graph where x is not present.

The intuition underlying the equivalence principle is that the strength of an argument in an argumentation graph should depend only on the strength of its direct attackers. The intuition underlying the reinforcement principle adds to this that increasing the strength of an attacker should increase the impact of its attack. To show how these intuitions may not generally hold when assuming the overall strength of attackers in a PrAF, we present the following: consider the PrAF $\mathbf{F}$ shown in figure 3 with two induced graphs $\mathbf{G}$ and $\mathbf{G}'$ shown in subfigures a and b respectively. Recall that we want the overall strength of an argument to be based on the strengths assigned to it in each induced graph by a gradual semantics $\mathbf{S}$. We are interested in the overall strengths of arguments a and b and how they are affected by their respective attackers x and y. If we we were to select h-Categoriser (def. 2.7) as $\mathbf{S}$, we would have $\mathtt{Deg}^{\mathbf{S}}_{\mathbf{G}}(a) = 1/2$, $\mathtt{Deg}^{\mathbf{S}}_{\mathbf{G}'}(a) = 1$, and $\mathtt{Deg}^{\mathbf{S}}_{\mathbf{G}}(b) = \mathtt{Deg}^{\mathbf{S}}_{\mathbf{G}'}(b) = 2/3$. The strength of y would be $1/2$ in all induced graphs and we would have $\mathtt{Deg}^{\mathbf{S}}_{\mathbf{G}}(x) = 1$. Based on the neutrality principle's prescription that an argument with strength 0 contributes the same as no argument, we say $\mathtt{Deg}^{\mathbf{S}}_{\mathbf{G}'}(x) = 0$.

Now that we know the strength of the arguments in each induced graph, the question becomes how to aggregate these. Given that, in the constellations approach, we consider each induced graph as a possible world with some probability, we might approach the matter similarly to an expected value and say that the amount an argument's strength in one induced graph contributes to its overall strength should be directly proportional to the probability of that induced graph. That is to say we multiply each strength in an induced graph with that graph's probability and sum over graphs to find the overall strength. If we take this approach, we may alter the value of $P_{\mathcal{A}}(x)$, and by extension $P^I_{\mathbf{F}}(\mathbf{G})$ and $P^I_{\mathbf{F}}(\mathbf{G}')$, to create scenarios where we may not desire equivalence or reinforcement.

First consider equivalence: say $P_{\mathcal{A}}(x) = 1/2$. This gives $P^I_{\mathbf{F}}(\mathbf{G}) = P^I_{\mathbf{F}}(\mathbf{G}') = 1/2$. The overall strength of x and y is equal, with $\mathtt{Deg}^{\mathbb{S}}_{\mathbf{F}}(x) = 1 \cdot 1/2 + 0 \cdot 1/2 = 1/2$ and $\mathtt{Deg}^{\mathbb{S}}_{\mathbf{F}}(y) = 1/2 \cdot 1/2 + 1/2 \cdot 1/2 = 1/2$. The attackers of a and b are thus equally strong overall, and the certainty with which they are attacked is also equal, but we have $\mathtt{Deg}^{\mathbb{S}}_{\mathbf{F}}(a) = 1/2 \cdot 1/2 + 1 \cdot 1/2 = 3/4$ while $\mathtt{Deg}^{\mathbb{S}}_{\mathbf{F}}(b) = 2/3 \cdot 1/2 + 2/3 \cdot 1/2 = 2/3$; a and b do not have the same overall strength, nor do we necessarily want them to.

To demonstrate how reinforcement is not always desirable we say $P_{\mathcal{A}}(x) = 2/3$. We now end up with $\mathtt{Deg}^{\mathbb{S}}_{\mathbf{F}}(x) = 2/3 > \mathtt{Deg}^{\mathbb{S}}_{\mathbf{F}}(y) = 1/2$ and $\mathtt{Deg}^{\mathbb{S}}_{\mathbf{F}}(a) = \mathtt{Deg}^{\mathbb{S}}_{\mathbf{F}}(b) = 2/3$. The attacker of a is stronger overall than that of b, but it is reasonable to desire that a and b are equally strong overall.

Having seen which principles' intuitions do not transfer nicely to the constellations setting, we present seven principles based on those presented in [5]. For this, we require the notion of isomorphisms on PrAFs:

Definition 5.2 (PrAF Isomorphism)**.** *Let* $\mathbf{F} = \langle \mathcal{A}, P_{\mathcal{A}}, \mathcal{R}, P_{\mathcal{R}} \rangle$ *and* $\mathbf{F}' = \langle \mathcal{A}', P'_{\mathcal{A}}, \mathcal{R}', P'_{\mathcal{R}} \rangle$ *be two PrAFs. An isomorphism from* $\mathbf{F}$ *to* $\mathbf{F}'$ *is a bijective function* f *from* $\mathcal{A}$ *to* $\mathcal{A}'$ *such that:*

- $\forall a \in \mathcal{A}, P_{\mathcal{A}}(a) = P'_{\mathcal{A}}(f(a))$*;*
- $\forall a, b \in \mathcal{A}, a\mathcal{R}b$ *iff* $f(a)\mathcal{R}'f(b)$*;*
- $\forall (a, b) \in \mathcal{R}, P_{\mathcal{R}}((a, b)) = P'_{\mathcal{R}}((f(a), f(b)))$

If $\mathbf{F} = \mathbf{F}'$*, we call any isomorphism from* $\mathbf{F}$ *to* $\mathbf{F}'$ *an automorphism.*

The first principle, PrAF Anonymity, expresses that an argument's evaluation does not depend on it's identity or the structure underlying it's abstract representation.

Principle 5.3 (PrAF Anonymity)**.** *A generalised semantics* $\mathbb{S}$ *satisfies anonymity iff for any two PrAFs* $\mathbf{F} = \langle \mathcal{A}, P_{\mathcal{A}}, \mathcal{R}, P_{\mathcal{R}} \rangle$ *and* $\mathbf{F}' = \langle \mathcal{A}', P'_{\mathcal{A}}, \mathcal{R}', P'_{\mathcal{R}} \rangle$*, and any isomorphism* f *from* $\mathbf{F}$ *to* $\mathbf{F}'$*, the following holds:* $\forall a \in \mathcal{A}, \mathtt{Deg}^{\mathbb{S}}_{\mathbf{F}}(a) = \mathtt{Deg}^{\mathbb{S}}_{\mathbf{F}'}(f(a))$*.*

The PrAF Independence principle dictates that the strength of an argument cannot depend on arguments to which the argument is not connected.

Principle 5.4 (PrAF Independence)**.** *A generalised semantics* $\mathbb{S}$ *satisfies independence iff for any two PrAFs* $\mathbf{F} = \langle \mathcal{A}, P_{\mathcal{A}}, \mathcal{R}, P_{\mathcal{R}} \rangle$ *and* $\mathbf{F}' = \langle \mathcal{A}', P'_{\mathcal{A}}, \mathcal{R}', P'_{\mathcal{R}} \rangle$ *where* $\mathcal{A} \cap \mathcal{A}' = \emptyset$*, it holds that:* $\forall a \in \mathcal{A}, \mathtt{Deg}^{\mathbb{S}}_{\mathbf{F}}(a) = \mathtt{Deg}^{\mathbb{S}}_{\mathbf{F} \oplus \mathbf{F}'}(a)$*.*

The third principle, PrAF Directionality, states that the strength of an argument a in a PrAF can only depend on an argument b if a path from b to a, that is a finite non-empty sequence of arguments $\langle x_1, \ldots, x_n \rangle$ s.t. $x_1 = b, x_n = a$, and $\forall i < n,\ x_i \mathcal{R} x_{i+1}$, exists in the PrAF.

Principle 5.5 (PrAF Directionality)**.** *A generalised semantics* $\mathbb{S}$ *satisfies directionality iff for any two PrAFs* $\mathbf{F} = \langle \mathcal{A}, P_{\mathcal{A}}, \mathcal{R}, P_{\mathcal{R}} \rangle$ *and* $\mathbf{F}' = \langle \mathcal{A}, P_{\mathcal{A}}, \mathcal{R}', P'_{\mathcal{R}} \rangle$ *where* $\mathcal{R}' = \mathcal{R} \cup \{(a, b)\}$ *and* $\forall r \in \mathcal{R}, P'_{\mathcal{R}}(r) = P_{\mathcal{R}}(r)$*, the following holds: for any* $x \in \mathcal{A}$*, if there is no path from* b *to* x*,* $\mathtt{Deg}^{\mathbb{S}}_{\mathbf{F}}(x) = \mathtt{Deg}^{\mathbb{S}}_{\mathbf{F}'}(x)$*.*

For PrAF Maximality to hold, the strength of any argument that is not attacked should be equal to its probability.

Principle 5.6 (PrAF Maximality)**.** *A generalised semantics* $\mathbb{S}$ *satisfies probability maximality iff for any PrAF* $\mathbf{F} = \langle \mathcal{A}, P_{\mathcal{A}}, \mathcal{R}, P_{\mathcal{R}} \rangle$*,* $\forall a \in \mathcal{A}$*, it holds that: if* $\mathtt{Att}_{\mathbf{F}}(a) = \emptyset$*, then* $\mathtt{Deg}^{\mathbb{S}}_{\mathbf{F}}(a) = P_{\mathcal{A}}(a)$*.*

PrAF Weakening requires that the strength of an argument is decreased when it is attacked by an argument that is not worthless.

Principle 5.7 (PrAF Weakening)**.** *A generalised semantics* $\mathbb{S}$ *satisfies weakening iff for any PrAF* $\mathbf{F} = \langle \mathcal{A}, P_{\mathcal{A}}, \mathcal{R}, P_{\mathcal{R}} \rangle$*,* $\forall a \in \mathcal{A}$*, it holds that: if* $\exists b \in \mathtt{Att}_{\mathbf{F}}(a)$ *s.t.* $\mathtt{Deg}^{\mathbb{S}}_{\mathbf{F}}(b) > 0$ *then* $\mathtt{Deg}^{\mathbb{S}}_{\mathbf{F}}(a) < P_{\mathcal{A}}(a)$*.*

PrAF Weakening Soundness goes beyond PrAF Weakening in that it enforces that attacks are the only source of strength loss.

Principle 5.8 (PrAF Weakening Soundness)**.** *A generalised semantics* $\mathbb{S}$ *satisfies weakening soundness iff for any PrAF* $\mathbf{F} = \langle \mathcal{A}, P_{\mathcal{A}}, \mathcal{R}, P_{\mathcal{R}} \rangle$*,* $\forall a \in \mathcal{A}$*, the following holds: if* $\mathtt{Deg}^{\mathbb{S}}_{\mathbf{F}}(a) < P_{\mathcal{A}}(a)$ *then* $\exists b \in \mathtt{Att}_{\mathbf{F}}(a)$ *s.t.* $\mathtt{Deg}^{\mathbb{S}}_{\mathbf{F}}(b) > 0$*.*

The PrAF Resilience principle requires any argument present in a PrAF be assigned a positive strength.

Principle 5.9 (PrAF Resilience)**.** *A generalised semantics* $\mathbb{S}$ *satisfies resilience iff for any PrAF* $\mathbf{F} = \langle \mathcal{A}, P_{\mathcal{A}}, \mathcal{R}, P_{\mathcal{R}} \rangle$*,* $\forall a \in \mathcal{A}$*,* $\mathtt{Deg}^{\mathbb{S}}_{\mathbf{F}}(a) > 0$*.*

We have already discussed that if we follow the constellation approach, we should not expect that the Equivalence principle holds. Now we turn to a special case of Equivalence proposed in [8], called Symmetry. Originally it states that two arguments a and b that have the same sets of attackers also have equal strengths. Similarly as for Equivalence, in the case of PrAFs that "symmetry" between a and b breaks easily when we zoom in to induced graphs. For example, if a and b are not certain and there are different (asymmetric) attacks from a and b towards their own attackers, the induced graph in which only a appears will be very different from those in which b occurs. In order to enforce symmetry we require that a and b interact with their attackers, attackers of their attackers, and so on in a symmetric way.

Definition 5.10 (Attack Structure)**.** *For PrAF* $\mathbf{F} = \langle \mathcal{A}, P_{\mathcal{A}}, \mathcal{R}, P_{\mathcal{R}} \rangle$ *and* $a \in \mathcal{A}$*, the attack structure of* a *in* $\mathbf{F}$ *is* $\mathtt{Str}_{\mathbf{F}}(a) = \{a\} \cup \{c \in \mathcal{A} \mid$ *there is a path from* c *to* $a\}$*. We denote by* $\mathtt{Str}_{\mathbf{F}}(a, b)$ *the set* $\mathtt{Str}_{\mathbf{F}}(a) \cup \mathtt{Str}_{\mathbf{F}}(b)$*.*

Using attack structures, we may now define PrAF Symmetry.

Principle 5.11 (PrAF Symmetry)**.** *A generalised semantics* $\mathbb{S}$ *satisfies PrAF Symmetry iff for every PrAF* $\mathbf{F} = \langle \mathcal{A}, P_{\mathcal{A}}, \mathcal{R}, P_{\mathcal{R}} \rangle$*,* $\forall a, b \in \mathcal{A}$*, the following holds: if* $f : \mathbf{F}_{|\mathtt{Str}_{\mathbf{F}}(a,b)} \to \mathbf{F}_{|\mathtt{Str}_{\mathbf{F}}(a,b)}$ *s.t.* $f(a) = b$*,* $f(b) = a$ *and* $f(x) = x$ *otherwise, is an automorphism, then* $\mathtt{Deg}^{\mathbb{S}}_{\mathbf{F}}(a) = \mathtt{Deg}^{\mathbb{S}}_{\mathbf{F}}(b)$*.*

The principles Neutrality and Counting build on the idea of symmetry, and say that if an attack is added to one of two arguments in a symmetric situation, the attack will additionally harm the target if the strength of that attacker is positive, otherwise it will not. We apply that intuition directly to our notion of PrAF Symmetry. First, we consider the intuition about attackers with zero strength to get PrAF Neutrality.

Principle 5.12 (PrAF Neutrality)**.** *A generalised semantics* $\mathbb{S}$ *satisfies Praf Neutrality iff for every PrAF* $\mathbf{F} = \langle \mathcal{A}, P_{\mathcal{A}}, \mathcal{R}, P_{\mathcal{R}} \rangle$*,* $\forall a, b \in \mathcal{A}$*, the following holds: if there exists* c *such that* $(c, b) \in \mathcal{R}$*,* $(c, b) \notin \mathcal{R}$*, and* $f : \mathbf{F}_{|\mathtt{Str}_{\mathbf{F}}(a,b) \setminus \{c\}} \to \mathbf{F}_{|\mathtt{Str}_{\mathbf{F}}(a,b) \setminus \{c\}}$ *s.t.* $f(a) = b$*,* $f(b) = a$ *and* $f(x) = x$ *otherwise, is an automorphism, then* $\mathtt{Deg}^{\mathbb{S}}_{\mathbf{F}}(c) = 0$ *implies* $\mathtt{Deg}^{\mathbb{S}}_{\mathbf{F}}(a) = \mathtt{Deg}^{\mathbb{S}}_{\mathbf{F}}(b)$*.*

Now, we apply the intuition about attackers with positive strength to get PrAF Counting.

Principle 5.13 (PrAF Counting)**.** *A generalised semantics* $\mathbb{S}$ *satisfies PrAF Counting iff for every PrAF* $\mathbf{F} = \langle \mathcal{A}, P_{\mathcal{A}}, \mathcal{R}, P_{\mathcal{R}} \rangle$*,* $\forall a, b \in \mathcal{A}$*, the following holds: if there exists* c *such that* $(c, b) \in \mathcal{R}$*,* $(c, b) \notin \mathcal{R}$*, and* $f : \mathbf{F}_{|\mathtt{Str}_{\mathbf{F}}(a,b) \setminus \{c\}} \to \mathbf{F}_{|\mathtt{Str}_{\mathbf{F}}(a,b) \setminus \{c\}}$ *s.t.* $f(a) = b$*,* $f(b) = a$ *and* $f(x) = x$ *otherwise, is an automorphism, then* $\mathtt{Deg}^{\mathbb{S}}_{\mathbf{F}}(c) > 0$ *implies* $\mathtt{Deg}^{\mathbb{S}}_{\mathbf{F}}(a) > \mathtt{Deg}^{\mathbb{S}}_{\mathbf{F}}(b)$*.*

The following result shows that PrAF Symmetry is already a consequence of a subset of other principles.

Theorem 2. *If a generalised semantics* $\mathbb{S}$ *satisfies PrAF Anonymity, PrAF Independence and PrAF Directionality, then* $\mathbb{S}$ *also satisfies PrAF Symmetry.*

Proof. Let $\mathbf{F} = \langle \mathcal{A}, P_{\mathcal{A}}, \mathcal{R}, P_{\mathcal{R}} \rangle$ be a PrAF and $\mathbb{S}$ be a generalised semantics that satisfies PrAF Anonymity, PrAF Independence, and PrAF Directionality. Suppose that $a, b \in \mathcal{A}$ and there is an automorphism $f : \mathbf{F}_{|\mathtt{Str}_{\mathbf{F}}(a,b)} \to \mathbf{F}_{|\mathtt{Str}_{\mathbf{F}}(a,b)}$ s.t. $f(a) = b$, $f(b) = a$ and $f(x) = x$ otherwise. Since $\mathbb{S}$ satisfies PrAF Anonymity and f is an automorphism, we know $\mathtt{Deg}^{\mathbb{S}}_{\mathbf{F}_{|\mathtt{Str}_{\mathbf{F}}(a,b)}}(a) = \mathtt{Deg}^{\mathbb{S}}_{\mathbf{F}_{|\mathtt{Str}_{\mathbf{F}}(a,b)}}(b)$. Let $\mathtt{Str}_{\mathbf{F}}(a, b) = \langle \mathcal{A}', P'_{\mathcal{A}}, \mathcal{R}', P'_{\mathcal{R}} \rangle$. Additionally, let $\mathbf{F}' = \langle \mathcal{A} \setminus \mathcal{A}', (P_{\mathcal{A}})_{|\mathcal{A} \setminus \mathcal{A}'}, \mathcal{R}_{|(\mathcal{A} \setminus \mathcal{A}') \times (\mathcal{A} \setminus \mathcal{A}')}, (P_{\mathcal{R}})_{|(\mathcal{R}_{|(\mathcal{A} \setminus \mathcal{A}') \times (\mathcal{A} \setminus \mathcal{A}')})} \rangle$

be $\mathbf{F}$ restricted to the complement of the attack structure of a and b. Since $\mathbf{G}$ satisfies PrAF Independence, we know $\mathtt{Deg}^{\mathbb{S}}_{\mathbf{F}_{|\mathtt{Str}_{\mathbf{F}}(a)} \oplus \mathbf{F}'}(a) = \mathtt{Deg}^{\mathbb{S}}_{\mathbf{F}_{|\mathtt{Str}_{\mathbf{F}}(a,b)}}(a) = \mathtt{Deg}^{\mathbb{S}}_{\mathbf{F}_{|\mathtt{Str}_{\mathbf{F}}(a,b)}}(b) = \mathtt{Deg}^{\mathbb{S}}_{\mathbf{F}_{|\mathtt{Str}_{\mathbf{F}}(a)} \oplus \mathbf{F}'}(b)$. One by one, we may now add the attacks between the attack structure of a and b and their probabilities present in $\mathbf{F}$ to $\mathbf{F}_{|\mathtt{Str}_{\mathbf{F}}(a)} \oplus \mathbf{F}'$ to finally return to $\mathbf{F}$. Since we know none of these attacks creates a path to a or b as they are not part of the attack structure of a and b, we know that at each step $\mathbb{S}$'s satisfaction of PrAF Directionality ensures the degrees of a and b do not change. This gives that $\mathtt{Deg}^{\mathbb{S}}_{\mathbf{F}}(a) = \mathtt{Deg}^{\mathbb{S}}_{\mathbf{F}}(b)$. □

We now present the result which claims that from a subset of the principles it follows that an argument's strength is bounded by its probability.

Theorem 3. *If a generalised semantics $\mathbb{S}$ satisfies PrAF Independence, PrAF Maximality, PrAF Weakening, PrAF Neutrality, and PrAF Directionality, then for any $\mathbf{F} = \langle \mathcal{A}, P_{\mathcal{A}}, \mathcal{R}, P_{\mathcal{R}} \rangle \in \mathtt{PrAF}$, for any $a \in \mathcal{A}$, $\mathtt{Deg}^{\mathbb{S}}_{\mathbf{F}}(a) \leq P_{\mathcal{A}}(a)$.*

Proof. Let $\mathbf{F} = \langle \mathcal{A}, P_{\mathcal{A}}, \mathcal{R}, P_{\mathcal{R}} \rangle$ be a PrAF with $a \in \mathcal{A}$ and let $\mathbb{S}$ be a generalised semantics that satisfies PrAF Independence, PrAF Maximality, PrAF Weakening, PrAF Neutrality, and PrAF Directionality. For a, we have either $\mathtt{Att}_{\mathbf{F}}(a) = \emptyset$ or $\mathtt{Att}_{\mathbf{F}}(a) \neq \emptyset$. If $\mathtt{Att}_{\mathbf{F}}(a) = \emptyset$, we get through $\mathbb{S}$ satisfying PrAF Maximality that $\mathtt{Deg}^{\mathbb{S}}_{\mathbf{F}}(a) = P_{\mathcal{A}}(a)$. If $\mathtt{Att}_{\mathbf{F}}(a) \neq \emptyset$ we have either $\exists b \in \mathtt{Att}_{\mathbf{F}}(a)$, $\mathtt{Deg}^{\mathbb{S}}_{\mathbf{F}}(b) > 0$ or $\forall b \in \mathtt{Att}_{\mathbf{F}}(a)$, $\mathtt{Deg}^{\mathbb{S}}_{\mathbf{F}}(b) = 0$. If $\exists b \in \mathtt{Att}_{\mathbf{F}}(a)$, s.t. $\mathtt{Deg}^{\mathbb{S}}_{\mathbf{F}}(b) > 0$, $\mathbb{S}$ satisfying PrAF Weakening gives us $\mathtt{Deg}^{\mathbb{S}}_{\mathbf{F}}(a) < P_{\mathcal{A}}(a)$. In the case $\forall b \in \mathtt{Att}_{\mathbf{F}}(a)$, $\mathtt{Deg}^{\mathbb{S}}_{\mathbf{F}}(b) = 0$, we continue by induction on the number of attackers of a. If a has one attacker, then let us consider the framework obtained from $\mathbf{F}$ by adding only one argument b with the same probability as a, ie. $P_{\mathcal{A}}(b) = P_{\mathcal{A}}(a)$. By PrAF Anonimity and Directionality, the degrees of all arguments from $\mathbf{F}$ will stay the same as in the original graph, while dhe degree of b will be $P_{\mathcal{A}}(b) = P_{\mathcal{A}}(a)$ by PrAF Maximality. Finally, by PrAF Neutrality we have that degrees of a and b must be the same. For the induction step, let $\mathtt{Att}_{\mathbf{F}}(a) = \{x_1, \ldots, x_{n+1}\}$. Reasoning as above, we extend the framework with a new argument b and n attacks toward b: from x_1, x_2,…,x_n. By induction hypothesis, degree of b will be zero, and the degree of a must be the same as the degree of b by PrAF Neutrality. □

Our next formal result states that there is always a subset of arguments in a probabilistic argumentation framework which impacts the strength of a given argument, namely its attack structure.

Theorem 4. *If a semantics $\mathbb{S}$ satisfies PrAF-Independence and PrAF-Directionality, then for any $\mathbf{F} = \langle \mathcal{A}, P_{\mathcal{A}}, \mathcal{R}, P_{\mathcal{R}} \rangle$, for any $a \in \mathcal{A}$, the following holds: $\mathtt{Deg}^{\mathbb{S}}_{\mathbf{F}}(a) = \mathtt{Deg}^{\mathbb{S}}_{\mathbf{F}_{|\mathtt{Str}_{\mathbf{F}}(a)}}(a)$.*

Proof. Let $\mathbf{F} = \langle \mathcal{A}, P_{\mathcal{A}}, \mathcal{R}, P_{\mathcal{R}} \rangle$ be a PrAF with $a \in \mathcal{A}$ and let $\mathbb{S}$ be a generalised semantics that satisfies PrAF Independence and PrAF Directionality.

$$\mathbf{F}_{|\mathtt{Str}_{\mathbf{F}}(a)} = \langle \mathtt{Str}_{\mathbf{F}}(a), (P_{\mathcal{A}})_{|\mathtt{Str}_{\mathbf{F}}(a)}, \mathcal{R}_{|\mathtt{Str}_{\mathbf{F}}(a) \times \mathtt{Str}_{\mathbf{F}}(a)}, (P_{\mathcal{R}})_{|(\mathcal{R}_{|\mathtt{Str}_{\mathbf{F}}(a) \times \mathtt{Str}_{\mathbf{F}}(a)})} \rangle.$$

Let

$$\mathbf{F}' = \langle \mathcal{A} \setminus \mathtt{Str}_{\mathbf{F}}(a), (P_{\mathcal{A}})_{|\mathcal{A} \setminus \mathtt{Str}_{\mathbf{F}}(a)}, \mathcal{R}_{|(\mathcal{A} \setminus \mathtt{Str}_{\mathbf{F}}(a)) \times (\mathcal{A} \setminus \mathtt{Str}_{\mathbf{F}}(a))}, (P_{\mathcal{R}})_{|(\mathcal{R}_{|(\mathcal{A} \setminus \mathtt{Str}_{\mathbf{F}}(a)) \times (\mathcal{A} \setminus \mathtt{Str}_{\mathbf{F}}(a))})} \rangle$$

be the restriction of $\mathbf{F}$ to all arguments but the attack structure of a. From $\mathbf{S}$ satisfying PrAF Independence we get $\mathtt{Deg}^{\mathbb{S}}_{\mathbf{F}_{|\mathtt{Str}_{\mathbf{F}}(a)}}(a) = \mathtt{Deg}^{\mathbb{S}}_{\mathbf{F}_{|\mathtt{Str}_{\mathbf{F}}(a)} \oplus \mathbf{F}'}(a)$. $\mathbf{F}_{|\mathtt{Str}_{\mathbf{F}}(a)} \oplus \mathbf{F}'$ differs from $\mathbf{F}$ in the attacks between the attack structure of a and its complement, none of which, per Definition 5.10, form a path to a that is not already in $\mathbf{F}_{|\mathtt{Str}_{\mathbf{F}}(a)}$. From $\mathbb{S}$ satisfying PrAF Directionality, we see that we may add these attacks to $\mathbf{F}_{|\mathtt{Str}_{\mathbf{F}}(a)} \oplus \mathbf{F}'$ so that we get $\mathbf{F}$ without altering the strength of a. This gives that $\mathtt{Deg}^{\mathbb{S}}_{\mathbf{F}}(a) = \mathtt{Deg}^{\mathbb{S}}_{\mathbf{F}_{|\mathtt{Str}_{\mathbf{F}}(a)}}(a)$. $\square$

5.2. **Expected Strength Semantics.** With a set of principles—or desirable properties of a generalised semantics—laid out, let us consider one way of specifying a generalised semantics $\mathbb{S}$ using a gradual semantics $\mathbf{S}$ and consider how this specification relates to the different principles. When we were discussing equivalence and reinforcement, we weighted the contribution of an argument's strength in an induced graph with that induced graph's probability. Formalising this gives:

Definition 5.14 (Expected Strength Semantics)**.** *Given a gradual semantics* $\mathbf{S}$*, the expected strength semantics based on* $\mathbf{S}$*, denoted* $E(\mathbf{S})$*, is the generalised semantics such that* $\forall \mathbf{F} = \langle \mathcal{A}, P_{\mathcal{A}}, \mathcal{R}, P_{\mathcal{R}} \rangle \in \mathtt{PrAF}, \forall a \in \mathcal{A}$,

$$\mathtt{Deg}^{E(\mathbf{S})}_{\mathbf{F}}(a) = \sum_{\mathbf{G} = \langle \mathcal{A}', \mathcal{R}' \rangle \in \mathtt{I}(\mathbf{F}), a \in \mathcal{A}'} P^{I}_{\mathbf{F}}(\mathbf{G}) \cdot \mathtt{Deg}^{\mathbf{S}}_{\mathbf{G}}(a)$$

The next results show that if the underlying gradual semantics $\mathbf{S}$ satisfies some principles from the literature (presented Section 2), then $E(\mathbf{S})$ satisfies the principles for generalised semantics proposed in this section.

$E(\mathbf{S})$'s satisfaction of the first four principles follows directly from $\mathbf{S}$ satisfying these principles' gradual semantics counterparts.

Theorem 5. *Let* $\mathbf{S}$ *be a gradual semantics.*

- *If* $\mathbf{S}$ *satisfies Anonymity, then* $E(\mathbf{S})$ *satisfies PrAF Anonymity.*

- *If* $\mathbf{S}$ *satisfies Independence, then* $E(\mathbf{S})$ *satisfies PrAF Independence.*
- *If* $\mathbf{S}$ *satisfies Directionality, then* $E(\mathbf{S})$ *satisfies PrAF Directionality.*
- *If* $\mathbf{S}$ *satisfies Maximality, then* $E(\mathbf{S})$ *satisfies PrAF Maximality.*

Proof. We prove the four items in Theorem 5 in turn:

- Let $\mathbf{F}' = \langle \mathcal{A}', P'_{\mathcal{A}}, \mathcal{R}', P'_{\mathcal{R}} \rangle$ be a PrAF s.t. there is an isomorphism f from $\mathbf{F}$ to $\mathbf{F}'$ and suppose $\mathbf{S}$ satisfies Anonymity. f implies the existence a bijective function g from $\mathtt{I}(\mathbf{F})$ to $\mathtt{I}(\mathbf{F}')$ s.t. for all $\mathbf{G} \in \mathtt{I}(\mathbf{F})$, there is an isomorphism from $\mathbf{G}$ to $g(\mathbf{G})$ and $P^I_{\mathbf{F}}(\mathbf{G}) = P^I_{\mathbf{F}'}(g(\mathbf{G}))$. $\mathbf{S}$ satisfying Anonymity now means that $\mathtt{Deg}^{\mathbf{S}}_{\mathbf{G}}(a) = \mathtt{Deg}^{\mathbf{S}}_{\mathbf{G}'}(a)$. Combining this with the fact that $P^I_{\mathbf{F}}(\mathbf{G}) = P^I_{\mathbf{F}'}(g(\mathbf{G}))$, through Definition 5.14, we see that $\mathtt{Deg}^{E(\mathbf{S})}_{\mathbf{F}}(a) = \mathtt{Deg}^{E(\mathbf{S})}_{\mathbf{F}'}(a)$.
- Let $\mathbf{F}' = \langle \mathcal{A}', P'_{\mathcal{A}}, \mathcal{R}', P'_{\mathcal{R}} \rangle$ be a PrAF s.t. $\mathcal{A} \cap \mathcal{A}' = \emptyset$ and suppose $\mathbf{S}$ satisfies Independence. Since the $\mathcal{A}$ and $\mathcal{A}'$ are disjoint, there exists a function f from $\mathtt{I}(\mathbf{F})$ to the power set of $\mathtt{I}(\mathbf{F} \oplus \mathbf{F}')$ s.t. $\forall \mathbf{G}'' = \langle \mathcal{A}'', \mathcal{R}'' \rangle \in \mathtt{I}(\mathbf{F})$, $f(\mathbf{G}'') = \{\mathbf{G}''' = \langle \mathcal{A}''', \mathcal{R}''' \rangle \in \mathtt{I}(\mathbf{F} \oplus \mathbf{F}') | \forall a \in \mathcal{A}, a \in \mathcal{A}''' \text{ iff } a \in \mathcal{A}'' \wedge \forall r \in \mathcal{R}, r \in \mathcal{R}''' \text{ iff } r \in \mathcal{R}''\}$. For any two $\mathbf{G}, \mathbf{G} \in \mathtt{I}(\mathbf{F})$, $f(\mathbf{G}) \cap f(\mathbf{G}') = \emptyset$ and $\bigcup_{\mathbf{G} \in \mathtt{I}(\mathbf{F})} f(\mathbf{G}) = \mathtt{I}(\mathbf{F} \oplus \mathbf{F}')$. Definition 2.22 gives that for any $\mathbf{G} \in \mathtt{I}(\mathbf{F})$,

$$P^I_{\mathbf{F}}(\mathbf{G}) = \sum_{\mathbf{G}' \in f(\mathbf{G})} P^I_{\mathbf{F} \oplus \mathbf{F}'}(\mathbf{G}') \tag{4}$$

  For any $\mathbf{G}'' = \langle \mathcal{A}'', \mathcal{R}'' \rangle \in \mathtt{I}(\mathbf{F})$, each $\mathbf{G}''' = \langle \mathcal{A}''', \mathcal{R}''' \rangle \in f(\mathbf{G})$ can be written as $\mathbf{G}'' \oplus \mathbf{G}''''$ with $\mathbf{G}''$ and $\mathbf{G}''''$ having disjoint sets of arguments. Since $\mathbf{S}$ satisfies Independence we now see $\forall a \in \mathcal{A}$, $\forall \mathbf{G} \in \mathtt{I}(\mathbf{F})$, $\forall \mathbf{G}' \in \mathtt{I}(\mathbf{F}')$, $\mathtt{Deg}^{\mathbf{S}}_{\mathbf{G}}(a) = \mathtt{Deg}^{\mathbf{S}}_{\mathbf{G}'}(a)$. Combining this with Equation 4 proves the result.
- Let $\mathbf{F}' = \langle \mathcal{A}, P_{\mathcal{A}}, \mathcal{R}', P'_{\mathcal{R}} \rangle$ be a PrAFs with $a, b, x \in \mathcal{A}$ where $\mathcal{R}' = \mathcal{R} \cup \{(a, b)\}$ and $\forall r \in \mathcal{R}, P'_{\mathcal{R}}(r) = P_{\mathcal{R}}(r)$. Assume $\mathbf{S}$ satisfies Directionality and there is no path from b to x. Since there is no path from b to x in $\mathbf{F}$ or $\mathbf{F}'$, there exists no $\mathbf{G} \in \mathtt{I}(\mathbf{F}) \cup \mathtt{I}(\mathbf{F}')$ where there is a path from b to x. In much the same way as we did in the previous bullet, we identify a function f from $\mathtt{I}(\mathbf{F})$ to the power set of $\mathtt{I}(\mathbf{F}')$. Since $\mathbf{S}$ satisfies Directionality and there are no induced graphs with a path from b to x, we get that $\forall \mathbf{G} \in \mathtt{I}(\mathbf{F})$, $\forall \mathbf{G}' \in f(\mathbf{G})$, $\mathtt{Deg}^{\mathbf{S}}_{\mathbf{G}}(x) = \mathtt{Deg}^{\mathbf{S}}_{\mathbf{G}'}(x)$. The result follows from this.
- Let $a \in \mathcal{A}$ s.t. $\mathtt{Att}_{\mathbf{F}}(a)\emptyset$. Assume $\mathbf{S}$ satisfies Maximality. Since $\mathtt{Att}_{\mathbf{F}}(a)\emptyset$, $\forall \mathbf{G} \in \mathtt{I}(\mathbf{F})$, $\mathtt{Att}_{\mathbf{G}}(a) = \emptyset$. $\mathbf{S}$ satisfying Maximality now means that $\forall \mathbf{G} \in \mathtt{I}(\mathbf{F})$, $\mathtt{Deg}^{\mathbf{S}}_{\mathbf{G}}(a) = 1$. Equation 2 combines with Definition 5.14 to give $\forall a \in \mathcal{A}$, $\mathtt{Deg}^{E(\mathbf{S})}_{\mathbf{F}}(a) = P_{\mathcal{A}}(a)$.

□

Together with Resilience, ensuring that attackers are assigned positive strength in all induced graphs, $\mathbf{S}$ satisfying Weakening implies that $E(\mathbf{S})$ satisfies PrAF Weakening.

Theorem 6. *Let* $\mathbf{S}$ *be a gradual semantics. If* $\mathbf{S}$ *satisfies Weakening and Resilience, then* $E(\mathbf{S})$ *satisfies PrAF Weakening.*

Proof. Let $a, b \in \mathcal{A}$ with $b \in \mathtt{Att}_{\mathbf{F}}(a)$ and $\mathtt{Deg}_{\mathbf{F}}^{E(\mathbf{S})}(b) > 0$. Assume $\mathbf{S}$ satisfies Weakening and Resilience. It follows from Equation 2 and Definition 5.14 that if $\exists \mathbf{G}' = \langle \mathcal{A}', \mathcal{R}' \rangle \in \mathtt{I}(\mathbf{F})$ s.t. $a \in \mathcal{A}'$ and $\mathtt{Deg}_{\mathbf{G}'}^{\mathbf{S}}(a) < 1$ then $\mathtt{Deg}_{\mathbf{F}}^{E(\mathbf{S})} < P_{\mathcal{A}}(a)$. Since $b \in \mathtt{Att}_{\mathbf{F}}(a)$, there is at least one $\mathbf{G}' = \langle \mathcal{A}', \mathcal{R}' \rangle \in \mathtt{I}(\mathbf{F})$ with $a, b \in \mathcal{A}$ and $(b, a) \in \mathcal{R}'$. Since $\mathbf{S}$ satisfies Resilience, we know that $\mathtt{Deg}_{\mathbf{G}'}^{\mathbf{S}}(b) > 0$. From here, $\mathbf{S}$ satisfying Weakening means $\mathtt{Deg}_{\mathbf{G}'}^{\mathbf{S}}(a) < 1$. We know see $\mathtt{Deg}_{\mathbf{F}}^{E(\mathbf{S})}(a) < P_{\mathcal{A}}(a)$. □

The next two principles again follow directly from their gradual semantics counterparts.

Theorem 7. *Let* $\mathbf{S}$ *be a gradual semantics.*

- *If* $\mathbf{S}$ *satisfies Weakening Soundness, then* $E(\mathbf{S})$ *satisfies PrAF Weakening Soundness.*
- *If* $\mathbf{S}$ *satisfies Resilience, then* $E(\mathbf{S})$ *satisfies PrAF Resilience.*

Proof. We prove the two parts of Theorem 7 in turn

- Contrapositive; suppose $\mathbf{S}$ satisfies Weakening Soundness, $a \in \mathcal{A}$ and there is no $b \in \mathtt{Att}_{\mathbf{F}}(a)$ with $\mathtt{Deg}_{\mathbf{F}}^{E(\mathbf{S})}(b) > 0$. It now follows through Definition 5.14 that there is no $\mathbf{G}' = \langle \mathcal{A}', \mathcal{R}' \rangle \in \mathtt{I}(\mathbf{F})$ with $a, b \in \mathcal{A}$, $(b, a) \in \mathcal{R}'$, and $\mathtt{Deg}_{\mathbf{G}'}^{\mathbf{S}}(b) > 0$. Since $\mathbf{S}$ satisfies Weakening Soundness, this means $\forall \mathbf{G}' = \langle \mathcal{A}', \mathcal{R}' \rangle \in \mathtt{I}(\mathbf{F})$, s.t. $a \in \mathcal{A}'$, $\mathtt{Deg}_{\mathbf{G}'}^{\mathbf{S}}(a) = 1$. It follows that $\mathtt{Deg}_{\mathbf{F}}^{E(\mathbf{S})}(a) = P_{\mathcal{A}}(a)$.
- Contradiction; Suppose $\mathbf{S}$ satisfies Resilience and $a \in \mathcal{A}$ s.t. not $\mathtt{Deg}_{\mathbf{F}}^{E(\mathbf{S})}(a) > 0$ (i.e. $\mathtt{Deg}_{\mathbf{F}}^{E(\mathbf{S})}(a) = 0$). Through Definition 5.14 this implies $\forall \mathbf{G}' = \langle \mathcal{A}', \mathcal{R}' \rangle \in \mathtt{I}(\mathbf{F})$, s.t. $a \in \mathcal{A}'$, $\mathtt{Deg}_{\mathbf{G}'}^{\mathbf{S}}(a) = 0$, but $\mathbf{S}$ satisfies Resilience so $\forall \mathbf{G}' = \langle \mathcal{A}', \mathcal{R}' \rangle \in \mathtt{I}(\mathbf{F})$, s.t. $a \in \mathcal{A}'$, $\mathtt{Deg}_{\mathbf{G}'}^{\mathbf{S}}(a) > 0$.

□

Mirroring Theorem 2, the following result shows that $E(\mathbf{S})$ satisfies PrAF Symmetry when $\mathbf{S}$ satisfies Anonymity, Independence, and Directionality.

Theorem 8. *Let* $\mathbf{S}$ *be a gradual semantics. If* $\mathbf{S}$ *satisfies Anonymity, Independence and Directionality, then* $E(\mathbf{S})$ *satisfies PrAF Symmetry.*

Proof. The result follows from Theorem 2 and Theorem 5 combined with the first three claims stated in the formulation of this Theorem. □

By adding Neutrality, we also get satisfaction of PrAF Neutrality.

Theorem 9. *Let* $\mathbf{S}$ *be a gradual semantics. If* $\mathbf{S}$ *satisfies Anonymity, Independence, Directionality and Neutrality, then* $E(\mathbf{S})$ *satisfies PrAF Neutrality.*

Proof. Let $a, b, c \in \mathcal{A}$ s.t. $(c, b) \in \mathcal{R}$, $f : \mathbf{F}_{|\mathtt{Str}_{\mathbf{F}}(a,b)\setminus\{c\}} \to \mathbf{F}_{|\mathtt{Str}_{\mathbf{F}}(a,b)\setminus\{c\}}$ s.t. $f(a) = b$, $f(b) = a$ and $f(x) = x$ otherwise, is an automorphism, and $\mathtt{Deg}_{\mathbf{F}}^{E(\mathbf{S})}(c) = 0$. Suppose $\mathbf{S}$ satisfies Anonymity, Independence, Directionality, and Neutrality. Our second remark is that from the fact that f is an automorphism we obtain that a and b have the same attackers in $\mathbf{F}_{|\mathtt{Str}_{\mathbf{F}}(a,b)\setminus\{c\}}$. $\mathtt{Deg}_{\mathbf{F}}^{E(\mathbf{S})}(c) = 0$ implies that $\forall \mathbf{G} \in \mathtt{I}(\mathbf{F})$, if c is an argument of $\mathbf{G}$ then $\mathtt{Deg}_{\mathbf{G}}^{\mathbf{S}}(c) = 0$. Let us first consider the induced graphs of the PrAF $\mathbf{F}_{|\mathtt{Str}_{\mathbf{F}}(a,b)\setminus\{c\}}$. Obviously, the isomorphism f defines a bijection between induced graphs of $\mathbf{F}_{|\mathtt{Str}_{\mathbf{F}}(a,b)\setminus\{c\}}$ which contain a and those that contain b, mapping each argument x to $f(x)$ and each attack (x, y) to $(f(x), f(y))$. For simplicity we denote that bijection with f as well. From Anonymity it follows that for any induced graph G of $\mathbf{F}_{|\mathtt{Str}_{\mathbf{F}}(a,b)\setminus\{c\}}$ that contains a, we have $\mathtt{Deg}_{\mathbf{G}}^{\mathbf{S}}(a) = \mathtt{Deg}_{f(\mathbf{G})}^{\mathbf{S}}(b)$. Moreover, $P^I_{\mathbf{F}_{|\mathtt{Str}_{\mathbf{F}}(a,b)\setminus\{c\}}}(\mathbf{G}) = P^I_{\mathbf{F}_{|\mathtt{Str}_{\mathbf{F}}(a,b)\setminus\{c\}}}(f(\mathbf{G}))$. Consequently, the expected strength semantics will assign the same strength to a and b in $\mathbf{F}_{|\mathtt{Str}_{\mathbf{F}}(a,b)\setminus\{c\}}$.

Now let us consider the induced graph of the original PrAF $\mathbf{F}$. Clearly, each such graph that includes a (respectively b) extends *exactly one* of the induced graphs $\mathbf{G}$ from $\mathbf{F}_{|\mathtt{Str}_{\mathbf{F}}(a,b)\setminus\{c\}}$, if we consider the extensions that do not add to $\mathbf{G}$ novel arguments and attacks from $\mathbf{F}_{|\mathtt{Str}_{\mathbf{F}}(a,b)\setminus\{c\}}$. Moreover, it is easy to show using Definition 2.22 that the sum of probabilities of all those extensions of $\mathbf{G}$ in $\mathbf{F}_{|\mathtt{Str}_{\mathbf{F}}(a,b)\setminus\{c\}}$ is exactly $P^I_{\mathbf{F}_{|\mathtt{Str}_{\mathbf{F}}(a,b)\setminus\{c\}}}(\mathbf{G})$. The proof of this fact is straightforward and deals with independence of probabilities of elements of the framework in the same way as the proof of Proposition 2.23. Therefore, to finish the proof it is sufficient to show that the degree of a (resp. b) in each extension $\mathbf{G}'$ of the graph $\mathbf{G}$ that contains a (resp. b) is exactly $\mathtt{Deg}_{\mathbf{G}}^{\mathbf{S}}(a)$ (resp. $\mathtt{Deg}_{\mathbf{G}}^{\mathbf{S}}(b)$). For a this follows from the fact that the attack structures of a are the same in $\mathbf{G}'$ and $\mathbf{G}$ and the Theorem 2 from [8]. For the argument b we consider two cases. If $\mathbf{G}'$ does not contain the attack from c to b, then we apply the same reasoning as for a. If the attack from c to b is present in the extension, we simply use Neutrality (recall that we know that degree of c is zero in every induced graph). $\square$

By adding Counting and Resilience instead of Neutrality, we get PrAF Counting.

Theorem 10. *Let* $\mathbf{S}$ *be a gradual semantics. If* $\mathbf{S}$ *satisfies Anonymity, Independence, Directionality, Counting and Resilience, then* $E(\mathbf{S})$ *satisfies PrAF Counting.*

Proof. This result is proved in much the same way as the previous one, with the main difference being that Counting is used to show that the attack from c harms b instead of Neutrality being used to show that it is not.

Let $a, b, c \in \mathcal{A}$ s.t. $(c, b) \in \mathcal{R}$, $f : \mathbf{F}_{|\mathtt{Str}_{\mathbf{F}}(a,b)\setminus\{c\}} \to \mathbf{F}_{|\mathtt{Str}_{\mathbf{F}}(a,b)\setminus\{c\}}$ s.t. $f(a) = b$, $f(b) = a$ and $f(x) = x$ otherwise, is an automorphism, and $\mathtt{Deg}_{\mathbf{F}}^{E(\mathbf{S})}(c) > 0$. Suppose $\mathbf{S}$ satisfies Anonymity, Independence, Directionality, Resilience, and Counting. As in the previous proof, we obtain that a and b have the same attackers in $\mathbf{F}_{|\mathtt{Str}_{\mathbf{F}}(a,b)\setminus\{c\}}$ from the fact that f is an automorphism. $\mathbf{S}$ satisfying Resilience implies that $\forall \mathbf{G} \in \mathtt{I}(\mathbf{F})$, if c is an argument of $\mathbf{G}$ then $\mathtt{Deg}_{\mathbf{G}}^{\mathbf{S}}(c) > 0$. As above, let us first consider the induced graphs of the PrAF $\mathbf{F}_{|\mathtt{Str}_{\mathbf{F}}(a,b)\setminus\{c\}}$. Once again, the isomorphism f defines a bijection between induced graphs of $\mathbf{F}_{|\mathtt{Str}_{\mathbf{F}}(a,b)\setminus\{c\}}$ which contain a and those that contain b, mapping each argument x to $f(x)$ and each attack (x, y) to $(f(x), f(y))$. Once more, let us denote this bijection f as well. From Anonymity it follows that for any induced graph $\mathbf{G}$ of $\mathbf{F}_{|\mathtt{Str}_{\mathbf{F}}(a,b)\setminus\{c\}}$ that contains a, we have $\mathtt{Deg}_{\mathbf{G}}^{\mathbf{S}}(a) = \mathtt{Deg}_{f(\mathbf{G})}^{\mathbf{S}}(b)$. Moreover, $P^I_{\mathbf{F}_{|\mathtt{Str}_{\mathbf{F}}(a,b)\setminus\{c\}}}(\mathbf{G}) = P^I_{\mathbf{F}_{|\mathtt{Str}_{\mathbf{F}}(a,b)\setminus\{c\}}}(f(\mathbf{G}))$. Consequently, the expected strength semantics will assign the same strength to a and b in $\mathbf{F}_{|\mathtt{Str}_{\mathbf{F}}(a,b)\setminus\{c\}}$.

Let us now consider the induced graph of the original PrAF $\mathbf{F}$. We again see that each such graph that includes a (respectively b) extends *exactly one* of the induced graphs $\mathbf{G}$ from $\mathbf{F}_{|\mathtt{Str}_{\mathbf{F}}(a,b)\setminus\{c\}}$, if we consider the extensions that do not add to $\mathbf{G}$ novel arguments and attacks from $\mathbf{F}_{|\mathtt{Str}_{\mathbf{F}}(a,b)\setminus\{c\}}$. As in the previous proof, it is again easy to show using Definition 2.22 that the sum of probabilities of all those extensions of $\mathbf{G}$ in $\mathbf{F}_{|\mathtt{Str}_{\mathbf{F}}(a,b)\setminus\{c\}}$ is exactly $P^I_{\mathbf{F}_{|\mathtt{Str}_{\mathbf{F}}(a,b)\setminus\{c\}}}(\mathbf{G})$.

To finish the proof we show that the degree of a in each extension $\mathbf{G}'$ of the graph $\mathbf{G}$ that contains a is exactly $\mathtt{Deg}_{\mathbf{G}}^{\mathbf{S}}(a)$, while the degree of b in each extension $\mathbf{G}'$ of the graph $\mathbf{G}$ that contains b is no higher than $\mathtt{Deg}_{\mathbf{G}}^{\mathbf{S}}(b)$ and there exist such extensions where the degree of b is lower than $\mathtt{Deg}_{\mathbf{G}}^{\mathbf{S}}(b)$. For a this follows from the fact that the attack structures of a are the same in $\mathbf{G}'$ and $\mathbf{G}$ and the Theorem 2 from [8]. For the argument b we consider two cases. First, if $\mathbf{G}'$ does not contain the attack from c to b, then we apply the same reasoning as for a. Second, if the attack from c to b is present in the extension, Counting shows that the degree of b in the extension is strictly lower than $\mathtt{Deg}_{\mathbf{G}}^{\mathbf{S}}(b)$ (recall that the degree of any argument in an induced graph is positive because of Resilience). □

Let us recall that the h-categoriser semantics satisfies all the principles from the Appendix [5]. Together with the theorems 5 through 10, that gives us the following result, which verifies compatibility of all the principles we proposed.

Theorem 11. *$E(\mathtt{Hbs})$ satisfies all the principles proposed in Sec. 5.1.*

Proof. The result follows from Theorems 5 through 10 and `Hbs` satisfying Anonymity, Independence, Directionality, Maximality, Weakening, Resilience, and Weakening Soundness [5]. □

6. Discussion

6.1. **Related Work.** While this thesis is the first instance of gradual semantics being applied to probabilistic argumentation frameworks, the two concepts have met in the literature once before; in [43], Thimm, Cerutti and Rienstra propose a novel gradual semantics, Probabilistic Graded Semantic, for use with argumentation graphs based on the constellations approach to probabilistic argumentation. In their approach, an extension semantics, an inference mode, and a single probability value are used to assign each argument in an ordinary argumentation graph a unique strength value; first, the AG is transformed into a PrAF by assigning each argument in it the same given probability value and assuming all attacks are certain. Then, the acceptability degree of each argument is equated to the probability of acceptability under the chosen extension semantics and inference mode as per Definition 2.24. In their work, the strength assigned to an argument is said to correspond to the argument's resilience to changes in the graph's topology. This resilience also plays a role in the values assigned by the Expected Strength Semantics presented in Section 5. It should be noted, however, that in assigning an equal probability to all arguments—instead of taking the probabilities of arguments and attacks as a constituent of the framework, as is done in this thesis—the information encoded in relative probabilities between elements, a major part of the PrAF's value, cannot be represented. Instead, the single probability value now controls globally whether attackers or defenders are more important to the acceptability degree of an argument. It should also be noted that, in using the probability of acceptability under extension semantics to determine the strength of an argument, Probabilistic Graded Semantics cannot benefit from the rich evaluative scale offered by the gradual semantics used for Expected Strength Semantics. As was done for the semantics proposed in this thesis, the semantics proposed in [43] was studied on the basis of principles, with Thimm et al. opting to use the rationality postulates for ranking semantics proposed by Amgoud and Ben-Naim [2]. The second contribution of [43] is the analysis of the proposed semantics as a consistency measure for argumentation graphs.

Following the original approach of Li, Oren, and Norman [33], in this thesis, we made an assumption of independence between the probabilities of belonging to a graph of different arguments and attacks. This allowed us to calculate the probability of each induced graph using the joint probabilities of independent variables (Def. 2.22). It has been noted in the literature, however, that making this independence assumption is not always appropriate [28]—for example when two attacks attack the same premise in two arguments, meaning that one appearing should increase the likelihood of the other appearing as well—and some works pertaining to the constellations approach to probabilistic argumentation

[28, 29], in fact, do not make this assumption. Instead, these works assume a more general case where a probability distribution over induced graphs is taken as a constituent of the framework. The work in this thesis can be easily adapted to this more general case; instead of calculating the probability of each induced graph with Definition 2.22, these values would be assumed to be given and used directly to calculate the probability of acceptability ($P^{\mathbf{S},t}_{\mathbf{F}}(\mathcal{X})$), the probability of a ranking query holding ($P_{\mathbf{F},\mathbf{S}}(\alpha)$), or the acceptability degree ($\mathtt{Deg}^{E(\mathbf{S})}_{\mathbf{F}}(a)$). The probabilities of individual elements appearing in the graph could be calculated through summation of the probabilities of the induced graph containing said element.

Mantadelis and Bistarelli [35] proposed the Probabilistic Attack Normal Form for PrAFs as a means of avoiding implicit dependence between attacks and the arguments they connect, by removing uncertainty of arguments and replacing it with added uncertain attacks from an added perfect argument. However, this "solution" relies on the fact that, under extension-based semantics, an argument that is attacked may not be accepted. Under the gradual semantics used in this thesis it is possible that attacks merely weaken an argument, meaning this approach could not be easily combined with the work we presented.

Given the variety of possible gradual semantics, it is no surprise that—like this thesis—many works on the topic include (e.g. [36, 32]) or are dedicated to (e.g. [3, 20]) the study of principles that such semantics should or may satisfy. When proposing the set of principles for generalised semantics in Section 5, we generalised an existing set of principles from [5] which, in turn, refine the first set of principles for gradual semantics [3], and extend them to consider weights on arguments. This set of principles from [5] was itself extended to also consider weights on attacks in [7] and, in part, simplified in [8].

As all these sets of principles describe the same intuitions, we found no benefit to using the original principles from [3] over those from [5], even though we used a non-weighted framework in this thesis. Meanwhile, as we will touch on later, using principles that consider weights on arguments puts us in a better position for future work. With regard to the later sets of principles, that presented in [7] offered no advantage either and that from [8] seemed to translate less easily to the probabilistic setting.

Multiple sets of principles describing the same intuitions is not unique to those sets originating in [3]. As noted by Baroni, Rago, and Toni in [11] many of the principles found in the various works on gradual argumentation correspond to the same basic ideas. For example, the Maximality principle—which corresponds to the basic idea that the strength of an argument may only differ from its base score if the argument is attacked—is stated in [3] as it is in Section 2 and slightly changed to include argument weights in [5],[7], and [8], while equivalent properties are used in [32] and [13].

6.2. **Future Work.** Based on the work presented in this thesis, we identify two clear opportunities for future work. First is the approximation of values produced, and second is the addition of weights to the probabilistic argumentation frameworks used.

As is evident from Definition 2.21, the number of induced graphs belonging to a PrAF grows exponentially with regards to the number of arguments and attacks in the PrAF. This clearly makes it impractical to calculate the probability of acceptability ($P_{\mathbf{F}}^{\mathbf{S},t}(\mathcal{X})$), the probability of a ranking query holding ($P_{\mathbf{F},\mathbf{S}}(\alpha)$), or the acceptability degree ($\mathtt{Deg}_{\mathbf{F}}^{E(\mathbf{S})}(a)$) exactly when working with large PrAFs. Li, Oren, and Norman made the same observation about calculating the probability of acceptability under extension semantics ($P_{\mathbf{F}}^{\mathbf{S},i}(\mathcal{X})$) [33] and the complexity of finding extensions and determining the probability of acceptability in the constellations approach to probabilistic argumentation using extension semantics has since been studied extensively (e.g. [24]). Li, Oren, and Norman's solution to this problem is to approximate the value using Monte-Carlo simulation, and we see no reason to believe the same solution may not be applied in our case. Such a simulation has three basic steps: first, inputs are taken randomly from a probability distribution over the domain. In our case, this is the probability distribution over induced graphs ($P_{\mathbf{F}}^{I}$). Second, some computation is performed using the selected inputs. In our case this would be calculating the strengths of arguments in an induced graph and determining acceptability, determining ranking query satisfaction, or multiplying with the induced graph's probability. Finally, the results of repeatedly performing the first two steps is aggregated. Further motivating this approach are the promising experimental results from [8] that indicate that the number of iterations needed to closely approximate the acceptability degree of arguments under h-categoriser and other gradual semantics may be constant w.r.t. the size of the graph and that the time needed for this approximation is low.

In this thesis we followed the constellations approach to probabilistic argumentation and worked with probabilistic argumentation frameworks in which arguments and attacks are given probabilities of belonging to a graph. In doing so, we assumed that any two arguments in an induced graph, disregarding attacks, are equivalent. There are a number of factors that may still set such arguments apart, however. For example, such as in the epistemic approach to probabilistic argumentation, we may assign each argument a different probability of us believing it [27, 31, 42]. Such differences in probability of belief between arguments—like differences in certainty of the argument's reasons [16], number of user votes given to the argument [32], importance of a value promoted by the argument [14], or trustworthiness of the argument's source [38]—may be represented by a weighting on arguments that is carried over from the PrAF into its induced graphs. Seeing how weights can provide meaningful information, even when combined with probabilities of belonging to a graph, extending the work in this thesis to include weights emerges as promising future work. This extension would involve changing to gradual semantics that consider weights on arguments, like those in [8], and further generalising the semantics and corresponding set of principles presented in Section 5 to meaningfully use these weights which we accommodated for already by basing our principles on those from [5].

6.3. **Relevance to Artificial Intelligence.** Argumentation's relevance as a field in Artificial Intelligence is well-established [9, 15]; its applications include decision making [46], reasoning under inconsistency [17], non-monotonic reasoning [34], and providing explainability to artificially intelligent systems [44, 22], and it finds use in various domains including law, medicine, robotics, the semantic web, and security [44], and e-government [9].

Where extension semantics focus on sets of arguments and allow us to answer questions about the acceptability of arguments, gradual semantics provide a richer evaluative scale, focusing on individual arguments and allowing us to answer questions about the strength of arguments. Another difference between these classes of semantics that differentiates the types of problems to which they may naturally be applied is that under extensions semantics an attack generally leads to the immediate rejection of its target, while under gradual semantics an attack may weaken its target without ruling it out altogether.

When applying extension or gradual semantics to argumentation graphs, certainty about which arguments ought to appear in said graph is assumed. Assuming such certainty limits the practical use of the framework, however, as situations in which uncertainty about the topology of the graph arise abound; for example, the language in which arguments are presented may be ambiguous, claims or premises may be left implicit, or arguments may be presented with explicit uncertainty [28]. Probabilistic argumentation frameworks allow us to quantify such uncertainty.

The application of extension semantics to probabilistic argumentation frameworks has been studied extensively, allowing us to answer questions about arguments' acceptability when facing uncertainty about a graph's topology, gradual semantics had not until now. This thesis, in being the first study of gradual semantics in probabilistic argumentation frameworks, provides the first means of answering questions about arguments' strengths when facing this uncertainty.

6.4. **Conclusion.** In this thesis, we conducted the first study of gradual semantics in probabilistic argumentation frameworks. We followed the constellations approach, associating each argument and attack with a probability of it appearing in a graph, allowing us to represent the uncertainty about a graph's topology that may for example result from ambiguities in natural language or implicit premises or claims of argument being presented. We defined the probability of an argument's acceptability with respect to an arbitrary strength threshold when evaluated by a gradual semantics and investigated how the choice of semantics impacts this probability by means of principles taken from the literature. Then, we introduced the notion of ranking queries, which may express complex requirements for a ranking on arguments produced by a gradual semantics, and defined the probability of such a query being satisfied. After this, we proposed a method for calculating the overall strength of each argument in a probabilistic argumentation framework, and we evaluated this method against a set of principles. Finally, we saw that our approach may be easily adapted to the more general case where probabilistic independence between elements of the framework is not assumed and noted two promising avenues for continuing this work into the future.

References

1. Leila Amgoud, *A Replication Study of Semantics in Argumentation.*, Proceedings of the Twenty-Eighth International Joint Conference on Artificial Intelligence, IJCAI 2019, Macao, China, August 10-16, 2019, 2019, pp. 6260–6266.
2. Leila Amgoud and Jonathan Ben-Naim, *Ranking-Based Semantics for Argumentation Frameworks.*, Scalable Uncertainty Management - 7th International Conference, SUM 2013, Washington, DC, USA, September 16-18, 2013. Proceedings, 2013, pp. 134–147.
3. ———, *Axiomatic Foundations of Acceptability Semantics.*, Principles of Knowledge Representation and Reasoning: Proceedings of the Fifteenth International Conference, KR 2016, Cape Town, South Africa, April 25-29, 2016., 2016, pp. 2–11.
4. ———, *Evaluation of arguments in weighted bipolar graphs.*, Int. J. Approx. Reason. **99** (2018), 39–55.
5. Leila Amgoud, Jonathan Ben-Naim, Dragan Doder, and Srdjan Vesic, *Acceptability Semantics for Weighted Argumentation Frameworks.*, Proceedings of the Twenty-Sixth International Joint Conference on Artificial Intelligence, IJCAI 2017, Melbourne, Australia, August 19-25, 2017, 2017, pp. 56–62.
6. Leila Amgoud and Dragan Doder, *Gradual Semantics for Weighted Graphs: An Unifying Approach.*, Principles of Knowledge Representation and Reasoning: Proceedings of the Sixteenth International Conference, KR 2018, Tempe, Arizona, 30 October - 2 November 2018., 2018, pp. 613–614.
7. ———, *Gradual Semantics Accounting for Varied-Strength Attacks.*, Proceedings of the 18th International Conference on Autonomous Agents and MultiAgent Systems, AAMAS '19, Montreal, QC, Canada, May 13-17, 2019, 2019, pp. 1270–1278.
8. Leila Amgoud, Dragan Doder, and Srdjan Vesic, *Evaluation of argument strength in attack graphs: Foundations and semantics.*, Artif. Intell. **302** (2022), 103607.
9. Katie Atkinson, Pietro Baroni, Massimiliano Giacomin, Anthony Hunter, Henry Prakken, Chris Reed, Guillermo Ricardo Simari, Matthias Thimm, and Serena Villata, *Towards Artificial Argumentation.*, AI Mag. **38** (2017), no. 3, 25–36.
10. Pietro Baroni, Dov M. Gabbay, Massimiliano Giacomin, and Leendert van der Torre (eds.), *Handbook of formal argumentation*, College Publications, 2018.
11. Pietro Baroni, Antonio Rago, and Francesca Toni, *How Many Properties Do We Need for Gradual Argumentation?*, Proceedings of the Thirty-Second AAAI Conference on Artificial Intelligence, (AAAI-18), the 30th innovative Applications of Artificial Intelligence (IAAI-18), and the 8th AAAI Symposium on Educational Advances in Artificial Intelligence (EAAI-18), New Orleans, Louisiana, USA, February 2-7, 2018, 2018, pp. 1736–1743.
12. ———, *From fine-grained properties to broad principles for gradual argumentation: A principled spectrum.*, Int. J. Approx. Reason. **105** (2019), 252–286.
13. Pietro Baroni, Marco Romano, Francesca Toni, Marco Aurisicchio, and Giorgio Bertanza, *Automatic evaluation of design alternatives with quantitative argumentation.*, Argument Comput. **6** (2015), no. 1, 24–49.
14. Trevor J. M. Bench-Capon, *Persuasion in Practical Argument Using Value-based Argumentation Frameworks.*, J. Log. Comput. **13** (2003), no. 3, 429–448.
15. Trevor J. M. Bench-Capon and Paul E. Dunne, *Argumentation in artificial intelligence.*, Artif. Intell. **171** (2007), no. 10-15, 619–641.
16. Salem Benferhat, Didier Dubois, and Henri Prade, *Argumentative inference in uncertain and inconsistent knowledge bases.*, UAI '93: Proceedings of the Ninth Annual Conference on Uncertainty in Artificial Intelligence, The Catholic

University of America, Providence, Washington, DC, USA, July 9-11, 1993, 1993, pp. 411–419.

17. Philippe Besnard and Anthony Hunter, *A logic-based theory of deductive arguments.*, Artif. Intell. **128** (2001), no. 1-2, 203–235.
18. Stefano Bistarelli, Theofrastos Mantadelis, Francesco Santini, and Carlo Taticchi, *Probabilistic Argumentation Frameworks with MetaProbLog and ConArg.*, IEEE 30th International Conference on Tools with Artificial Intelligence, ICTAI 2018, 5-7 November 2018, Volos, Greece., 2018, pp. 675–679.
19. Stefano Bistarelli and Francesco Santini, *A Definition of Sceptical Semantics in the Constellations Approach.*, Logic Programming and Nonmonotonic Reasoning - 16th International Conference, LPNMR 2022, Genova, Italy, September 5-9, 2022, Proceedings, 2022, pp. 62–74.
20. Elise Bonzon, Jérôme Delobelle, Sébastien Konieczny, and Nicolas Maudet, *A Comparative Study of Ranking-Based Semantics for Abstract Argumentation.*, Proceedings of the Thirtieth AAAI Conference on Artificial Intelligence, February 12-17, 2016, Phoenix, Arizona, USA., 2016, pp. 914–920.
21. Claudette Cayrol and Marie-Christine Lagasquie-Schiex, *Graduality in Argumentation.*, J. Artif. Intell. Res. **23** (2005), 245–297.
22. Kristijonas Cyras, Antonio Rago, Emanuele Albini, Pietro Baroni, and Francesca Toni, *Argumentative XAI: A Survey.*, Proceedings of the Thirtieth International Joint Conference on Artificial Intelligence, IJCAI 2021, Virtual Event / Montreal, Canada, 19-27 August 2021., 2021, pp. 4392–4399.
23. Phan Minh Dung, *On the Acceptability of Arguments and its Fundamental Role in Nonmonotonic Reasoning, Logic Programming and n-Person Games.*, Artif. Intell. **77** (1995), no. 2, 321–358.
24. Bettina Fazzinga, Sergio Flesca, and Filippo Furfaro, *Complexity of fundamental problems in probabilistic abstract argumentation: Beyond independence.*, Artif. Intell. **268** (2019), 1–29.
25. Dov M. Gabbay and Odinaldo Rodrigues, *Equilibrium States in Numerical Argumentation Networks.*, Logica Universalis **9** (2015), no. 4, 411–473.
26. Anthony Hunter, *Some Foundations for Probabilistic Abstract Argumentation.*, Computational Models of Argument - Proceedings of COMMA 2012, Vienna, Austria, September 10-12, 2012, 2012, pp. 117–128.
27. ______, *A probabilistic approach to modelling uncertain logical arguments.*, Int. J. Approx. Reason. **54** (2013), no. 1, 47–81.
28. ______, *Probabilistic qualification of attack in abstract argumentation.*, Int. J. Approx. Reason. **55** (2014), no. 2, 607–638.
29. Anthony Hunter and Kawsar Noor, *Aggregation of Perspectives Using the Constellations Approach to Probabilistic Argumentation.*, The Thirty-Fourth AAAI Conference on Artificial Intelligence, AAAI 2020, The Thirty-Second Innovative Applications of Artificial Intelligence Conference, IAAI 2020, The Tenth AAAI Symposium on Educational Advances in Artificial Intelligence, EAAI 2020, New York, NY, USA, February 7-12, 2020., 2020, pp. 2846–2853.
30. Anthony Hunter and Matthias Thimm, *Optimization of dialectical outcomes in dialogical argumentation.*, Int. J. Approx. Reason. **78** (2016), 73–102.
31. ______, *Probabilistic Reasoning with Abstract Argumentation Frameworks.*, J. Artif. Intell. Res. **59** (2017), 565–611.
32. João Leite and João G. Martins, *Social Abstract Argumentation.*, IJCAI 2011, Proceedings of the 22nd International Joint Conference on Artificial Intelligence, Barcelona, Catalonia, Spain, July 16-22, 2011, 2011, pp. 2287–2292.

33. Hengfei Li, Nir Oren, and Timothy J. Norman, *Probabilistic Argumentation Frameworks.*, Theorie and Applications of Formal Argumentation - First International Workshop, TAFA 2011. Barcelona, Spain, July 16-17, 2011, Revised Selected Papers, 2011, pp. 1–16.
34. Fangzhen Lin and Yoav Shoham, *Argument Systems: A Uniform Basis for Nonmonotonic Reasoning.*, Proceedings of the 1st International Conference on Principles of Knowledge Representation and Reasoning (KR'89). Toronto, Canada, May 15-18 1989., 1989, pp. 245–255.
35. Theofrastos Mantadelis and Stefano Bistarelli, *Probabilistic abstract argumentation frameworks, a possible world view.*, Int. J. Approx. Reason. **119** (2020), 204–219.
36. Paul-Amaury Matt and Francesca Toni, *A Game-Theoretic Measure of Argument Strength for Abstract Argumentation.*, Logics in Artificial Intelligence, 11th European Conference, JELIA 2008, Dresden, Germany, September 28 - October 1, 2008. Proceedings, 2008, pp. 285–297.
37. Kawsar Noor and Anthony Hunter, *Analysing Product Reviews Using Probabilistic Argumentation.*, Computational Models of Argument - Proceedings of COMMA 2020, Perugia, Italy, September 4-11, 2020., 2020, pp. 295–306.
38. Célia da Costa Pereira, Andrea Tettamanzi, and Serena Villata, *Changing One's Mind: Erase or Rewind?*, IJCAI 2011, Proceedings of the 22nd International Joint Conference on Artificial Intelligence, Barcelona, Catalonia, Spain, July 16-22, 2011, 2011, pp. 164–171.
39. Sylwia Polberg and Anthony Hunter, *Empirical evaluation of abstract argumentation: Supporting the need for bipolar and probabilistic approaches.*, Int. J. Approx. Reason. **93** (2018), 487–543.
40. Nico Potyka, *Extending Modular Semantics for Bipolar Weighted Argumentation.*, Proceedings of the 18th International Conference on Autonomous Agents and MultiAgent Systems, AAMAS '19, Montreal, QC, Canada, May 13-17, 2019, 2019, pp. 1722–1730.
41. ———, *Attractor - A Java Library for Gradual Bipolar Argumentation.*, Computational Models of Argument - Proceedings of COMMA 2022, Cardiff, Wales, UK, 14-16 September 2022, 2022, pp. 369–370.
42. Matthias Thimm, *A Probabilistic Semantics for abstract Argumentation.*, ECAI 2012 - 20th European Conference on Artificial Intelligence. Including Prestigious Applications of Artificial Intelligence (PAIS-2012) System Demonstrations Track, Montpellier, France, August 27-31 , 2012, 2012, pp. 750–755.
43. Matthias Thimm, Federico Cerutti, and Tjitze Rienstra, *Probabilistic Graded Semantics.*, Computational Models of Argument - Proceedings of COMMA 2018, Warsaw, Poland, 12-14 September 2018, 2018, pp. 369–380.
44. Alexandros Vassiliades, Nick Bassiliades, and Theodore Patkos, *Argumentation and explainable artificial intelligence: a survey.*, Knowl. Eng. Rev. **36** (2021), e5.
45. Bruno Yun and Srdjan Vesic, *Gradual Semantics for Weighted Bipolar SETAFs.*, Symbolic and Quantitative Approaches to Reasoning with Uncertainty - 16th European Conference, ECSQARU 2021, Prague, Czech Republic, September 21-24, 2021, Proceedings, 2021, pp. 201–214.
46. Qiaoting Zhong, Xiuyi Fan, Xudong Luo, and Francesca Toni, *An explainable multi-attribute decision model based on argumentation.*, Expert Syst. Appl. **117** (2019), 42–61.