Journal of Artificial Intelligence Research 59 (2017) 565-611 Submitted 11/16; published 08/17

Probabilistic Reasoning with Abstract Argumentation Frameworks

Anthony Hunter ANTHONY.HUNTER@UCL.AC.UK
University College London, UK

Matthias Thimm THIMM@UNI-KOBLENZ.DE
University of Koblenz-Landau, Germany

Abstract

Abstract argumentation offers an appealing way of representing and evaluating arguments and counterarguments. This approach can be enhanced by considering probability assignments on arguments, allowing for a quantitative treatment of formal argumentation. In this paper, we regard the assignment as denoting the degree of belief that an agent has in an argument being acceptable. While there are various interpretations of this, an example is how it could be applied to a deductive argument. Here, the degree of belief that an agent has in an argument being acceptable is a combination of the degree to which it believes the premises, the claim, and the derivation of the claim from the premises. We consider constraints on these probability assignments, inspired by crisp notions from classical abstract argumentation frameworks and discuss the issue of probabilistic reasoning with abstract argumentation frameworks. Moreover, we consider the scenario when assessments on the probabilities of a subset of the arguments are given and the probabilities of the remaining arguments have to be derived, taking both the topology of the argumentation framework and principles of probabilistic reasoning into account. We generalise this scenario by also considering inconsistent assessments, i.e., assessments that contradict the topology of the argumentation framework. Building on approaches to inconsistency measurement, we present a general framework to measure the amount of conflict of these assessments and provide a method for inconsistency-tolerant reasoning.

1. Introduction

Uncertain reasoning usually differentiates between *qualitative* and *quantitative* uncertainty. Approaches to qualitative uncertain reasoning focus on issues such as defeasibility, default-assumptions, and contradictions. These include approaches such as defeasible logics (Nute, 1994), default logics (Reiter, 1980), logic programming (Gelfond & Leone, 2002), and computational models of argumentation (Rahwan & Simari, 2009). One central feature of these approaches is that they provide inferences in a qualitative manner, that is, some statement is either acceptable or not acceptable (in some approaches there is also a third option of "do not know"). We can see this even in abstract argumentation where we can associate each abstract argument with a statement describing the argument. Approaches to quantitative uncertain reasoning, on the other hand, focus on the problem of quantifying the acceptance status of statements and include approaches such as probabilistic reasoning (Pearl, 1988; Paris, 1994), Dempster-Shafer theory (Shafer, 1976), and fuzzy logics (Cintula, Hájek, & Noguera, 2011). As truth and correct decisions are noisy concepts in real-world scenarios these approaches aim at modelling and reasoning with those in a more appropriate manner.

Combining approaches to qualitative and quantitative uncertain reasoning is a natural way to benefit from the advantages of both areas. In this paper, we address the challenge of combining abstract argumentation frameworks (Dung, 1995) with probabilistic reasoning capabilities, which has recently gained some attention in the community of formal argumentation (Hunter, 2016a; Sun & Liao, 2016; Fazzinga, Flesca, & Furfaro, 2016; Bex & Renooij, 2016; Riveret & Governatori, 2016; Hunter & Thimm, 2016b; Hunter, 2016b; Thimm & Gabbay, 2016), see also Section 8.1 for a thorough discussion. An abstract argumentation framework is a directed graph with the nodes being the arguments and edges indicating attack between arguments. Work in this field w.r.t. probabilistic reasoning can be divided (Hunter, 2013) into the *constellations* approach (see e.g. Li, Oren, & Norman, 2011) and the *epistemic* approach (see e.g. Thimm, 2012).

In the constellations approach, uncertainty in the topology of the graph (probabilities on arguments and attacks) is used to make probabilistic assessments on the acceptance of arguments. In the epistemic approach, the topology of the graph is fixed but probabilistic assessments on the acceptance of arguments are evaluated w.r.t. the relations of the arguments in the graph. The core idea of the epistemic approach is that the more likely one is to believe in an argument, the less likely one is to believe in an argument attacking it. The epistemic approach is useful for modelling the beliefs that an opponent might have in the arguments that could be presented, which is useful for example when deciding on the best arguments to present in order to persuade that opponent (Hunter, 2015). The approach is also useful for modelling agents who are unable to directly add or change the argument graph, for instance when considering the beliefs of the audience of a debate.

Here we follow the epistemic approach to probabilistic argumentation and provide a comprehensive account of our framework developed in previous works (Hunter & Thimm, 2014c, 2014b, 2014d, 2016a). To give an overview, our approach to probabilistic reasoning in abstract argumentation frameworks is as follows. We regard assignments of probabilities to arguments as denoting the belief that an agent has that an argument is acceptable. Often, we will just abbreviate our phraseology so that for example instead of talking about *belief in an argument being acceptable*, we will just refer *belief in an argument*. So for a probability function P, and an argument $\mathcal{A}$, $P(\mathcal{A}) > 0.5$ denotes that the argument is believed (to the degree given by $P(\mathcal{A})$), $P(\mathcal{A}) < 0.5$ denotes that the argument is disbelieved (to the degree given by $P(\mathcal{A})$), and $P(\mathcal{A}) = 0.5$ denotes that the argument is neither believed or disbelieved. This approach leads to the notion of an epistemic extension: This is the subset of the arguments in the graph that are believed to be acceptable to some degree (i.e. the arguments such that $P(\mathcal{A}) > 0.5$). Since this is a very general idea, our aim in this paper is to consider various properties (i.e. constraints) that hold for classes of probability functions, and for the resulting epistemic extensions. We structure our presentation on two views as follows:

Standard view on using probability of arguments. In this view, we provide properties for the probability function that ensure that the epistemic extensions coincide with Dung's definitions for extensions. Key properties include *coherence* (if $\mathcal{A}$ attacks $\mathcal{B}$, then $P(\mathcal{A}) \leq 1 - P(\mathcal{B})$) and *foundation* (if $\mathcal{A}$ has no attackers, then $P(\mathcal{A}) = 1$). The advantage of using a probability function instead of Dung's definitions is that we can also specify the degree to which each argument is believed.

Non-standard view on using probability of arguments. In this view, we consider alternative properties for the probability function. This means that the resulting epistemic extensions may not coincide with Dung's definitions for extensions.

The epistemic approach extends abstract argumentation. The notion of an abstract argument graph is very general in that there is no formal restriction on what constitutes an argument or what constitutes an attack. This generality is an advantage in that very diverse kinds of argumentative situations can be modelled using abstract argumentation. At the core of abstract argumentation is the idea that acceptable sets of arguments can be drawn from an argument graph. Various kinds of semantics, starting with Dung's proposals for grounded, preferred, stable, and complete semantics, provide options for determining what constitutes an acceptable set of arguments. Various proposals for extending abstract argumentation, such as value-based argumentation, ranking-based semantics, and weighted argumentation frameworks, introduce extra information to enable the selection of acceptable sets of arguments, and the epistemic approach to argumentation is another proposal in this vein.

As we stated above, for an argument A, P(A) represents the degree of belief that A is acceptable. How we might determine whether an argument is acceptable depends on the kinds of arguments we are dealing with and the kind of application. However, to give an indication, if we are dealing with deductive arguments (i.e. structured arguments where each argument has a set of logical formulae as premises and a logical formula as a claim), then we could specify that a deductive argument is acceptable when its premises are believed, its claim is believed, and the derivation of the claim from the premises is believed. So if there is uncertainty in any of those three dimensions, then this is reflected in the degree of belief that the argument is acceptable. In the work of Hunter (2013), the epistemic approach is applied to classical logic arguments. Each argument has a set of classical logic formulae as premises and a classical logic formula as claim, and the claim is derived from the premises using the classical consequence relation. Uncertainty was captured by a probability distribution over the models of the language, and the probability of an argument was defined as the sum of the probability of the models that satisfy the premises. So the probability of the claim is never less than the probability of the premises, and there is no uncertainty in the derivation of the claim from the premises. In contrast, if we were to use a non-monotonic logic in structured argumentation, then we may have uncertainty in all three dimensions.

To give another example of defining acceptability, we could consider inductive arguments. Here, an inductive argument is a set of examples (as premises) from which a general statement is obtained (as a claim) by a process of induction (i.e. generalisation). Then we could specify that an inductive argument is acceptable when its premises are believed, its claim is believed, and the inductive process by which the claim is obtained from its premises is believed. As there is uncertainty in one or more of these three dimensions, this is reflected in the degree of belief that the argument is acceptable being less than one.

As a third example of defining acceptability, we could consider analogical arguments. In an analogical argument, perceived similarity between two situations is used to claim that some feature of the first situation will hold for the second situation. So we could specify that an analogical argument is acceptable when the first situation does indeed exist and that it has the feature, the two situations are indeed similar, and as a result that the second situation also has the feature. Again there is uncertainty in one or more of these dimensions, and this is reflected in the degree of belief that the argument is acceptable being less than one.

The framework that we present in this paper is appealing theoretically as it provides further insights into semantics for abstract argumentation, and it offers a finer-grained representation of uncertainty in arguments. Perhaps more importantly, our framework for probability functions is appealing practically because we can better handle the following situations.

Modelling an audience judging arguments Consider how a member of the audience of a discussion hears arguments and counterarguments, but is unable (or does not want) to express arguments. Here it is natural to consider how that member of the audience considers which arguments she believes, thereby constructing an epistemic extension. For example, suppose we hear one of our friends saying argument $\mathcal{A}$ = "*John suffers from hay fever, and so a picnic in the hay field will be unpleasant for him*" and we hear another of our friends saying argument $\mathcal{B}$ = "*John has taken a homeopathic medicine for hay fever and therefore he won't suffer from hay fever.*" We are the audience of this discussion, and perhaps for diplomatic motives, we do not want to add any counterarguments. Yet we may wish to judge the arguments that have been presented by our friends. If we regard homeopathic medicine as just water, then we will have high belief in argument A and low belief in argument $\mathcal{B}$ (e.g. $P(\mathcal{A}) = 0.9$ and $P(\mathcal{B}) = 0$), leading to an epistemic extension containing just $\mathcal{A}$. In practice, we can make these judgments when hearing arguments presented in natural language. We can assess the degree to which we believe the premises, the claim, and the derivation of the claim from the premises, and we can then use those evaluations to obtain an overall value for the belief we have in an argument being acceptable. Furthermore, we may choose to adjust this overall value when taking into account other arguments. For instance, if we learn of a counterargument that we assign a high degree of belief in it being acceptable, we may wish to decrease the belief in the original argument being acceptable. We give a larger example in Figure 1.

Modelling an opponent in a dialogue Consider how one agent in an argumentation dialogue will have a model of the other agents in the dialogue. This modelling may include what arguments the other agents believe, and this may be used for a better choice of move in the dialogue. For example, politicians at election time often select arguments to present to a specific group of voters depending on the type of voters. If the voters are business people, then arguments concerning increased expenditure on infrastructure and skills training might be presented since it may be more likely that these would be believed by this audience, whereas if the voters are retired people, then arguments concerning increased expenditure on healthcare might be presented. Furthermore, the politician may take care to not present arguments for which the group of voters might believe counterarguments. So, modelling the beliefs of an opponent in the arguments that could arise in a dialogue may be used by the proponent in a strategy for winning the dialogue. See the work of Hunter (2015) for an application of the epistemic approach to modelling a persuadee in persuasion dialogues.

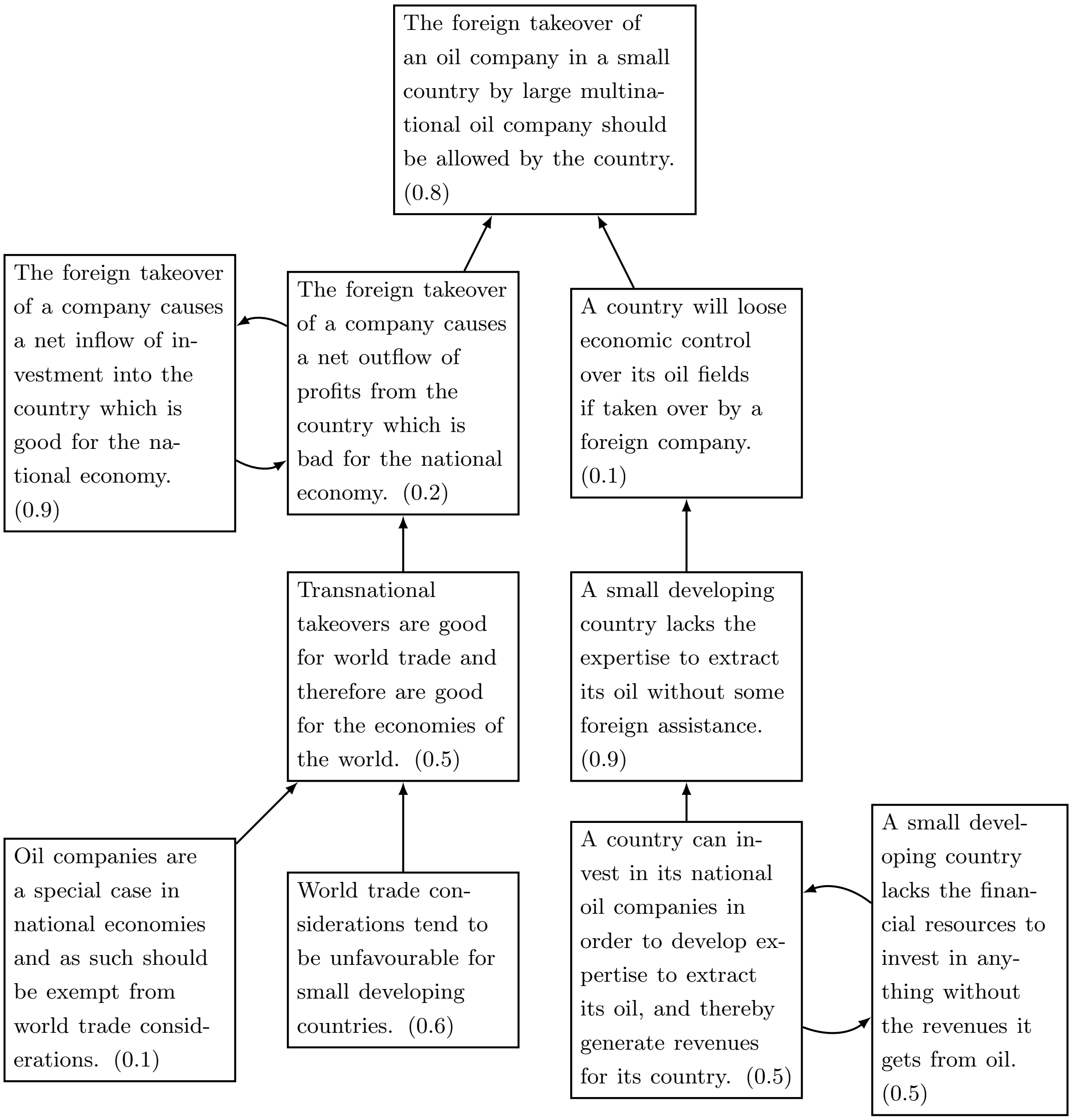

Figure 1: Consider a member of the audience listening to a radio documentary about the takeover of oil production companies in small developing countries by large multinationals. The documentary may be exploring the question of whether such small countries should permit these foreign takeovers. To explore the question, the documentary includes a number of interviews with experts from small developing countries, from multinational oil companies, and from financial institutes. Suppose the member of the audience records ten arguments, and puts them into the argument graph shown. For someone who is reasonably optimistic about multinational oil companies playing a beneficial role in developing countries, the probability value given for each argument (given in brackets in each box) may reflect their belief in the acceptability of each argument.

We extend our framework by also considering the case when probability assessments are either incomplete or contradictory (or both) and the challenge of completing and consolidating them (Hunter & Thimm, 2016a). The central challenge in this investigation is, given probabilistic assessments on arguments that are not meaningful w.r.t. the constraints established in the first part of this paper, how can these probabilities be modified to comply with these conditions? For this purpose and motivated by similar approaches to inconsistency measurement for classical and probabilistic logics (Hunter & Konieczny, 2010; Thimm, 2013; De Bona & Finger, 2015), we present inconsistency measures for evaluating the appropriateness of (partial) probability assessments and a general approach to use those measures to consolidate these assessments.

In summary, the contributions of our work are as follows:

1. We lay out the building blocks for our basic probabilistic framework and investigate the notion of epistemic extensions (Section 3).

2. We discuss several properties for *standard epistemic extensions* and show that these probabilistic concepts coincide with their corresponding concepts from abstract argumentation (Section 4).

3. We introduce non-standard epistemic extensions and a corresponding set of properties as a means to extend the standard view and provide a complete picture of the relationships between our different probabilistic properties (Section 5).

4. We present the concept of partial probability assessments and an approach to complete them by using the principle of maximum entropy (Section 6).

5. We investigate the case of contradictory probability assessments (Section 7), in particular:

    (a) We introduce inconsistency measures for evaluating the significance of a partial probability assessment violating the rationality conditions (Section 7.1).

    (b) We use the inconsistency measures to define consolidation operators for partial probability assessments (Sections 7.2 and 7.3).

Furthermore, we provide some necessary preliminaries in Section 2, discuss related works in Section 8, and conclude with a discussion in Section 9.

This paper builds on previous works (Hunter & Thimm, 2014c, 2014b, 2014d, 2016a) but extends it in several ways and provides a coherent view on the issue of probabilistic reasoning in abstract argumentation. More precisely, the short paper (Hunter & Thimm, 2014d) provides a general overview on the ideas of this paper and preliminary versions of the material in Sections 3–5 are available as a technical report (Hunter & Thimm, 2014c) and a workshop paper (Hunter & Thimm, 2014b). Sections 6 and 7 contains material from the conference paper (Hunter & Thimm, 2016a), extended with formal proofs and more discussion.

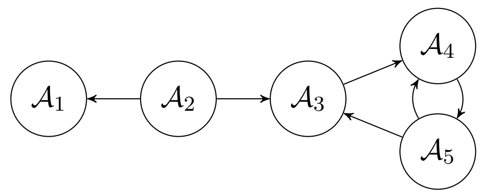

Figure 2: The argumentation framework AF from Example 4

2. Preliminaries

Abstract argumentation frameworks (Dung, 1995) take a very simple view on argumentation as they do not presuppose any internal structure of an argument. Abstract argumentation frameworks only consider the interactions of arguments by means of an attack relation between arguments.

Definition 1. *An* abstract argumentation framework AF *is a tuple* $\mathsf{AF} = (\mathsf{Arg}, \rightarrow)$ *where* Arg *is a set of arguments and* $\rightarrow$ *is a relation* $\rightarrow \subseteq \mathsf{Arg} \times \mathsf{Arg}$.

Let $\mathbb{A}$ denote the set of all abstract argumentation frameworks. For two arguments $\mathcal{A}, \mathcal{B} \in \mathsf{Arg}$ the relation $\mathcal{A} \rightarrow \mathcal{B}$ means that argument $\mathcal{A}$ attacks argument $\mathcal{B}$. For $\mathcal{A} \in \mathsf{Arg}$ define $\mathsf{Att}_{\mathsf{AF}}(\mathcal{A}) = \{\mathcal{B} \mid \mathcal{B} \rightarrow \mathcal{A}\}$. Abstract argumentation frameworks can be concisely represented by directed graphs, where arguments are represented as nodes and edges model the attack relation. Note that we only consider finite argumentation frameworks here, i.e., argumentation frameworks with a finite number of arguments.

Example 1. *Consider the abstract argumentation framework* $\mathsf{AF} = (\mathsf{Arg}, \rightarrow)$ *depicted in Figure 4. Here it is* $\mathsf{Arg} = \{\mathcal{A}_1, \mathcal{A}_2, \mathcal{A}_3, \mathcal{A}_4, \mathcal{A}_5\}$ *and* $\rightarrow\; = \{(\mathcal{A}_2, \mathcal{A}_1), (\mathcal{A}_2, \mathcal{A}_3), (\mathcal{A}_3, \mathcal{A}_4), (\mathcal{A}_4, \mathcal{A}_5), (\mathcal{A}_5, \mathcal{A}_4), \mathcal{A}_5, \mathcal{A}_3)\}$

Semantics are given to abstract argumentation frameworks by means of extensions (Dung, 1995) or labellings (Wu & Caminada, 2010). In this work, we use the latter.

Definition 2. *Let* $\mathsf{AF} = (\mathsf{Arg}, \rightarrow)$ *be an abstract argumentation framework. A* labelling L *for* AF *is a function* $L : \mathsf{Arg} \rightarrow \{\mathsf{in}, \mathsf{out}, \mathsf{undec}\}$.

A labeling L assigns to each argument $\mathcal{A} \in \mathsf{Arg}$ either the value in, meaning that the argument is accepted, out, meaning that the argument is not accepted, or undec, meaning that the status of the argument is undecided. Let $\mathsf{in}(L) = \{\mathcal{A} \mid L(\mathcal{A}) = \mathsf{in}\}$ and $\mathsf{out}(L)$ resp. $\mathsf{undec}(L)$ be defined analogously. The set $\mathsf{in}(L)$ for a labelling L is also called *extension* (Dung, 1995). A labelling L is called *conflict-free* if for no $\mathcal{A}, \mathcal{B} \in \mathsf{in}(L)$ we have that $\mathcal{A} \rightarrow \mathcal{B}$.

Arguably, the most important property of a semantics is its admissibility. A conflict-free labelling L is called *admissible* if and only if for all arguments $\mathcal{A} \in \mathsf{Arg}$

1. if $L(\mathcal{A}) = \mathsf{out}$ then there is $\mathcal{B} \in \mathsf{Arg}$ with $L(\mathcal{B}) = \mathsf{in}$ and $\mathcal{B} \rightarrow \mathcal{A}$, and

2. if $L(\mathcal{A}) = \mathsf{in}$ then $L(\mathcal{B}) = \mathsf{out}$ for all $\mathcal{B} \in \mathsf{Arg}$ with $\mathcal{B} \rightarrow \mathcal{A}$,

and it is called *complete* if, additionally, it satisfies

3. if $L(\mathcal{A}) = \mathsf{undec}$ then there is no $\mathcal{B} \in \mathsf{Arg}$ with $\mathcal{B} \rightarrow \mathcal{A}$ and $L(\mathcal{B}) = \mathsf{in}$ and there is a $\mathcal{B}' \in \mathsf{Arg}$ with $\mathcal{B}' \rightarrow \mathcal{A}$ and $L(\mathcal{B}') \neq \mathsf{out}$.

The intuition behind admissibility is that an argument can only be accepted if there are no attackers that are accepted and if an argument is not accepted then there has to be some reasonable grounds. The idea behind the completeness property is that the status of an argument is only undec if it cannot be classified as in or out. Different types of classical semantics (Dung, 1995; Caminada, 2006; Baroni, Caminada, & Giacomin, 2011) can be phrased by imposing further constraints. Let $\mathsf{AF} = (\mathsf{Arg}, \rightarrow)$ be an abstract argumentation framework and $L : \mathsf{Arg} \rightarrow \{\mathsf{in}, \mathsf{out}, \mathsf{undec}\}$ a complete labelling. Then

- L is *grounded* if and only if $\mathsf{in}(L)$ is minimal,
- L is *preferred* if and only if $\mathsf{in}(L)$ is maximal,
- L is *stable* if and only if $\mathsf{undec}(L) = \emptyset$, and
- L is *semi-stable* if and only if $\mathsf{undec}(L)$ is minimal.

All statements on minimality/maximality are meant to be with respect to set inclusion. Note that a grounded labelling is uniquely determined and always exists (Dung, 1995).

Example 2. *Consider again the argumentation framework* AF *in Figure 4. The labeling L_1 defined via*

$$L_1(\mathcal{A}_1) = \mathsf{out} \quad L_1(\mathcal{A}_2) = \mathsf{in} \quad L_1(\mathcal{A}_3) = \mathsf{out} \quad L_1(\mathcal{A}_4) = \mathsf{undec} \quad L_1(\mathcal{A}_5) = \mathsf{undec}$$

is complete and grounded. The labeling L_2 defined via

$$L_2(\mathcal{A}_1) = \mathsf{out} \quad L_2(\mathcal{A}_2) = \mathsf{in} \quad L_2(\mathcal{A}_3) = \mathsf{out} \quad L_2(\mathcal{A}_4) = \mathsf{in} \quad L_2(\mathcal{A}_5) = \mathsf{out}$$

is complete, preferred, stable, and semi-stable.

Abstract argumentation frameworks are arguably the most investigated formalism for formal argumentation. However, there are also formalisms for structured argumentation, such as deductive argumentation (Besnard & Hunter, 2008), ASPIC$^+$ (Modgil & Prakken, 2014), ABA (Toni, 2014), and defeasible logic programming (Garcia & Simari, 2004). In structured argumentation, arguments are a set of (e.g. propositional) formulas (the support of an argument) that derive a certain conclusion (the claim of an argument). The attack relation between arguments is then derived from logical inconsistency.

3. A Probabilistic Framework for Abstract Argumentation

We now go beyond classical three-valued semantics of abstract argumentation and turn to probabilistic interpretations of the status of arguments. Let $2^{\mathcal{X}}$ denote the power set of a set $\mathcal{X}$. For our purposes we define a probability function as follows.

Definition 3. *Let $\mathcal{X}$ be some finite set. A* probability function *P on $\mathcal{X}$ is a function $P : 2^{\mathcal{X}} \to [0,1]$ that satisfies*

$$\sum_{X \subseteq \mathcal{X}} P(X) = 1$$

Here, a probability function is a function on the set of subsets of some (finite) set which is *normalized*, i.e., the sum of the probabilities of all subsets is one. Let $\mathcal{P}$ be the set of all probability functions.

We use the concept of *subjective probability* (Paris, 1994) for interpreting probabilities. There, a probability $P(X)$ for some $X \subseteq \mathcal{X}$ denotes the *degree of belief* we put into X. Then a probability function P can be seen as an epistemic state of some agent that has uncertain beliefs with respect to $\mathcal{X}$. In probabilistic reasoning (Pearl, 1988; Paris, 1994), this interpretation of probability is widely used to model uncertain knowledge representation and reasoning.

In the following, we consider probability functions on sets of arguments of an abstract argumentation framework. Let $\mathsf{AF} = (\mathsf{Arg}, \to)$ be some fixed abstract argumentation framework and let $\mathcal{P}(\mathsf{AF})$ be the set of probability functions of the form $P : 2^{\mathsf{Arg}} \to [0,1]$. For $P \in \mathcal{P}(\mathsf{AF})$ and $\mathcal{A} \in \mathsf{Arg}$ we abbreviate

$$P(\mathcal{A}) = \sum_{\mathcal{A} \in E \subseteq \mathsf{Arg}} P(E) \quad .$$

Given some probability function P, the probability $P(\mathcal{A})$ represents the degree of belief that $\mathcal{A}$ is *acceptable* wrt. P. In order to bridge the gap between probability functions and labelings, consider the following definition from Hunter (2013).

Definition 4. *Let $\mathsf{AF} = (\mathsf{Arg}, \to)$ be an abstract argumentation framework and $P : 2^{\mathsf{Arg}} \to [0,1]$ a probability function on Arg. The labelling $L_P : \mathsf{Arg} \to \{\mathsf{in}, \mathsf{out}, \mathsf{undec}\}$ defined via the following constraints is called the* epistemic labelling *of P:*

- $L_P(\mathcal{A}) = \mathsf{in}$ *iff* $P(\mathcal{A}) > 0.5$
- $L_P(\mathcal{A}) = \mathsf{out}$ *iff* $P(\mathcal{A}) < 0.5$
- $L_P(\mathcal{A}) = \mathsf{undec}$ *iff* $P(\mathcal{A}) = 0.5$

In other words, an argument $\mathcal{A}$ is labelled in in L_P when it is believed to some degree (which we identify as $P(\mathcal{A}) > 0.5$), it is labelled out when it is disbelieved to some degree (which we identify as $P(\mathcal{A}) < 0.5$), and it is labelled undec when it is neither believed nor disbelieved (which we identify as $P(\mathcal{A}) = 0.5$). Furthermore, the *epistemic extension* of P is the set of arguments that are labelled in by the epistemic labelling, i.e., X is an epistemic extension iff $X = \mathsf{in}(L_P)$. We say that a labelling L and a probability function P are *congruent*, denoted by $L \sim P$, if for all $\mathcal{A} \in \mathsf{Arg}$ we have $L(\mathcal{A}) = \mathsf{in} \Leftrightarrow P(\mathcal{A}) = 1$, $L(\mathcal{A}) = \mathsf{out} \Leftrightarrow P(\mathcal{A}) = 0$, and $L(\mathcal{A}) = \mathsf{undec} \Leftrightarrow P(\mathcal{A}) = 0.5$. Note that if $L \sim P$ then $L = L_P$, i.e., if a labelling L and a probability function P are congruent then L is also the epistemic labelling of P.

An epistemic labelling can be used to give either a standard semantics (as we will investigate in Section 4) or a non-standard semantics (as we will investigate in Section 5).

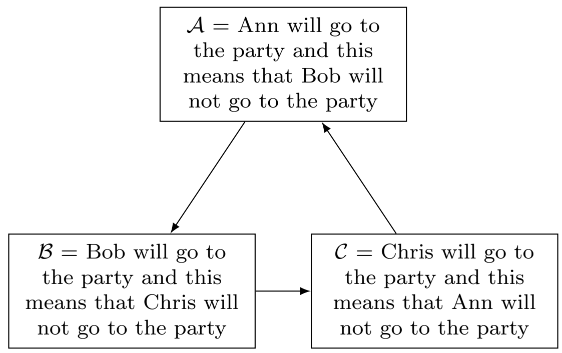

Figure 3: Example of three arguments in a simple cycle.

Example 3. *To further illustrate epistemic labelings and extensions, consider the graph given in Figure 3. Here, we may believe that, say, $\mathcal{A}$ is valid and that $\mathcal{B}$ and $\mathcal{C}$ are not valid. In which case, with this extra epistemic information about the arguments, we can resolve the conflict and so take the set $\{\mathcal{A}\}$ as the "epistemic" extension. In contrast, there is only one admissible set which is the empty set. So by having this extra epistemic information, we get a more informed extension (in the sense that it has harnessed the extra information in a sensible way).*

In general, we want epistemic extensions to allow us to better model the audience of argumentation. Consider, for example, when a member of the audience of a TV debate listens to the debate at home, she can produce the abstract argument graph based on the arguments and counterarguments exchanged. Then she can identify a probability function to represent the belief she has in each of the arguments. So she may disbelieve some of the arguments based on what she knows about the topic. Furthermore, she may disbelieve some of the arguments that are unattacked. As an extreme, she is at liberty of completely disbelieving all of the arguments (so to assign probability 0 to all of them). If we want to model audiences, where the audience either does not want to or is unable to add counterarguments to an argument graph being constructed in some form of argumentation, we need to take the beliefs of the audience into account, and we need to consider which arguments they believe or disbelieve.

4. Standard Epistemic Extensions

We now consider some constraints on the probability function which may take different aspects of the structure of the argument graph into account. We will show how these constraints are consistent with Dung's notions of admissibility.

For the remainder of this paper let $\mathsf{AF} = (\mathsf{Arg}, \rightarrow)$ be an abstract argumentation framework and $P : 2^{\mathsf{Arg}} \rightarrow [0,1]$ a probability function. Consider the following properties (note that COH is from Hunter, 2013 and JUS is from Thimm, 2012):

COH P is *coherent* wrt. AF if for every $\mathcal{A}, \mathcal{B} \in \mathsf{Arg}$, if $\mathcal{A} \rightarrow \mathcal{B}$ then $P(\mathcal{A}) \leq 1 - P(\mathcal{B})$.

SFOU P is *semi-founded* wrt. AF if $P(\mathcal{A}) \geq 0.5$ for every $\mathcal{A} \in \mathsf{Arg}$ with $\mathsf{Att}_{\mathsf{AF}}(\mathcal{A}) = \emptyset$.

FOU P is *founded* wrt. AF if $P(\mathcal{A}) = 1$ for every $\mathcal{A} \in \mathsf{Arg}$ with $\mathsf{Att}_{\mathsf{AF}}(\mathcal{A}) = \emptyset$.

SOPT P is *semi-optimistic* wrt. AF if $P(\mathcal{A}) \geq 1 - \sum_{\mathcal{B} \in \mathsf{Att}_{\mathsf{AF}}(\mathcal{A})} P(\mathcal{B})$ for every $\mathcal{A} \in \mathsf{Arg}$ with $\mathsf{Att}_{\mathsf{AF}}(\mathcal{A}) \neq \emptyset$.

OPT P is *optimistic* wrt. AF if $P(\mathcal{A}) \geq 1 - \sum_{\mathcal{B} \in \mathsf{Att}_{\mathsf{AF}}(\mathcal{A})} P(\mathcal{B})$ for every $\mathcal{A} \in \mathsf{Arg}$.

JUS P is *justifiable* wrt. AF if P is coherent and optimistic.

TER P is *ternary* wrt. AF if $P(\mathcal{A}) \in \{0, 0.5, 1\}$ for every $\mathcal{A} \in \mathsf{Arg}$.

The intuition behind these properties is as follows. COH ensures that if argument $\mathcal{A}$ attacks argument $\mathcal{B}$, then the degree to which $\mathcal{A}$ is believed caps the degree to which $\mathcal{B}$ can be believed. SFOU ensures that if an argument is not attacked, then the argument is not disbelieved (i.e. $P(\mathcal{A}) \geq 0.5$). FOU ensures that if an argument is not attacked, then the argument is believed without doubt (i.e. $P(\mathcal{A}) = 1$). SOPT ensures that the belief in $\mathcal{A}$ is bounded from below if the belief in its attackers is not high. OPT ensures that if an argument is not attacked, then the argument is believed without doubt (i.e. $P(\mathcal{A}) = 1$) and that the belief in $\mathcal{A}$ is bounded from below if the belief in its attackers is not high. In particular, the term $\sum_{\mathcal{B} \in \mathsf{Att}_{\mathsf{AF}}(\mathcal{A})} P(\mathcal{B})$ can be interpreted as the upper bound on the probability that *some* attacker of $\mathcal{A}$ is acceptable. Note that this condition enforces $P(\mathcal{A}) = 1$ if $P(\mathcal{B}) = 0$ for all attackers $\mathcal{B}$ of $\mathcal{A}$ and thus models a probabilistic version of (part of) the admissibility condition for ordinary abstract argumentation frameworks. JUS combines COH and OPT to provide bounds on the belief in an argument based on the belief in its attackers and attackees, and TER ensures that the probability assignment is a three-valued assignment.

Example 4. *Consider the abstract argumentation framework* $\mathsf{AF} = (\mathsf{Arg}, \rightarrow)$ *depicted in Fig. 4 and the probability functions depicted in Table 1 (note that these functions are only partially defined by giving the probabilities of arguments). The following observations can be made:*

- P_1 *is semi-founded, founded, but neither coherent, optimistic, semi-optimistic, ternary, nor justifiable,*
- P_2 *is coherent and semi-optimistic, but neither semi-founded, founded, optimistic, ternary, nor justifiable,*
- P_3 *is coherent, semi-optimistic, semi-founded, founded, optimistic, and justifiable, but not ternary,*
- P_4 *is semi-founded, founded, optimistic, and semi-optimistic, but neither coherent, justifiable, nor ternary, and*
- P_5 *is coherent, semi-founded, semi-optimistic, and ternary but neither optimistic, justifiable, nor founded.*

Example 5. *Consider the graph given in Figure 3. Suppose, we have an assignment* $P(\mathcal{A}) = 0.5$*,* $P(\mathcal{B}) = 0.5$*, and* $P(\mathcal{C}) = 0.5$*. This assignment, which makes no commitment to believe*

	$\mathcal{A}_1$	$\mathcal{A}_2$	$\mathcal{A}_3$	$\mathcal{A}_4$	$\mathcal{A}_5$	$\mathcal{A}_6$
P_1	0.2	0.7	0.6	0.3	0.6	1
P_2	0.7	0.3	0.5	0.5	0.2	0.4
P_3	0.7	0.3	0.7	0.3	0	1
P_4	0.7	0.8	0.9	0.8	0.7	1
P_5	0.5	0.5	0.5	0.5	0.5	0.5

Table 1: Some probability functions for Example 4

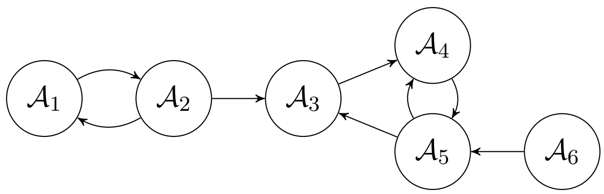

Figure 4: A simple argumentation framework

or disbelieve any of the arguments, satisfies the coherent, semi-founded, founded, semi-optimistic, optimistic, justifiable, and ternary properties. Now consider the assignment $P'(\mathcal{A}) = 0.8$, $P'(\mathcal{B}) = 0.2$, *and* $P'(\mathcal{C}) = 0.2$. *This assignment satisfies the coherent, semi-founded, and founded properties, but it does not satisfy the semi-optimistic, optimistic, justifiable, or ternary properties. Failure of the semi-optimistic, optimistic, and justifiable properties comes from the* $P'(\mathcal{C})$ *being too low given* $P'(\mathcal{B})$ *being so low.*

Example 6. *Consider Figure 1. Here, there is the unattacked argument "Oil companies are a special case in national economies and as such should be exempt from world trade considerations" with a probability assignment of* 0.1. *So the member of the audience has not believed this argument. Possibly they may have an argument against it or they may simply disbelieve it without having a reason against. In any case, this assignment violates the founded and semi-founded properties. So even though these are important properties for reflecting abstract argumentation theory, it may also be appropriate to suspend them in practice for applications such as this.*

Recall that $\mathcal{P}(\mathsf{AF})$ is the set of all probability functions on Arg. Let $\mathcal{P}_t(\mathsf{AF})$ be the set of all t-probability functions with $t \in \{\mathsf{COH},\mathsf{SFOU},\mathsf{FOU},\mathsf{OPT},\mathsf{SOPT},\mathsf{JUS},\mathsf{TER}\}$. For $T \subseteq \{\mathsf{COH},\mathsf{SFOU},\mathsf{FOU},\mathsf{OPT},\mathsf{SOPT},\mathsf{JUS},\mathsf{TER}\}$ we abbreviate

$$\mathcal{P}_T(\mathsf{AF}) = \bigcap_{t \in T} \mathcal{P}_t(\mathsf{AF})$$

We obtain the following relationships between the different classes of probability functions.

Proposition 1. *Let* $\mathsf{AF} = (\mathsf{Arg}, \to)$ *be an abstract argumentation framework.*

1. $\emptyset \subsetneq \mathcal{P}_{\mathsf{JUS}}(\mathsf{AF}) \subsetneq \mathcal{P}_{\mathsf{COH}}(\mathsf{AF}) \subsetneq \mathcal{P}(\mathsf{AF})$

2. $\mathcal{P}_{\mathsf{OPT}}(\mathsf{AF}) = \mathcal{P}_{\{\mathsf{SOPT},\mathsf{FOU}\}}(\mathsf{AF})$.

3. $\mathcal{P}_{\mathsf{FOU}}(\mathsf{AF}) \subsetneq \mathcal{P}_{\mathsf{SFOU}}(\mathsf{AF})$.

4. $\emptyset \subsetneq \mathcal{P}_{\mathsf{TER}}(\mathsf{AF}) \subsetneq \mathcal{P}(\mathsf{AF})$.

For the proof of item 1.) of the above proposition see the works by Thimm (2012) and Hunter (2013). The remaining relationships follow directly from these definitions.

For all probability functions P such that L_P is admissible in the classical sense, we have that P assigns some degree of belief to each argument that is unattacked, thereby P satisfies the SFOU constraint.

Proposition 2. *For all probability functions P, if L_P is admissible then $P \in \mathcal{P}_{\mathsf{SFOU}}(\mathsf{AF})$.*

Proof. Assume L_P is admissible. Therefore, if $L_P(\mathcal{A}) = \mathsf{out}$, then there is an argument $\mathcal{B}$ such that $\mathcal{B} \to \mathcal{A}$ and $L_P(\mathcal{B}) = \mathsf{in}$. Therefore, if $\mathsf{Att}_{\mathsf{AF}}(\mathcal{A}) = \emptyset$, then $L_P(\mathcal{A}) \neq \mathsf{out}$. Therefore, if $\mathsf{Att}_{\mathsf{AF}}(\mathcal{A}) = \emptyset$, then $P(\mathcal{A}) \geq 0.5$. Therefore, $P \in \mathcal{P}_{\mathsf{SFOU}}(\mathsf{AF})$. □

We can further constrain a probability assignment so that the epistemic labelling straightforwardly captures the standard semantics (i.e. Dung's semantics). By setting the probability function appropriately, its epistemic labelling corresponds to grounded, complete, stable, preferred, or semi-stable labellings. All we require is a three-valued probability function that simulates each complete labelling function. For this, we provide the following definition that provides the counter-part in our framework for a complete labelling.

Definition 5. *Let $\mathsf{AF} = (\mathsf{Arg}, \to)$ be an argumentation framework. Then a* complete probability function *$P \in \mathcal{P}(\mathsf{AF})$ for AF is a probability function P such that for every $\mathcal{A} \in \mathsf{Arg}$ the following conditions hold:*

1. *$P \in \mathcal{P}_{\mathsf{TER}}(\mathsf{AF})$;*

2. *if $P(\mathcal{A}) = 1$ then $P(\mathcal{B}) = 0$ for all $\mathcal{B} \in \mathsf{Arg}$ with $\mathcal{B} \to \mathcal{A}$;*

3. *if $P(\mathcal{B}) = 0$ for all $\mathcal{B}$ with $\mathcal{B} \to \mathcal{A}$ then $P(\mathcal{A}) = 1$;*

4. *if $P(\mathcal{A}) = 0$ then there is $\mathcal{B} \in \mathsf{Arg}$ with $P(\mathcal{B}) = 1$ and $\mathcal{B} \to \mathcal{A}$;*

5. *if $P(\mathcal{B}) = 1$ for some $\mathcal{B}$ with $\mathcal{B} \to \mathcal{A}$ then $P(\mathcal{A}) = 0$.*

Note that the above definition straightforwardly follows the definition of completeness for classical semantics. Therefore, we have that P is a complete probability function if and only if there is a complete labeling L and $P \sim L$.

In the same way that Caminada & Gabbay (2009) showed that different semantics can be obtained by imposing further restrictions on the choice of labelling, we can obtain the different semantics by imposing further restrictions on the choice of complete probability function. These constraints, as shown in the following result, involve minimizing or maximizing particular assignments. So for instance, if the assignment of 1 to arguments is maximized, then a preferred labelling is obtained.

Restriction on a complete probability function P	Classical semantics
No restriction	complete extensions
No arguments $\mathcal{A}$ such that $P(\mathcal{A}) = 0.5$	stable
Maximal no. of $\mathcal{A}$ such that $P(\mathcal{A}) = 1$	preferred
Maximal no. of $\mathcal{A}$ such that $P(\mathcal{A}) = 0$	preferred
Maximal no. of $\mathcal{A}$ such that $P(\mathcal{A}) = 0.5$	grounded
Minimal no. of $\mathcal{A}$ such that $P(\mathcal{A}) = 1$	grounded
Minimal no. of $\mathcal{A}$ such that $P(\mathcal{A}) = 0$	grounded
Minimal no. of $\mathcal{A}$ such that $P(\mathcal{A}) = 0.5$	semi-stable

Table 2: Correspondences between probabilistic and classical semantics

Proposition 3. *Let* $\mathsf{AF} = (\mathsf{Arg}, \rightarrow)$ *be an abstract argumentation framework and* $P \in \mathcal{P}(\mathsf{AF})$*. If* P *is a complete probability function for* AF *and the restriction specified in Table 2 holds for* P*, then the corresponding type of epistemic labelling is obtained.*

Proof. Let L and P be congruent, i.e., $L \sim P$. So L is a complete labelling iff P is a complete probability assignment. Therefore, each restriction in Section 2 holds for L iff the corresponding restriction in Table 2 holds for P. For instance, "Maximal number of arguments $\mathcal{A}$ such that $L(\mathcal{A}) = \mathsf{undec}$" holds iff "Maximal number of arguments $\mathcal{A}$ such that $P(\mathcal{A}) = 0.5$" holds. Therefore, the corresponding type of extension for the restriction on L (as listed in Section 2 and proven to hold in Caminada & Gabbay, 2009), also hold for the equivalent restriction on P in Table 2. □

For an argumentation framework AF we can identify specific probability functions in $P \in \mathcal{P}_{\mathsf{JUS}}(\mathsf{AF})$ that are congruent with admissible labellings, grounded labellings, or stable labellings, for AF as follows.

Proposition 4. *(Thimm, 2012) Let* $\mathsf{AF} = (\mathsf{Arg}, \rightarrow)$ *be an abstract argumentation framework.*

1. *For every admissible* L *there is* $P \in \mathcal{P}_{\mathsf{JUS}}(\mathsf{AF})$ *with* $L \sim P$.
2. *Let* L *be the grounded labelling and let*[1] $P = \arg\max_{Q \in \mathcal{P}_{\mathsf{JUS}}(\mathsf{Arg})} H(Q)$*. Then* $L \sim P$.
3. *Let stable labellings exist for* AF *and let* L *be a stable labelling. Then there is* $P \in \arg\min_{Q \in \mathcal{P}_{\mathsf{JUS}}(\mathsf{Arg})} H(Q)$ *with* $L \sim P$.

So Proposition 3 and Proposition 4 provide two ways to identify probability functions that capture specific types of labellings. Each of these results show that standard notions of classical semantics (i.e. admissibility and the definitions for different kinds of labelling such as grounded labellings, stable labellings, etc.) can be captured using probability functions.

The next result shows that using probability functions to capture labellings gives a finer-grained formalization of classical semantics.

1. Define the *entropy* $H(P)$ of P as $H(P) = -\sum_{E \subseteq \mathsf{Arg}} P(E) \log P(E)$

Proposition 5. *For each complete labelling L, if there is an argument $\mathcal{A}$ such that $L(\mathcal{A}) \neq$ undec, then there are infinitely many probability functions P such that $L_P = L$.*

Proof. Let L be a complete labeling such that there is an argument $\mathcal{A}$ with $L(\mathcal{A}) \neq \mathsf{undec}$. Without loss of generality assume that $L(\mathcal{A}) = \mathsf{in}$. Then every probability function with $P(\mathcal{B}) = 0.5$ iff $L(\mathcal{B}) = \mathsf{undec}$, $P(\mathcal{B}) = 0$ iff $L(\mathcal{B}) = \mathsf{out}$, $P(\mathcal{B}) = 1$ iff $L(\mathcal{B}) = \mathsf{in}$ and $\mathcal{B} \neq \mathcal{A}$, and $P(\mathcal{A}) \in (0.5, 1]$ yields $L_P = L$. $\square$

Obviously, for every probability function P, there is by definition exactly one epistemic labelling L_P. This means that using a probability function to identify which arguments are in, undec, or out, subsumes using labels. Furthermore, the probability function captures more information about the arguments. The granularity can differentiate between for example a situation where $\mathcal{A}$ is believed (i.e. it is in) with certainty by $P(\mathcal{A}) = 1$ from a situation where $\mathcal{A}$ is only just believed (i.e. it is only just in) for example by $P(\mathcal{A}) = 0.51$. Similarly, we can differentiate a situation where an attack by $\mathcal{B}$ on $\mathcal{A}$ is undoubted when $P(\mathcal{B}) = 1$ and $P(\mathcal{A}) = 0$ from a situation where an attack by $\mathcal{B}$ on $\mathcal{A}$ is very much doubted when for example $P(\mathcal{B}) = 0.55$ and $P(\mathcal{A}) = 0.45$.

In conclusion, we have shown how axioms can be used to constrain the probability function, and thereby constrain the epistemic labelings and the epistemic extensions. This allows us to subsume Dung's notions of extensions as epistemic extensions. Furthermore, we get a finer-grained representation of the labelling of arguments by representing the belief in each of the arguments.

5. Non-standard Epistemic Extensions

Before exploring the notion of non-standard epistemic extensions, we will augment the set of properties we introduced in the previous section with the following properties. Let $\mathsf{AF} = (\mathsf{Arg}, \rightarrow)$ be an abstract argumentation framework and $P : 2^{\mathsf{Arg}} \rightarrow [0, 1]$.

RAT P is *rational* wrt. AF if for every $\mathcal{A}, \mathcal{B} \in \mathsf{Arg}$, if $\mathcal{A} \rightarrow \mathcal{B}$ then $P(\mathcal{A}) > 0.5$ implies $P(\mathcal{B}) \leq 0.5$.

NEU P is *neutral* wrt. AF if $P(\mathcal{A}) = 0.5$ for every $\mathcal{A} \in \mathsf{Arg}$.

INV P is *involutary* wrt. AF if for every $\mathcal{A}, \mathcal{B} \in \mathsf{Arg}$, if $\mathcal{A} \rightarrow \mathcal{B}$, then $P(\mathcal{A}) = 1 - P(\mathcal{B})$.

MAX P is *maximal* wrt. AF if $P(\mathcal{A}) = 1$ for every $\mathcal{A} \in \mathsf{Arg}$.

MIN P is *minimal* wrt. AF if $P(\mathcal{A}) = 0$ for every $\mathcal{A} \in \mathsf{Arg}$.

The intuition behind these properties is as follows. RAT ensures that if argument $\mathcal{A}$ attacks argument $\mathcal{B}$, and $\mathcal{A}$ is believed (i.e. $P(\mathcal{A}) > 0.5$), then $\mathcal{B}$ is not believed (i.e. $P(\mathcal{B}) \leq 0.5$); NEU ensures that all arguments are neither believed nor disbelieved (i.e. $P(\mathcal{A}) = 0.5$ for all arguments); INV ensures that if argument $\mathcal{A}$ attacks argument $\mathcal{B}$, then the belief in $\mathcal{A}$ is the inverse of the belief in $\mathcal{B}$; MAX ensures that all arguments are completely believed; and MIN ensures that all arguments are completely disbelieved.

Example 7. *We continue Example 4, the abstract argumentation framework* $\mathsf{AF} = (\mathsf{Arg}, \rightarrow)$ *depicted in Fig. 4, and the probability functions depicted in Table 1. The following observations can be made:*

1. *P_2 and P_3 are rational but neither neutral, involutary, maximal, nor minimal,*

2. *P_1 and P_4 are neither rational, neutral, involutary, maximal, nor minimal, and*

3. *P_5 is rational, neutral, and involutary but neither maximal nor minimal.*

Example 8. *We return to Figure 1 to illustrate the applicability of the rationality postulate. In all cases, whenever an attacker is believed, the attackee is disbelieved. Hence, the structure of the graph has been taken into account when assigning the probability value for each argument. However, the neutral, involutary, maximal and minimal postulates do not hold for this example. We justify this as follows: satisfaction of the neutral postulate would show that the audience neither believed nor disbelieved any of the arguments; satisfaction of involutary would force the audience to have the same value for all attackers of an argument; satisfaction of maximal postulate would force the audience to completely believe all arguments, and satisfaction of the minimal postulate would force the audience to completely disbelieve all arguments.*

As before let $\mathcal{P}_t(\mathsf{AF})$ be the set of all t-probability functions with $t \in \{\mathsf{COH},\mathsf{SFOU},\mathsf{FOU},\mathsf{SOPT},\mathsf{OPT},\mathsf{JUS},\mathsf{TER},\mathsf{RAT},\mathsf{NEU},\mathsf{INV},\mathsf{MAX},\mathsf{MIN}\}$ and let $\mathcal{P}_T(\mathsf{AF})$ for a set T of conditions be defined as before. We extend the classification from Proposition 1 as follows.

Proposition 6. *Let* $\mathsf{AF} = (\mathsf{Arg}, \rightarrow)$ *be an abstract argumentation framework.*

1. $\emptyset \subsetneq \mathcal{P}_{\mathsf{JUS}}(\mathsf{AF}) \subsetneq \mathcal{P}_{\mathsf{COH}}(\mathsf{AF}) \subsetneq \mathcal{P}_{\mathsf{RAT}}(\mathsf{AF}) \subsetneq \mathcal{P}(\mathsf{AF})$

2. $\emptyset \subsetneq \mathcal{P}_{\mathsf{NEU}}(\mathsf{AF}) \subseteq \mathcal{P}_{\mathsf{INV}}(\mathsf{AF}) \subsetneq \mathcal{P}_{\mathsf{COH}}(\mathsf{AF})$

3. $\emptyset \subsetneq \mathcal{P}_{\mathsf{INV}}(\mathsf{AF}) \subsetneq \mathcal{P}_{\mathsf{SOPT}}(\mathsf{AF})$

4. $\emptyset \subsetneq \mathcal{P}_{\mathsf{MIN}}(\mathsf{AF}) \subsetneq \mathcal{P}_{\mathsf{COH}}(\mathsf{AF})$

5. $\emptyset \subsetneq \mathcal{P}_{\mathsf{MAX}}(\mathsf{AF}) \subsetneq \mathcal{P}_{\mathsf{OPT}}(\mathsf{AF})$

Proof. We only give the proof for 2.). The proofs for 1.) can be found in the works by Thimm (2012) and Hunter (2013), the remaining proofs are straightforward.

The probability function P with $P(E) = 1/|2^{\mathsf{Arg}}|$ for all $E \subseteq \mathsf{Arg}$ has $P(\mathcal{A}) = 0.5$ for all $\mathcal{A} \in \mathsf{Arg}$ and is therefore neutral. It follows that $\mathcal{P}_{\mathsf{NEU}}(\mathsf{AF}) \neq \emptyset$ for every AF. Furthermore, if $P \in \mathcal{P}_{\mathsf{NEU}}(\mathsf{AF})$ then for every $\mathcal{A},\mathcal{B} \in \mathsf{Arg}$, if $\mathcal{A} \rightarrow \mathcal{B}$ we have trivially $P(\mathcal{A}) = 1 - P(\mathcal{B})$, so $P \in \mathcal{P}_{\mathsf{INV}}(\mathsf{AF})$ and then also $P(\mathcal{A}) \leq 1 - P(\mathcal{B})$, i..e, $P \in \mathcal{P}_{\mathsf{COH}}(\mathsf{AF})$. Finally, for $\mathsf{AF} = (\mathsf{Arg}, \rightarrow)$ with $\mathsf{Arg} = \{\mathcal{A},\mathcal{B}\}$ and $\rightarrow = \{(\mathcal{A},\mathcal{B})\}$ any probability function P with $P(\mathcal{A}) = 0.4$ and $P(\mathcal{B}) = 0.4$ is coherent but not involutary. □

Together with Examples 4 and 7 we obtain the strict classification of classes of probability functions as depicted in Figure 5.

The RAT constraint is a weaker version of the COH constraint, and it can be used to capture each admissible labelling as a probability function.

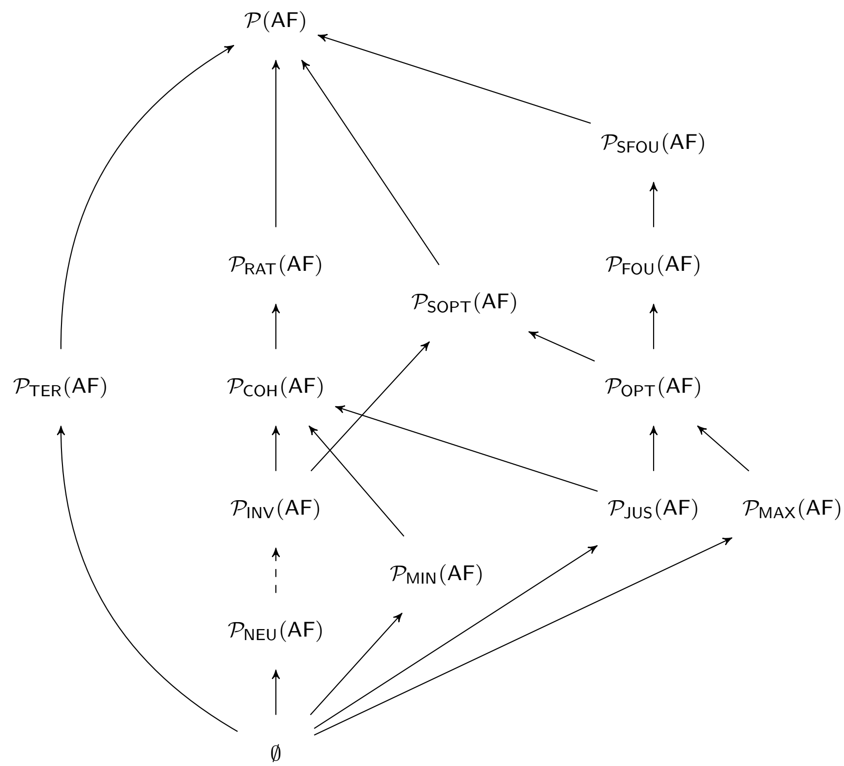

Figure 5: Classes of probability functions (a normal arrow $\rightarrow$ indicates a strict subset relation, a dashed arrow $\dashrightarrow$ indicates a subset relation)

Proposition 7. *If L is an admissible labelling, then there is a $P \in \mathcal{P}_{\mathsf{RAT}}(\mathsf{AF})$ such that $L \sim P$.*

Proof. Let L be an admissible labelling and let P be such that $L \sim P$. Let $\mathcal{A}, \mathcal{B} \in \mathsf{Arg}$ with $\mathcal{A} \rightarrow \mathcal{B}$ and assume $P(\mathcal{A}) > 0.5$. As $L \sim P$ it follows $P(\mathcal{A}) = 1$ and $L(\mathcal{A}) = \mathsf{in}$. As $\mathcal{A} \rightarrow \mathcal{B}$ it follows $L(\mathcal{B}) = \mathsf{out}$ and therefore $P(\mathcal{B}) = 0 \leq 0.5$, showing that P satisfies RAT. $\square$

Furthermore, the epistemic labelling corresponding to each probability function that satisfies the RAT property is conflict-free, which has already been shown b¡ Hunter (2013).

Proposition 8. *Let $\mathsf{AF} = (\mathsf{Arg}, \rightarrow)$ be an abstract argumentation framework. For each $P \in P_{\mathsf{RAT}}(AF)$, $\mathsf{in}(L_P)$ is a conflict-free set of arguments in* AF.

When the argument graph has odd cycles, there is no probability function that is involutary, apart from a neutral probability function.

Proposition 9. *Let $\mathsf{AF} = (\mathsf{Arg}, \rightarrow)$ be an abstract argumentation framework. If* AF *contains an odd cycle (i.e. there is a sequence of attacks $A_1 \rightarrow A_2 \rightarrow ...... \rightarrow A_k$ where $A_1 = A_k$ and k is an even number), and $P \in \mathcal{P}_{\mathsf{INV}}(\mathsf{AF})$ then $P \in P_{\mathsf{NEU}}(\mathsf{AF})$.*

Proof. Assume that there is a sequence of attacks $A_1 \rightarrow A_2 \rightarrow ...... \rightarrow A_k$ where $A_1 = A_k$ and k is an even number. Let $P(A_1) = \alpha$. Hence, $P(A_2) = 1 - \alpha$, $P(A_3) = \alpha$, $\ldots$, $P(A_{k-1}) = \alpha$, $P(A_k) = 1 - \alpha$. Therefore, $P(A_1) = \alpha$ and $P(A_k) = 1 - \alpha$. Yet $A_1 = A_k$. This is only possible if $\alpha = 0.5$. Hence, $P \in P_{\mathsf{NEU}}(\mathsf{AF})$. □

Even when the graph is acyclic, it may be the case that there is no involutary probability function (apart from the neutral probability function). Consider for example an argument graph containing three arguments $\mathcal{A}$, $\mathcal{B}$ and $\mathcal{C}$, with $\mathcal{A}$ attacking both $\mathcal{B}$ and $\mathcal{C}$, and $\mathcal{B}$ attacking $\mathcal{C}$. For this, there is no involutary probability function (apart from the neutral probability function). If we restrict consideration to trees, then we are guaranteed to have a probability function that is involutary and not neutral. But even in this case there are constraints such as "siblings must have the same assignment" as captured in the next result.

Proposition 10. *If $P \in \mathcal{P}_{\mathsf{INV}}(\mathsf{AF})$, then for all $\mathcal{B}_i, \mathcal{B}_j \in \mathsf{Att}_{\mathsf{AF}}(\mathcal{A})$ we have $P(\mathcal{B}_i) = P(\mathcal{B}_j)$.*

Proof. Let $\mathcal{B}_i \rightarrow \mathcal{A}$ and $\mathcal{B}_j \rightarrow \mathcal{A}$ be attacks. Assume $P \in \mathcal{P}_{\mathsf{INV}}(\mathsf{AF})$. Therefore, $P(\mathcal{A}) = 1 - P(\mathcal{B}_i)$ and $P(\mathcal{A}) = 1 - P(\mathcal{B}_j)$. Hence, $P(\mathcal{B}_i) = P(\mathcal{B}_j)$. □

When $P \in \mathcal{P}_{\mathsf{MAX}}(\mathsf{AF})$, the probability function does not take the structure of the graph into account. Hence, there is an incompatibility between a probability function being maximal and a probability function being either rational or coherent (as shown in the proposition below). However, there is compatibility between a probability function being maximal and a probability function being founded since each $P \in \mathcal{P}_{\mathsf{MAX}}(\mathsf{AF})$ is in $\mathcal{P}_{\mathsf{FOU}}(\mathsf{AF})$.

Proposition 11. *Let $\mathsf{AF} = (\mathsf{Arg}, \rightarrow)$ be an abstract argumentation framework. If there are $\mathcal{A}, \mathcal{B} \in \mathsf{Arg}$ such that $\mathcal{A} \rightarrow \mathcal{B}$, then $\mathcal{P}_{\mathsf{RAT}}(AF) \cap \mathcal{P}_{\mathsf{MAX}}(AF) = \emptyset$.*

Proof. Assume there is an attack $\mathcal{A} \rightarrow \mathcal{B}$. So for all $P \in \mathcal{P}_{\mathsf{RAT}}(AF)$, if $P(\mathcal{A}) = 1$, then $P(\mathcal{B}) \leq 0.5$, and $P(\mathcal{B}) = 1$, then $P(\mathcal{A}) \leq 0.5$. And for all $P \in \mathcal{P}_{\mathsf{MAX}}(AF)$, $P(\mathcal{A}) = 1$ and $P(\mathcal{B}) = 1$. So $\mathcal{P}_{\mathsf{RAT}}(AF) \cap \mathcal{P}_{\mathsf{MAX}}(AF) = \emptyset$. □

In conclusion, we have identified epistemic extensions that are obtained from rational probability functions as being an alternative to extensions obtained by Dung's definitions. Rational probability functions are more general than coherent probability functions, and allow the audience more flexibility in expressing their beliefs in the arguments whilst taking the structure of the argument graph into account. We have also considered alternatives such as the involutary probability functions but these are over-constrained. By means of Example 5, Example 6, and Example 8 we showed scenarios where we might want certain standard or non-standard properties to hold or to not hold. So in general, we believe that it depends on the applications as to whether one would expect particular postulates to hold.

6. Partial Probability Assessments

The framework outlined so far allows us to assess whether probability functions reflect the topology of an argumentation framework and can be used for uncertain reasoning based on argumentation. In the following, we will investigate the case when we already have probabilistic information on some arguments and need to infer meaningful probabilities for the remaining arguments.

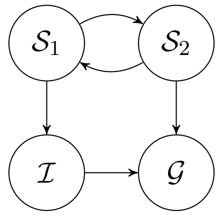

Figure 6: The argumentation framework from Example 9

Example 9. *Consider a court case where the defendant John is either innocent or guilty to have committed the murder of Frank. Footage from a surveillance camera at the crime scene gives evidence that a person looking like John was present at the time of the crime, giving a reason that John is not innocent. However, footage from another surveillance camera far away from the crime scene gives evidence that a person looking like John was not present at the time of the crime, giving a reason that John is not guilty. This scenario can be modelled with the argumentation framework depicted in Figure 6 where the arguments $\mathcal{I}$, $\mathcal{G}$, $\mathcal{S}_1$, and $\mathcal{S}_2$ are given as*

$\mathcal{I}$**:** *A person accused of murder is presumed innocent unless guilt is proven beyond reasonable doubt, so John is innocent.*

$\mathcal{G}$**:** *John is guilty because he had a motive.*

$\mathcal{S}_1$**:** *John is guilty because surveillance footage proves that he was at the crime scene, which proves murder beyond reasonable doubt.*

$\mathcal{S}_2$**:** *John is innocent because surveillance footage proves that he was not at the crime scene.*

Note that there is no attack from $\mathcal{G}$ to $\mathcal{I}$ in Figure 6 as $\mathcal{G}$ does not provide a proof beyond reasonable doubt that John is guilty.

Now the footage from the camera $\mathcal{S}_1$ is examined by a lab which assesses that the probability of the person in the pictures is indeed John is 0.7. So given $P(\mathcal{S}_1) = 0.7$ what are now adequate probabilities for the remaining arguments?

A partial function $\beta : \mathsf{Arg} \to [0,1]$ on Arg is called a *partial probability assignment*. Let Π denote the set of all partial probability assignments. Let $\mathsf{dom}\,\beta \subseteq \mathsf{Arg}$ be the domain of β, i.e., the arguments for which a probabilistic assessment is available. We are now interested in deriving probabilities for the remaining arguments $\mathsf{Arg} \setminus \mathsf{dom}\,\beta$, taking the information we already have in β and the argumentation framework AF into account.

A probability function $P \in \mathcal{P}(\mathsf{AF})$ is *β-compliant* if for every $\mathcal{A} \in \mathsf{dom}\,\beta$ we have $\beta(\mathcal{A}) = P(\mathcal{A})$. Let $\mathcal{P}^{\beta}(\mathsf{AF}) \subseteq \mathcal{P}(\mathsf{AF})$ be the set of all β-compliant probability functions. Observe that $\mathcal{P}^{\beta}(\mathsf{AF})$ is always non-empty.

Proposition 12. *For all partial $\beta : \mathsf{Arg} \to [0,1]$, $\mathcal{P}^{\beta}(\mathsf{AF}) \neq \emptyset$.*

Proof. Let $\beta : \mathsf{Arg} \to [0,1]$ with $\mathsf{dom}\,\beta = \{\mathcal{A}_1, \ldots, \mathcal{A}_n\}$ and assume $\beta(\mathcal{A}_1) < \ldots < \beta(\mathcal{A}_n)$. Define $P : 2^{\mathsf{Arg}} \to [0,1]$ via (let $i = 1, \ldots, n$)

$$P(\{\mathcal{A}_1, \ldots, \mathcal{A}_i\}) = \beta(\mathcal{A}_i) - \sum_{j=1}^{i-1} \beta(\mathcal{A}_j) \tag{1}$$

$$P(\emptyset) = 1 - \beta(\mathcal{A}_n)$$

$$P(X) = 0 \qquad \text{for all remaining sets } X$$

It can be easily verified that P is a β-compliant probability function, hence $P \in \mathcal{P}^{\beta}(\mathsf{AF})$. If $\mathsf{dom}\,\beta = \{\mathcal{A}_1, \ldots, \mathcal{A}_n, \mathcal{B}\}$ with $\beta(\mathcal{A}_1) < \ldots < \beta(\mathcal{A}_n)$ and $\beta(\mathcal{B}) = \beta(\mathcal{A}_k)$ for some $k \in \{1, \ldots, n\}$, replace the left side of (1) for $i = k$ with $P(\{\mathcal{A}_1, \ldots, \mathcal{A}_i, \mathcal{B}\}))$. The general case follows by induction. $\square$

Of course, not all probability functions in $\mathcal{P}^{\beta}(\mathsf{AF})$ are adequate for reasoning as they may not take the actual argumentation framework into account. Recall that for $T \subseteq \{\mathsf{RAT}, \mathsf{COH}, \mathsf{SFOU}, \mathsf{FOU}, \mathsf{OPT}, \mathsf{SOPT}, \mathsf{JUS}\}$, the set $\mathcal{P}_T(\mathsf{AF})$ contains all probability functions which comply with all considered rationality conditions. Given a partial probability assessment β and some rationality conditions T, for the remainder of this section we assume $\mathcal{P}_T(\mathsf{AF}) \cap \mathcal{P}^{\beta}(\mathsf{AF}) \neq \emptyset$, i.e., there is at least one probability function that is both β-compliant and adheres to the set of rationality conditions (we address the case $\mathcal{P}_T(\mathsf{AF}) \cap \mathcal{P}^{\beta}(\mathsf{AF}) = \emptyset$ in the next section).

Define $\mathcal{P}^{\beta}_T(\mathsf{AF}) = \mathcal{P}_T(\mathsf{AF}) \cap \mathcal{P}^{\beta}(\mathsf{AF})$.

Definition 6. *Let β be a partial probability assignment and T a set of rationality conditions. Then the* possible probabilities *of $\mathcal{A} \in \mathsf{Arg} \setminus \mathsf{dom}\,\beta$, denoted as $\mathsf{p}^{\beta}_{T,\mathsf{AF}}(\mathcal{A})$, is defined as $\mathsf{p}^{\beta}_{T,\mathsf{AF}}(\mathcal{A}) = \{P(\mathcal{A}) \mid P \in \mathcal{P}^{\beta}_T(\mathsf{AF})\}$.*

Under the assumption $\mathcal{P}^{\beta}_T(\mathsf{AF}) \neq \emptyset$, it is clear that $\mathsf{p}^{\beta}_{T,\mathsf{AF}}(\mathcal{A}) \neq \emptyset$ as well.

Example 10. *We continue Example 9 with $\beta_1(\mathcal{S}_1) = 0.7$ and assume $T_1 = \{\mathit{COH}\}$. Then for the arguments $\mathcal{S}_2, \mathcal{I}, \mathcal{G}$ we obtain $\mathsf{p}^{\beta_1}_{T_1,\mathsf{AF}}(\mathcal{S}_2) = [0, 0.3]$, $\mathsf{p}^{\beta_1}_{T_1,\mathsf{AF}}(\mathcal{I}) = [0, 0.3]$, and $\mathsf{p}^{\beta_1}_{T_1,\mathsf{AF}}(\mathcal{G}) = [0, 0.7]$.*

We need some set theoretical notions before we can state our next result. A subset X of a topological space is *(path-)connected*, if for every two elements $x, y \in X$ there is a continuous function $f : [0,1] \to X$ with $f(0) = x$ and $f(1) = y$.[2] A set X is called *convex*, if for every two elements $x, y \in X$ and $\delta \in [0,1]$ we also have $\delta x + (1-\delta)y \in X$. A set X is *closed* if for every sequence $x_1, x_2, \ldots \in X$, if $\lim_{n\to\infty} x_i$ exists then $\lim_{n\to\infty} x_i \in X$.

Proposition 13. *Let AF be an abstract argumentation framework and β a partial probability assignment.*

1. *The set $\mathcal{P}^{\beta}(\mathsf{AF})$ is connected, convex, and closed.*
2. *The sets $\mathcal{P}(\mathsf{AF})$, $\mathcal{P}_{\mathsf{COH}}(\mathsf{AF})$, $\mathcal{P}_{\mathsf{SFOU}}(\mathsf{AF})$, $\mathcal{P}_{\mathsf{FOU}}(\mathsf{AF})$, $\mathcal{P}_{\mathsf{OPT}}(\mathsf{AF})$, $\mathcal{P}_{\mathsf{SOPT}}(\mathsf{AF})$, $\mathcal{P}_{\mathsf{JUS}}(\mathsf{AF})$ are connected, convex, and closed.*

2. Note that $\mathcal{P}(\mathsf{AF})$ is a topological space as it can be identified with a subspace of $[0,1]^n$ with $n = |2^{\mathsf{Arg}}|$.

3. *The set* $\mathcal{P}_{\mathsf{RAT}}(\mathsf{AF})$ *is connected and closed, but not convex in general.*

4. *For every* $T \subseteq \{\mathit{COH},\mathit{SFOU},\mathit{FOU},\mathit{OPT},\mathit{SOPT},\mathit{JUS}\}$ *the sets* $\mathcal{P}_T(\mathsf{AF})$ *and* $\mathcal{P}^{\beta}_T(\mathsf{AF})$ *are connected, convex, and closed.*

5. *For every* $T \subseteq \{\mathit{COH},\mathit{SFOU},\mathit{FOU},\mathit{OPT},\mathit{SOPT},\mathit{JUS}\}$ *and* $\mathcal{A} \in \mathsf{Arg}$ *the set* $\mathsf{p}^{\beta}_{T,\mathsf{AF}}(\mathcal{A})$ *is connected, convex, and closed.*

Proof.

1. Let $P_1, P_2 \in \mathcal{P}^{\beta}(\mathsf{AF})$ and consider the convex combination $P = \delta P_1 + (1-\delta)P_2$ for some $\delta \in [0,1]$. For every $\mathcal{A} \in \mathsf{dom}\,\beta$ we have $P(\mathcal{A}) = \delta P_1(\mathcal{A}) + (1-\delta)P_2(\mathcal{A}) = \delta\beta(\mathcal{A}) + (1-\delta)\beta(\mathcal{A}) = \beta(\mathcal{A})$ and therefore $P \in \mathcal{P}^{\beta}(\mathsf{AF})$. Closure of $\mathcal{P}^{\pi}(\mathsf{AF})$ is straightforward and connectedness follows from convexity.

2. Analogous to 1.).

3. Let $\mathsf{AF} = (\mathsf{Arg}, \rightarrow)$ be given by $\mathsf{Arg} = \{\mathcal{A}, \mathcal{B}\}$ and $\rightarrow = (\mathcal{A}, \mathcal{B})$. Consider $P_1, P_2 \in \mathcal{P}_{\mathsf{RAT}}(\mathsf{AF})$ with

$$\begin{aligned} P_1(\mathcal{A}) &= 1 & \qquad P_1(\mathcal{B}) &= 0.4 \\ P_2(\mathcal{A}) &= 0.4 & \qquad P_2(\mathcal{B}) &= 0.8 \end{aligned}$$

For the convex combination $P = 0.5P_1 + 0.5P_2$ we obtain $P(\mathcal{A}) = 0.7$ and $P(\mathcal{B}) = 0.6$, i.e. $P \notin \mathcal{P}_{\mathsf{RAT}}(\mathsf{AF})$. However, $\mathcal{P}_{\mathsf{RAT}}(\mathsf{AF})$ is closed as for every converging sequence $P_1, P_2, \ldots$ with $P_i \in \mathcal{P}_{\mathsf{RAT}}(\mathsf{AF})$ for all $i \in \mathbb{N}$ there is $N \in \mathbb{N}$ such that for all $\mathcal{A} \rightarrow \mathcal{B}$ either

- $P_j(\mathcal{A}) \leq 0.5$ for all $j > N$; then $\lim_{i\to\infty} P_j(\mathcal{A}) \leq 0.5$ as well and the condition of coherence is trivially satisfied, or
- $P_j(\mathcal{A}) > 0.5$ and consequently $P_j(\mathcal{B}) \leq 0.5$ for all $j > N$; then $\lim_{i\to\infty} P_j(\mathcal{B}) \leq 0.5$ as well and the condition of coherence is satisfied in any case.

Consider now $P_0 \in \mathcal{P}(\mathsf{AF})$ defined via $P(\emptyset) = 1$ and $P(E) = 0$ for all $E \subseteq \mathsf{Arg}$ with $E \neq \emptyset$. Then $P_0(\mathcal{A}) = 0$ for all $\mathcal{A} \in \mathsf{Arg}$ and therefore $P_0 \in \mathcal{P}_{\mathsf{RAT}}(\mathsf{AF})$. Let $P \in \mathcal{P}_{\mathsf{RAT}}(\mathsf{AF})$ and $P_\delta = \delta P_0 + (1-\delta)P$ be a convex combination of P_0 and P for any $\delta \in [0,1]$. Let $\mathcal{A} \rightarrow \mathcal{B}$ and assume $P_\delta(\mathcal{A}) \geq 0.5$. As $P_0(\mathcal{A}) = 0$ it follows $P(\mathcal{A}) \geq 0.5$ as well. Then $P(\mathcal{B}) < 0.5$ as P is rational. As both $P_0(\mathcal{B}) = 0 < 0.5$ and $P(\mathcal{B}) < 0.5$ it follows $P_\delta(\mathcal{B}) < 0.5$ as well as P_δ is a convex combination. It follows $P_\delta \in \mathcal{P}_{\mathsf{RAT}}(\mathsf{AF})$ for all $\delta \in [0,1]$. Therefore, there is a path from P_0 to every other $P \in \mathcal{P}_{\mathsf{RAT}}(\mathsf{AF})$ which implies that $\mathcal{P}_{\mathsf{RAT}}(\mathsf{AF})$ is connected.

4. This follows directly from the fact that the finite intersection of convex sets is convex and that the finite intersection of closed sets is closed.

5. This follows from the fact that the projection of a connected, convex, and closed set is again connected, convex, and closed. □

The final statement above is equivalent to saying that $\mathsf{p}^{\beta}_{T,\mathsf{AF}}(\mathcal{A})$ is an interval. Note also, that if $\mathsf{RAT} \in T$ the set $\mathcal{P}_T(\mathsf{AF})$ is closed, but not necessarily connected or convex. In the following, we focus on the cases where $T \subseteq \{\mathsf{COH},\mathsf{SFOU},\mathsf{FOU},\mathsf{OPT},\mathsf{SOPT},\mathsf{JUS}\}$.

Proposition 13 implies that the problem of determining $\mathsf{p}^{\beta}_{T,\mathsf{AF}}(\mathcal{A})$ is equivalent to the classical *probabilistic entailment problem* (Jaumard, Hansen, & Poggi, 1991; Hansen & Jaumard, 2000; Lukasiewicz, 2000). We can directly exploit this relationship to make some observations on the computational complexity of some problems related to $\mathsf{p}^{\beta}_{T,\mathsf{AF}}(\mathcal{A})$.

Proposition 14. *Let* AF *be an abstract argumentation framework,* β *a partial probability assignment, and* $T \subseteq \{\mathsf{COH},\mathsf{SFOU},\mathsf{FOU}, \mathsf{OPT},\mathsf{SOPT},\mathsf{JUS}\}$.

1. *Deciding* $p \in \mathsf{p}^{\beta}_{T,\mathsf{AF}}(\mathcal{A})$ *for some* $p \in [0,1]$ *is* NP*-complete.*
2. *Deciding* $[l,u] = \mathsf{p}^{\beta}_{T,\mathsf{AF}}(\mathcal{A})$ *for some* $l,u \in [0,1]$ *is* D^{P}*-complete.*
3. *Computing* $l,u \in [0,1]$ *such that* $[l,u] = \mathsf{p}^{\beta}_{T,\mathsf{AF}}(\mathcal{A})$ *is* $\mathsf{FP}^{\mathsf{NP}}$*-complete.*

Proof. The conditions imposed by β on the probabilities of arguments can be represented as linear constraints on P. Furthermore, the conditions of the properties in T can also be represented as linear constraints. Then we can represent these problems as probabilistic knowledge bases in the sense of Lukasiewic (2000) and the results follow directly from results by Lukasiewic (2000). □

Besides using $\mathsf{p}^{\beta}_{T,\mathsf{AF}}(\mathcal{A})$ to obtain bounds on the probabilities of the remaining arguments, we might also be interested in obtaining point probabilities for the remaining arguments that are as unbiased as possible, giving the probabilistic information of β. One can use the *principle of maximum entropy* (see also Proposition 4) for this purpose, which is thanks to Proposition 13 also applicable in our context.

Definition 7. *Let* AF *be an abstract argumentation framework,* β *a partial probability assignment, and* $T \subseteq \{\mathsf{COH},\mathsf{SFOU},\mathsf{FOU},\mathsf{OPT},\mathsf{SOPT},\mathsf{JUS}\}$*. Define the set* $\mathcal{P}^{\beta,\mathsf{AF},T}_{\mathsf{ME}}$ *via*

$$\mathcal{P}^{\beta,\mathsf{AF},T}_{\mathsf{ME}} = \arg \max_{Q \in \mathcal{P}^{\beta}_{T}(\mathsf{AF})} H(Q)$$

Proposition 15. *The set* $\mathcal{P}^{\beta,\mathsf{AF},T}_{\mathsf{ME}}$ *contains exactly one uniquely defined probability function.*

Proof. Due to Proposition 13 the set $\mathcal{P}_T(\mathsf{AF}) \cap \mathcal{P}^{\beta}(\mathsf{AF})$ is convex and closed. Maximizing entropy over a convex and closed set has a unique solution (Paris, 1994). □

Due to the above proposition we identify the singleton set $\mathcal{P}^{\beta,\mathsf{AF},T}_{\mathsf{ME}}$ with its only element, e.g., we write $\mathcal{P}^{\beta,\mathsf{AF},T}_{\mathsf{ME}}(\mathcal{A})$ to denote the probability $P(\mathcal{A})$ with $\{P\} = \mathcal{P}^{\beta,\mathsf{AF},T}_{\mathsf{ME}}$.

Example 11. *We continue Examples 9 and 10 with* $\beta_1(\mathcal{S}_1) = 0.7$ *and assume* $T_1 = \{\mathsf{COH}\}$*. Then we have* $\mathcal{P}^{\beta_1,\mathsf{AF},T_1}_{\mathsf{ME}}(\mathcal{S}_2) = 0.3$*,* $\mathcal{P}^{\beta_1,\mathsf{AF},T_1}_{\mathsf{ME}}(\mathcal{I}) = 0.3$*, and* $\mathcal{P}^{\beta_1,\mathsf{AF},T_1}_{\mathsf{ME}}(\mathcal{G}) = 0.5$*. Recall that* $\mathsf{p}^{\beta_1}_{T_1,\mathsf{AF}}(\mathcal{I}) = [0, 0.3]$ *(Example 10) and observe that* $\mathcal{P}^{\beta_1,\mathsf{AF},T_1}_{\mathsf{ME}}(\mathcal{I}) = 0.3$ *which is the maximal probability that can be assigned to* $\mathcal{I}$*. However, note also that this value is closest to* 0.5

which is the probability value with the least amount of information (in the information-theoretic sense). Indeed, it can be observed that all probabilities assigned above are those closest to 0.5 *which is a general feature of reasoning based on the principle of maximum entropy (note, however, that in more complex settings involving other rationality conditions the function* $\mathcal{P}_{\mathsf{ME}}^{\beta,\mathsf{AF},T}$ *is not always characterized as easily as that).*

Proposition 4 already pointed out the relationship of the principle of maximum entropy to grounded semantics. Note also that for β with $\mathsf{dom}\,\beta = \emptyset$ and $T = \{\mathsf{JUS}\}$ we have that $\mathcal{P}_{\mathsf{ME}}^{\beta,\mathsf{AF},T}$ corresponds to the grounded labelling of AF. Taking into account partial probabilistic information we therefore extended the notion of a grounded labelling and obtain a probability function that is "as grounded as possible". Similarly, if we exchange the maximum in Definition 7 by a minimum, we obtain a generalization of the notion of stable labelings, cf. Proposition 4 item 3.

7. Contradictory Probability Assessments

So far we assumed $\mathcal{P}_T^{\beta}(\mathsf{AF}) \neq \emptyset$. In this section, we investigate the general scenario without this assumption. Consider the following example.

Example 12. *We continue Example 9. New evidence obtained by analyzing the footage from camera* $\mathcal{S}_2$ *suggests that the probability of the person in those pictures is indeed John is* 0.4*. Therefore, the partial probability assessment* β_1' *is defined by* $\beta_1'(\mathcal{S}_1) = 0.7$ *and* $\beta_1'(\mathcal{S}_2) = 0.4$*. Considering the set of rationality conditions* $T = \{\mathsf{COH}\}$ *one can see that* $\mathcal{P}_T^{\beta_1'}(\mathsf{AF}) = \mathcal{P}_T(\mathsf{AF}) \cap \mathcal{P}^{\beta_1'}(\mathsf{AF}) = \emptyset$ *as the condition*

$$P(\mathcal{S}_1) \leq 1 - P(\mathcal{S}_2), \tag{2}$$

which is necessary for having $P \in \mathcal{P}_T(\mathsf{AF})$*, cannot be satisfied for any* β_1'*-compliant* P*. In this case, one would still be interested in obtaining a "reasonable" probability for example for* $\mathcal{I}$*.*

We address the issue outlined in the example above by, first, analyzing in a quantitative manner *how much* a partial probability assessment deviates from satisfying a given set of rationality conditions, and afterwards using this analysis to provide reasonable probabilities for the remaining arguments.

7.1 Inconsistency Measures for Contradictory Probability Assessments

We first address the question of how to measure the distance (or inconsistency) of a given partial probability assignment $\beta : \mathsf{Arg} \to [0,1]$ to the set $\mathcal{P}_T(\mathsf{AF})$ of probability functions. As before we restrict our attention to $T \subseteq \{\mathsf{COH}, \mathsf{SFOU}, \mathsf{FOU}, \mathsf{OPT}, \mathsf{SOPT}, \mathsf{JUS}\}$). Recall that Π denotes the set of all partial probability assignments and $\mathbb{A}$ denotes the set of all abstract argumentation frameworks.

Definition 8. *An* inconsistency measure $\mathcal{I}_T$ *is a function* $\mathcal{I}_T : \Pi \times \mathbb{A} \to [0, \infty)$*.*

The intuition behind an inconsistency measure $\mathcal{I}_T$ is that for a partial probability assessment β and an argumentation framework AF, the value $\mathcal{I}_T(\beta, \mathsf{AF})$ quantitatively assesses the severity of β violating the rationality conditions imposed by T in AF. In particular, larger values indicate greater violation while attaining the minimum $\mathcal{I}_T(\beta, \mathsf{AF})$ suggests that β does not violate the rationality conditions imposed by T in AF at all. Note that inconsistency measures have been investigated before mostly in the context of classical logic (Hunter & Konieczny, 2010).

Before formalizing the intuition behind an inconsistency measure, we need some further notation. Let β, β' be partial probability assignments. We say that β' is an *expansion* of β if $\mathsf{dom}\,\beta \subseteq \mathsf{dom}\,\beta'$ and $\beta(\mathcal{A}) = \beta'(\mathcal{A})$ for all $\mathcal{A} \in \mathsf{dom}\,\beta$. If $\mathsf{dom}\,\beta \cap \mathsf{dom}\,\beta' = \emptyset$ then define $(\beta \circ \beta')$ with $\mathsf{dom}\,\beta \circ \beta' = \mathsf{dom}\,\beta \cup \mathsf{dom}\,\beta'$ via $(\beta \circ \beta')(\mathcal{A}) = \beta(\mathcal{A})$ for $\mathcal{A} \in \mathsf{dom}\,\beta$ and $(\beta \circ \beta')(\mathcal{A}) = \beta'(\mathcal{A})$ for $\mathcal{A} \in \mathsf{dom}\,\beta'$. Two arguments $\mathcal{A}, \mathcal{B} \in \mathsf{Arg}$ are *simply connected* in AF if there is a path between them in the undirected version of AF. Let $CC(\mathsf{AF})$ be the set of all connected components of AF w.r.t. simple connectedness.

Now, some desirable properties for an inconsistency measure in our context—motivated by similar properties for inconsistency measures in classical logics (Hunter & Konieczny, 2010)—are as follows.

Consistency $\mathcal{I}_T(\beta, \mathsf{AF}) = 0$ iff $\mathcal{P}_T(\mathsf{AF}) \cap \mathcal{P}^\beta(\mathsf{AF}) \neq \emptyset$.

Monotonicity If β' is an expansion of β then $\mathcal{I}_T(\beta, \mathsf{AF}) \leq \mathcal{I}_T(\beta', \mathsf{AF})$.

Super-additivity If $\mathsf{dom}\,\beta \cap \mathsf{dom}\,\beta' = \emptyset$ then $\mathcal{I}_T(\beta \circ \beta', \mathsf{AF}) \geq \mathcal{I}_T(\beta, \mathsf{AF}) + \mathcal{I}_T(\beta', \mathsf{AF})$.

Separability $\mathcal{I}_T(\beta, \mathsf{AF}) = \sum_{\mathsf{AF}' \in CC(\mathsf{AF})} \mathcal{I}_T(\beta, \mathsf{AF}')$.

The property *consistency* states that an inconsistency measure must attain its minimal value if and only if there is at least one β-compliant probability function P that satisfies all conditions T w.r.t. AF. The property *monotonicity* states that the inconsistency cannot decrease when adding further constraints to a partial probability assessment. The property *super-additivity* means that the sum of the inconsistency values of two independent probability assessments cannot be larger than the inconsistency value of the joint probability assessment. Finally, the property *separability* demands that the inconsistency value decomposes on the connected components of an argumentation framework.

In order to implement inconsistency measures for our setting of probabilistic abstract argumentation, we base our measures on metrics on probability functions, see also the works by Thimm (2013), De Bona & Finger (2015), and Grant & Hunter (2013).

Definition 9. *A function $d : \mathcal{P} \times \mathcal{P} \to [0, \infty)$ is called a* pre-metrical distance measure *if it satisfies $d(P, P') = 0$ if and only if $P = P'$.*

In the following, we refer to *pre-metrical distance measures* simply by *distance measures* (note that we do not impose the properties *symmetry* and *triangle equality* of full distance measures). Examples of such distance measures are (let $p \geq 1$)

$$d_{\mathrm{KL}}(P, P') = \sum_{x \in \mathsf{dom}\,P \cap \mathsf{dom}\,P'} P(x) \log \frac{P(x)}{P'(x)}$$

$$d_p(P, P') = \sqrt[p]{\sum_{x \in \mathsf{dom}\,P \cap \mathsf{dom}\,P'} |P(x) - P'(x)|^p}$$

In the definition of d_{KL}, if $x = 0$ we assume $x \log x/y = 0$ and if $x \neq 0$ but $y = 0$ we assume $x \log x/y = x$. The measure d_{KL} is also called the *Kullback-Leibler divergence* (or *relative entropy*). The measure d_p is called the *p-norm distance*. In the following, we will use these two distance measures as examples to illustrate our approach. Note that any other pre-metrical distance measure can be used instead.

Note that both measures d_{KL} and d_p are defined over the set $\mathsf{dom}\, P \cap \mathsf{dom}\, P'$. This is only a technical necessity in order to have well-defined measures for all pairs of probability functions. In the following, distance measures are only applied on pairs of probability functions P and P' such that $P, P' \in \mathcal{P}(\mathsf{AF})$ for some AF, i.e., $\mathsf{dom}\, P = \mathsf{dom}\, P'$.

For a distance measure d, a probability function $P \in \mathcal{P}$ and closed sets $\mathcal{Q}, \mathcal{Q}' \subseteq \mathcal{P}$ we abbreviate

$$
\begin{aligned}
d(P, \mathcal{Q}) &= \min_{P' \in \mathcal{Q}} d(P, P') \\
d(\mathcal{Q}, P) &= \min_{P' \in \mathcal{Q}} d(P', P) \\
d(\mathcal{Q}, \mathcal{Q}') &= \min_{P' \in \mathcal{Q}} d(P', \mathcal{Q}')
\end{aligned}
$$

Using a distance measure d on probability functions we define a general inconsistency measure as follows.

Definition 10. *Let d be a distance measure and $T \subseteq \{\mathsf{COH}, \mathsf{SFOU}, \mathsf{FOU}, \mathsf{OPT}, \mathsf{SOPT}, \mathsf{JUS}\}$. The* distance-based *inconsistency measure $\mathcal{I}^d_T : \Pi \times \mathbb{A} \to [0, \infty)$ for T and d is defined via*

$$
\mathcal{I}^d_T(\beta, \mathsf{AF}) = d(\mathcal{P}^\beta(\mathsf{AF}), \mathcal{P}_T(\mathsf{AF}))
$$

In other words, $\mathcal{I}^d_T(\beta, \mathsf{AF})$ is the minimal distance w.r.t. d of a β-compliant probability function P_1 and a probability function P_2 that satisfies the rationality conditions of T w.r.t. AF.

Example 13. *We continue Example 12 with $T_1 = \{\mathsf{COH}\}$ and β'_1 defined by $\beta'_1(\mathcal{S}_1) = 0.7$ and $\beta'_1(\mathcal{S}_2) = 0.4$. For d_1 (the Manhattan distance) it can be easily seen that $\mathcal{I}^{d_1}_{T_1}(\beta'_1, \mathsf{AF}) = 0.1$ as this amounts to the absolute amount Equation (2) is violated. For d_2 (the Euclidean distance) we obtain $\mathcal{I}^{d_2}_{T_1}(\beta'_1, \mathsf{AF}) \approx 0.037$[3]. For d_{KL} (the Kullback-Leibler divergence) we obtain $\mathcal{I}^{d_{KL}}_{T_1}(\beta'_1, \mathsf{AF}) \approx 0.625$. A geometrical interpretation for both d_2 and d_{KL} is hard to provide but compare those values to the values obtained for β_2 defined by $\beta_2(\mathcal{S}_1) = 0.8$ and $\beta_2(\mathcal{S}_2) = 0.9$: $\mathcal{I}^{d_1}_{T_1}(\beta_2, \mathsf{AF}) = 0.7$, $\mathcal{I}^{d_2}_{T_1}(\beta_2, \mathsf{AF}) \approx 0.403$, and $\mathcal{I}^{d_{KL}}_{T_1}(\beta_2, \mathsf{AF}) \approx 0.312$. From an intuitive point of view β_2 seems more inconsistent than β'_1 (the constraint (2) is violated to a larger extent) and both $\mathcal{I}^{d_1}_{T_1}(\beta_2, \mathsf{AF})$ and $\mathcal{I}^{d_2}_{T_1}(\beta_2, \mathsf{AF})$ comply with this intuition as they assign larger inconsistency values to β_2 than to β'_1. For $\mathcal{I}^{d_{KL}}_{T_1}(\beta_2, \mathsf{AF})$ we obtain the opposite result, due to the fact that d_{KL} does not measure distance of probabilities but distance of information content.*

As can be seen by the following results, the family of inconsistency measures $\mathcal{I}^d_T$ complies with our formalization of a meaningful inconsistency measure.

3. Values of inconsistency measures were determined by using the OpenOpt optimization package `http://openopt.blogspot.de`

Proposition 16. *If d is a pre-metrical distance measure then $\mathcal{I}^d_T$ satisfies* consistency.

Proof. If $\mathcal{I}^d_T(\beta, \mathsf{AF}) = 0$ then $d(\mathcal{P}^\beta(\mathsf{AF}), \mathcal{P}_T(\mathsf{AF})) = 0$ and there are $P \in \mathcal{P}^\beta(\mathsf{AF})$ and $P' \in \mathcal{P}_T(\mathsf{AF})$ with $d(P, P') = 0$. As d is pre-metrical it follows $P = P'$ and $\mathcal{P}_T(\mathsf{AF}) \cap \mathcal{P}^\beta(\mathsf{AF}) \neq \emptyset$. The other direction is analogous. □

Proposition 17. *The function $\mathcal{I}^{d_{KL}}_T$ satisfies* consistency *and* monotonicity.

Proof. Consistency follows from Proposition 16 as d_{KL} is pre-metrical for discrete probability functions. Let β' be an expansion of β. Then $\mathcal{P}^{\beta'}(\mathsf{AF}) \subseteq \mathcal{P}^\beta(\mathsf{AF})$ and

$$d_{\mathrm{KL}}(\mathcal{P}^\beta(\mathsf{AF}), \mathcal{P}_T(\mathsf{AF})) \leq d_{\mathrm{KL}}(\mathcal{P}^{\beta'}(\mathsf{AF}), \mathcal{P}_T(\mathsf{AF}))$$

as $\min_{Q \in \mathcal{P}^\beta(\mathsf{AF})} d_{\mathrm{KL}}(Q, \mathcal{P}_T(\mathsf{AF})) \leq \min_{Q \in \mathcal{P}^{\beta'}(\mathsf{AF})} d_{\mathrm{KL}}(Q, \mathcal{P}_T(\mathsf{AF}))$. □

We previously conjectured that $\mathcal{I}^{d_{\mathrm{KL}}}_T$ also satisfies *super-additivity* and *separability* (2016a). This is not the case in general as the following example shows.

Example 14. *Consider an argumentation framework AF_1 consisting only of a self-attacking argument $\mathcal{A}$, i.e., $\mathsf{AF}_1 = (\mathsf{Arg}_1, \rightarrow_1)$ with $\mathsf{Arg}_1 = \{\mathcal{A}\}$ and $\rightarrow_1 = \{(\mathcal{A}, \mathcal{A})\}$. We consider $T = \{\mathsf{JUS}\}$. Observe that $\mathcal{P}_{\mathsf{JUS}}(\mathsf{AF}_1) = \{P_0\}$ with $P_0(\emptyset) = 0.5$ and $P_0(\{\mathcal{A}\}) = 0.5$. Consider a partial probability assessment β_1 with $\beta_1(\mathcal{A}) = 1$ and observe $\mathcal{P}^{\beta_1}(\mathsf{AF}_1) = \{P_1\}$ with $P_1(\emptyset) = 0$ and $P_1(\{\mathcal{A}\}) = 1$. Using the binary logarithm in d_{KL} it follows*

$$\begin{aligned}
\mathcal{I}^{d_{KL}}_{\mathsf{JUS}}(\beta_1, \mathsf{AF}_1) &= d_{KL}(\mathcal{P}^{\beta_1}(\mathsf{AF}_1), \mathcal{P}_{\mathsf{JUS}}(\mathsf{AF}_1)) \\
&= d_{KL}(P_1, P_0) \\
&= P_1(\emptyset) \log \frac{P_1(\emptyset)}{P_0(\emptyset)} + P_1(\{\mathcal{A}\}) \log \frac{P_1(\{\mathcal{A}\})}{P_0(\{\mathcal{A}\})} \\
&= 0 \log \frac{0}{0.5} + 1 \log \frac{1}{0.5} \\
&= 1
\end{aligned}$$

Now consider $\mathsf{AF}_2 = (\mathsf{Arg}_2, \rightarrow_2)$ with $\mathsf{Arg}_2 = \{\mathcal{B}\}$ and $\rightarrow_2 = \{(\mathcal{B}, \mathcal{B})\}$ and β_2 with $\beta_2(\mathcal{B}) = 1$. Analogously, we obtain $\mathcal{I}^{d_{KL}}_{\mathsf{JUS}}(\beta_2, \mathsf{AF}_2) = 1$ as well.

Define $\mathsf{AF} = (\mathsf{Arg}, \rightarrow)$ to be the union of AF_1 and AF_2, i.e., $\mathsf{Arg} = \mathsf{Arg}_1 \cup \mathsf{Arg}_2$ and $\rightarrow = \rightarrow_1 \cup \rightarrow_2$. Observe that AF decomposes into the two components AF_1 and AF_2 and that

$$\begin{aligned}
\mathcal{P}_{\mathsf{JUS}}(\mathsf{AF}) &= \{P \in \mathcal{P}(\mathsf{AF}) \mid P(\{\mathcal{A}\}) = P(\{\mathcal{B}\}), P(\{\mathcal{A}\}) + P(\{\mathcal{A}, \mathcal{B}\}) = 0.5\} \\
&= \{P \in \mathcal{P}(\mathsf{AF}) \mid P(\emptyset) = P(\{\mathcal{A}, \mathcal{B}\}) = 0.5 - x, P(\{\mathcal{A}\}) = P(\{\mathcal{B}\}) = x, x \in [0, 0.5]\}
\end{aligned}$$

For $x \in [0, 1]$, let P_x be defined via $P_x(\emptyset) = P_x(\{\mathcal{A}, \mathcal{B}\}) = 0.5 - x$, $P_x(\{\mathcal{A}\}) = P_x(\{\mathcal{B}\}) = x$. So $\mathcal{P}_{\mathsf{JUS}}(\mathsf{AF}) = \{P_x \mid x \in [0, 0.5]\}$. Consider now a partial probability assessment β with $\beta(\mathcal{A}) = \beta(\mathcal{B}) = 1$ (observe that $\beta = \beta_1 \circ \beta_2$) and observe $\mathcal{P}^\beta(\mathsf{AF}) = \{\hat{P}_1\}$ with $\hat{P}_1(\{\mathcal{A}, \mathcal{B}\}) = 1$

and $\hat{P}_1(\emptyset) = \hat{P}_1(\{\mathcal{A}\}) = \hat{P}_1(\{\mathcal{B}\}) = 0$. *It follows*

$$
\begin{aligned}
\mathcal{I}^{d_{KL}}_{\mathsf{JUS}}(\beta, \mathsf{AF}) &= d_{KL}(\mathcal{P}^{\beta}(\mathsf{AF}), \mathcal{P}_{\mathsf{JUS}}(\mathsf{AF})) \\
&= d_{KL}(\hat{P}_1, \{P_x \mid x \in [0, 0.5]\}) \\
&= \min_{x \in [0,0.5]} d_{KL}(\hat{P}_1, P_x) \\
&= \min_{x \in [0,0.5]} \hat{P}_1(\emptyset) \log \frac{\hat{P}_1(\emptyset)}{P_x(\emptyset)} + \hat{P}_1(\{\mathcal{A}\}) \log \frac{\hat{P}_1(\{\mathcal{A}\})}{P_x(\{\mathcal{A}\})} + \hat{P}_1(\{\mathcal{B}\}) \log \frac{\hat{P}_1(\{\mathcal{B}\})}{P_x(\{\mathcal{B}\})} \\
&\qquad + \hat{P}_1(\{\mathcal{A}, \mathcal{B}\}) \log \frac{\hat{P}_1(\{\mathcal{A}, \mathcal{B}\})}{P_x(\{\mathcal{A}, \mathcal{B}\})} \\
&= \min_{x \in [0,0.5]} \hat{P}_1(\{\mathcal{A}, \mathcal{B}\}) \log \frac{\hat{P}_1(\{\mathcal{A}, \mathcal{B}\})}{P_x(\{\mathcal{A}, \mathcal{B}\})} \\
&= \min_{x \in [0,0.5]} 1 \log \frac{1}{1-x} = 1
\end{aligned}
$$

disproving both super-additivity *and* separability*.*

For $\mathcal{I}^{d_p}_{T}$ we have a stronger result as follows.

Proposition 18. *For $p \geq 1$ the function $\mathcal{I}^{d_p}_{T}$ satisfies* consistency *and* monotonicity*. For $p = 1$ the function $\mathcal{I}^{d_p}_{T}$ also satisfies* separability *and* super-additivity*.*

Proof. Consistency follows from Proposition 16 and the proof of *monotonicity* is the same as in Proposition 17.

The proof of *super-additivity* for $p = 1$ is analogous to a similar result by Thimm (2013), Theorem 3.

For $p = 1$ and *separability*, let $CC(\mathsf{AF}) = \{\mathsf{AF}_1, \ldots, \mathsf{AF}_n\}$ and consider probability functions $P_1 \in \mathcal{P}^{\beta}(\mathsf{AF}_1), \ldots, P_n \in \mathcal{P}^{\beta}(\mathsf{AF}_n)$ with $d_1(P_i, \mathcal{P}_T(\mathsf{AF}_i)) = \mathcal{I}^{d_1}_{T}(\beta, \mathsf{AF}_i)$ for $i = 1, \ldots, n$. Construct a probability function P with $P(\mathcal{A}) = P_i(\mathcal{A})$ for all $\mathcal{A} \in \mathsf{Arg}$ and for $\mathcal{A}$ appearing in AF_i (as the AF_i are (maximal) connected components the i is uniquely defined). Observe that $d_1(P, \mathcal{P}_T(\mathsf{AF}_i)) = d_1(P_i, \mathcal{P}_T(\mathsf{AF}_i)) = \mathcal{I}^{d_1}_{T}(\beta, \mathsf{AF}_i)$ as only the values on $\mathcal{A}$ appearing in AF_i are taken into account. For $p = 1$ also observe that, given P_1, P_2, P_3 with $\mathsf{dom}\, P_1 = \mathsf{dom}\, P_2 \cup \mathsf{dom}\, P_3$, and $\mathsf{dom}\, P_2 \cap \mathsf{dom}\, P_3 = \emptyset$ we have ($\circ$ is again functional composition)

$$
\begin{aligned}
d_1(P_1, P_2 \circ P_3) &= \sum_{x \in \mathsf{dom}\, P_1} |P_1(x) - (P_2 \circ P_3)(x)| \\
&= \sum_{x \in \mathsf{dom}\, P_2} |P_1(x) - P_2(x)| + \sum_{x \in \mathsf{dom}\, P_3} |P_1(x) - P_3(x)| \\
&= d_1(P_1, P_2) + d_1(P_1, P_3)
\end{aligned}
$$

It follows

$$\begin{aligned} d_1(P, \mathcal{P}_T(\mathsf{AF})) &= \sum_{i=1}^{n} d_1(P, \mathcal{P}_T(\mathsf{AF}_i)) \\ &= \sum_{i=1}^{n} \mathcal{I}_T^{d_1}(\beta, \mathsf{AF}_i) \end{aligned}$$

yielding $\mathcal{I}_T(\beta, \mathsf{AF}) \leq \sum_{\mathsf{AF}' \in CC(\mathsf{AF})} \mathcal{I}_T(\beta, \mathsf{AF}')$.

It remains to show $\mathcal{I}_T(\beta, \mathsf{AF}) \geq \sum_{\mathsf{AF}' \in CC(\mathsf{AF})} \mathcal{I}_T(\beta, \mathsf{AF}')$. For that let $P \in \mathcal{P}^{\beta}(\mathsf{AF})$ with $d_1(P, \mathcal{P}_T(\mathsf{AF})) = \mathcal{I}_T^{d_1}(\beta, \mathsf{AF})$. Define $P_i \in \mathcal{P}(\mathsf{AF}_i)$ to be the projection of P onto AF_i, in particular $P_i(\mathcal{A}) = P(\mathcal{A})$ for all $\mathcal{A} \in \mathsf{dom}\, P_i$, for all $i = 1, \ldots, n$. It follows $P_i \in \mathcal{P}^{\beta}(\mathsf{AF}_i)$ and therefore $d_1(P, \mathcal{P}_T(\mathsf{AF}_i)) = d_1(P_i, \mathcal{P}_T(\mathsf{AF}_i))$ as before. Observe that $d_1(P_i, \mathcal{P}_T(\mathsf{AF}_i)) \geq \mathcal{I}_T(\beta, \mathsf{AF}_i)$ and due to $P = P_1 \circ \ldots \circ P_n$ we have

$$d_1(P, P') = \sum_{i=1}^{n} d_1(P_i, P')$$

for every probability function P'. It follows $\mathcal{I}_T(\beta, \mathsf{AF}) \geq \sum_{\mathsf{AF}' \in CC(\mathsf{AF})} \mathcal{I}_T(\beta, \mathsf{AF}')$. □

For $p > 1$ a relaxed version of *separability* holds.

Proposition 19. *For $p > 1$ the function $\mathcal{I}_T^{d_p}$ satisfies*

$$\mathcal{I}_T^{d_p}(\beta, \mathsf{AF}) \leq \sum_{\mathsf{AF}' \in CC(\mathsf{AF})} \mathcal{I}_T^{d_p}(\beta, \mathsf{AF}')$$

Proof. We use the same notation as in the proof of Proposition 18. Given P_1, P_2, P_3 with $\mathsf{dom}\, P_1 = \mathsf{dom}\, P_2 \cup \mathsf{dom}\, P_3$, and $\mathsf{dom}\, P_2 \cap \mathsf{dom}\, P_3 = \emptyset$ observe for $p > 1$:

$$\begin{aligned} d_p(P_1, P_2 \circ P_3) &= \sqrt[p]{\sum_{x \in \mathsf{dom}\, P_1} |P_1(x) - (P_2 \circ P_3)(x)|^p} \\ &= \sqrt[p]{\sum_{x \in \mathsf{dom}\, P_2} |P_1(x) - P_2(x)|^p + \sum_{x \in \mathsf{dom}\, P_3} |P_1(x) - P_3(x)|^p} \\ &\leq \sqrt[p]{\sum_{x \in \mathsf{dom}\, P_2} |P_1(x) - P_2(x)|^p} + \sqrt[p]{\sum_{x \in \mathsf{dom}\, P_3} |P_1(x) - P_3(x)|^p} \\ &= d_p(P_1, P_2) + d_p(P_1, P_3) \end{aligned}$$

which generalizes inductively to $\mathcal{I}_T(\beta, \mathsf{AF}) \leq \sum_{\mathsf{AF}' \in CC(\mathsf{AF})} \mathcal{I}_T(\beta, \mathsf{AF}')$. □

Inconsistency measures allow us to evaluate the degree to which a probability assessment deviates from rationality postulates. Determining this degree is important if we are to have systematic mechanisms for handling contradictory probability assessments. The proposals and results in this subsection, show that we have options for measuring inconsistency that have desirable properties.

7.2 Distance-Based Consolidation

The measure $\mathcal{I}_T^d$ allows us to quantitatively assess the violation of a partial probability assignment in the light of a given set of rationality conditions. However, Example 12 suggests that even in the presence of contradictory information, we want to be able to provide reasonable inference results. Following the idea of $\mathcal{I}_T^d$ we define the set of reasonable probability functions as those probability functions in $\mathcal{P}^\beta(\mathsf{AF})$ that are *closest* to satisfying the rationality conditions T.

Definition 11. *Define the set* $\Pi_{T,d,\mathsf{AF}}(\beta) \subseteq \mathcal{P}^\beta(\mathsf{AF})$ *via*

$$\Pi_{T,d,\mathsf{AF}}(\beta) = \{P \in \mathcal{P}^\beta(\mathsf{AF}) \mid d(P, \mathcal{P}_T(\mathsf{AF})) \textit{ is minimal}\}$$

In other words, the set $\Pi_{T,d,\mathsf{AF}}(\beta)$ is the set of "witnesses" of the inconsistency value $\mathcal{I}_T^d(\beta, \mathsf{AF})$, i.e., those probability functions P with $d(P, \mathcal{P}_T(\mathsf{AF})) = \mathcal{I}_T^d(\beta, \mathsf{AF})$ (we will also explore an alternative way of defining this set later in Section 7.3).

Our idea is now to use $\Pi_{T,d,\mathsf{AF}}(\beta)$ in the same way for reasoning as we used $\mathcal{P}_T^\beta(\mathsf{AF})$ in Section 6. In fact, it can be easily seen that under the assumption $\mathcal{P}_T^\beta(\mathsf{AF}) \neq \emptyset$ reasoning with $\Pi_{T,d,\mathsf{AF}}(\beta)$ coincides with our previous approach.

Proposition 20. *If* $\mathcal{P}_T^\beta(\mathsf{AF}) \neq \emptyset$ *then* $\Pi_{T,d,\mathsf{AF}}(\beta) = \mathcal{P}_T^\beta(\mathsf{AF})$ *for every pre-metrical distance measure* d*.*

Proof. If $\mathcal{P}_T^\beta(\mathsf{AF}) = \mathcal{P}_T(\mathsf{AF}) \cap \mathcal{P}^\beta(\mathsf{AF}) \neq \emptyset$ then for all $P \in \mathcal{P}_T^\beta(\mathsf{AF})$ we have that $d(P, \mathcal{P}_T(\mathsf{AF})) = 0$. This is the minimal value d can attain so $P \in \Pi_{T,d,\mathsf{AF}}(\beta)$. Furthermore, for every $P' \in \Pi_{T,d,\mathsf{AF}}(\beta)$ it must then hold $d(P', \mathcal{P}_T(\mathsf{AF})) = 0$ and therefore $P' \in \mathcal{P}_T(\mathsf{AF}) \cap \mathcal{P}^\beta(\mathsf{AF})$. □

Moreover, $\Pi_{T,d,\mathsf{AF}}(\beta)$ is a strict generalization of $\mathcal{P}_T^\beta(\mathsf{AF})$ as it always contains probability functions, even if $\mathcal{P}_T^\beta(\mathsf{AF}) = \emptyset$. Furthermore, $\Pi_{T,d,\mathsf{AF}}(\beta)$ features the same topological properties as $\mathcal{P}_T^\beta(\mathsf{AF})$ if the distance measure d is reasonably chosen, see below.

Proposition 21. $\Pi_{T,d,\mathsf{AF}}(\beta) \neq \emptyset$*.*

Proof. Let L be the grounded labeling and consider P_L with $P_L \sim L$. Then P_L complies with all rationality postulates in $\{\mathsf{RAT}, \mathsf{COH}, \mathsf{SFOU}, \mathsf{FOU}, \mathsf{OPT}, \mathsf{SOPT}, \mathsf{JUS}\}$. Therefore, the set $\mathcal{P}_T(\mathsf{AF})$ is non-empty. The set $\mathcal{P}^\beta(\mathsf{AF})$ is also non-empty, cf. Proposition 12. Therefore, there are probability functions $P_1 \in \mathcal{P}_T(\mathsf{AF})$ and $P_2 \in \mathcal{P}^\beta(\mathsf{AF})$ such that $d(P_1, P_2)$ is finite. It follows $\Pi_{T,d,\mathsf{AF}}(\beta) \neq \emptyset$. □

Proposition 22. *For strictly convex* d *and* $T \subseteq \{\mathsf{COH}, \mathsf{SFOU}, \mathsf{FOU}, \mathsf{OPT}, \mathsf{SOPT}, \mathsf{JUS}\}$ *the set* $\Pi_{T,d,\mathsf{AF}}(\beta)$ *is connected, convex, and closed.*

Proof. This follows from the connectedness, convexity, and closure of $\mathcal{P}_T(\mathsf{AF})$ and $\mathcal{P}^\beta(\mathsf{AF})$ (see Proposition 13) and the fact that minimizing a strictly convex function over convex and closed sets yields again a convex and closed set (Boyd & Vandenberghe, 2004). □

The above statement is true for our examples of distance measures, except for d_1 (the Manhattan distance) which is not strictly convex.

Corollary 1. *For $T \subseteq \{\mathit{COH},\mathit{SFOU},\mathit{FOU},\mathit{OPT},\mathit{SOPT},\mathit{JUS}\}$ and $d \in \{d_{KL}, d_p\}$ (for $p > 1$) we have that $\Pi_{t,d,\mathsf{AF}}(\beta)$ is a connected, convex, and closed set.*

The above results show that $\Pi_{T,d,\mathsf{AF}}(\beta)$ behaves exactly like $\mathcal{P}^{\beta}_{T}(\mathsf{AF})$ (in the topological sense) and is a strict generalization. We therefore extend the notion of *possible probabilities* $\mathsf{p}^{\beta}_{T,\mathsf{AF}}(\mathcal{A})$ to the general case.

Definition 12. *Let d be strictly convex and $T \subseteq \{\mathit{COH},\mathit{SFOU},\mathit{FOU},\mathit{OPT},\mathit{SOPT},\mathit{JUS}\}$. Define*

$$\pi^{\beta,d}_{T,\mathsf{AF}}(\mathcal{A}) = \{P(\mathcal{A}) \mid P \in \Pi_{T,d,\mathsf{AF}}(\beta)\}$$

Observe that by Proposition 20 we have $\pi^{\beta,d}_{T,\mathsf{AF}}(\mathcal{A}) = \mathsf{p}^{\beta}_{T,\mathsf{AF}}(\mathcal{A})$ if $\mathcal{P}^{\beta}_{T}(\mathsf{AF}) \neq \emptyset$ (for every pre-metrical distance measure d).

Example 15. *We continue Example 12 with $T_1 = \{\mathit{COH}\}$ and β'_1 defined by $\beta'_1(\mathcal{S}_1) = 0.7$ and $\beta'_1(\mathcal{S}_2) = 0.4$. For the Euclidean distance d_2 we obtain*

$$\pi^{\beta'_1,d_2}_{T_1,\mathsf{AF}}(\mathcal{I}) \approx [0.0284, 0.383] \qquad\qquad \pi^{\beta'_1,d_2}_{T_1,\mathsf{AF}}(\mathcal{G}) \approx [0.0270, 0.682]$$

which shows that beliefs in both $\mathcal{I}$ and $\mathcal{G}$ can be quite low (due to the conflict in the evidence) but that the belief in $\mathcal{G}$ can be up to 0.682 due to the stronger evidence in $\mathcal{S}_1$ and weaker evidence in $\mathcal{S}_2$.

Similarly as for $\mathcal{P}^{\beta}_{T}(\mathsf{AF})$ we can define reasoning based on maximum entropy on $\Pi_{T,d,\mathsf{AF}}(\beta)$ as follows.

Definition 13. *Let d be strictly convex and $T \subseteq \{\mathit{COH},\mathit{SFOU},\mathit{FOU},\mathit{OPT},\mathit{SOPT},\mathit{JUS}\}$. Define*

$$\mathcal{P}^{\beta,\mathsf{AF},T,d}_{\mathit{ME}} = \arg \max_{Q \in \Pi_{T,d,\mathsf{AF}}(\beta)} H(Q)$$

The validity of the following proposition follows also straightforwardly from our previous results.

Proposition 23. *Let AF be an abstract argumentation framework, β a partial probability assignment, d be strictly convex, and $T \subseteq \{\mathit{COH},\mathit{SFOU},\mathit{FOU},\mathit{OPT},\mathit{SOPT},\mathit{JUS}\}$. The set $\mathcal{P}^{\beta,\mathsf{AF},T,d}_{\mathit{ME}}$ contains exactly one uniquely defined probability function.*

We also write $\mathcal{P}^{\beta,\mathsf{AF},T,d}_{\mathsf{ME}}(\mathcal{A})$ to denote the probability $P(\mathcal{A})$ with $\{P\} = \mathcal{P}^{\beta,\mathsf{AF},T,d}_{\mathsf{ME}}$.

Example 16. *We continue Example 15 with $T_1 = \{\mathit{COH}\}$ and β'_1 defined by $\beta'_1(\mathcal{S}_1) = 0.7$ and $\beta'_1(\mathcal{S}_2) = 0.4$. For the Euclidean distance d_2 we obtain*

$$\mathcal{P}^{\beta'_1,\mathsf{AF},T_1,d_2}_{\mathit{ME}}(\mathcal{I}) \approx 0.3788 \qquad\qquad \mathcal{P}^{\beta'_1,\mathsf{AF},T_1,d_2}_{\mathit{ME}}(\mathcal{G}) \approx 0.4959$$

which gives an ambiguous picture on the innocence or guilt of John (due to the contradictory information), with a slight tendency towards guilt due to the slightly higher belief in $\mathcal{S}_1$.

Distance-based consolidation ensures that we get a probability function that is consistent with our partial probability assessment, and is as near as possible to being consistent with the selected rationality postulates. The results in this subsection ensure that this consolidation is well-behaved and viable.

7.3 An Alternative Point of View

In Definition 11 we defined $\Pi_{T,d,\mathsf{AF}}(\beta)$ to be a subset of probability functions of $\mathcal{P}^{\beta}(\mathsf{AF})$ that are closest to the set $\mathcal{P}_T(\mathsf{AF})$. The decision of defining $\Pi_{T,d,\mathsf{AF}}(\beta)$ like this was based on the need to consider only probability functions that are compliant with our observations but as rational as possible w.r.t. T. Consider now

$$\Pi^{*}_{T,d,\mathsf{AF}}(\beta) = \{P \in \mathcal{P}_T(\mathsf{AF}) \mid d(\mathcal{P}^{\beta}(\mathsf{AF}), P) \text{ is minimal}\}$$

The set $\Pi^{*}_{T,d,\mathsf{AF}}(\beta)$ contains those probability functions in $\mathcal{P}_T(\mathsf{AF})$ that are closest to the set $\mathcal{P}^{\beta}(\mathsf{AF})$, i.e., probability functions that are fully rational w.r.t. T and maximally compliant with our observations. It can be easily seen that $\Pi^{*}_{T,d,\mathsf{AF}}(\beta)$ behaves exactly like $\Pi_{T,d,\mathsf{AF}}(\beta)$ w.r.t. its topological properties.

Proposition 24.

1. *If $\mathcal{P}^{\beta}_{T}(\mathsf{AF}) \neq \emptyset$ then $\Pi^{*}_{T,d,\mathsf{AF}}(\beta) = \Pi_{T,d,\mathsf{AF}}(\beta) = \mathcal{P}^{\beta}_{T}(\mathsf{AF})$ for every pre-metrical distance measure d.*

2. $\Pi^{*}_{T,d,\mathsf{AF}}(\beta) \neq \emptyset$

3. *For strictly convex d and $T \subseteq \{\mathit{COH}, \mathit{SFOU}, \mathit{FOU}, \mathit{OPT}, \mathit{SOPT}, \mathit{JUS}\}$ the set $\Pi^{*}_{T,d,\mathsf{AF}}(\beta)$ is connected, convex, and closed.*

Consequently, we can define reasoning based on $\Pi^{*}_{T,d,\mathsf{AF}}(\beta)$ in the same way as we did on $\Pi_{T,d,\mathsf{AF}}(\beta)$.

If we view the $\Pi_{T,d,\mathsf{AF}}(\beta)$ and $\Pi^{*}_{T,d,\mathsf{AF}}(\beta)$ operators as repair operators, then they offer us two options:

1. $\Pi_{T,d,\mathsf{AF}}(\beta)$ is used when we want to preserve the prior information we have in β but want to get as close as possible to satisfying the rationality constraints in T; and

2. $\Pi^{*}_{T,d,\mathsf{AF}}(\beta)$ is used when we want to impose the rationality constraints we have in T but want to keep as much as possible from the prior information we have in β.

We can regard $\Pi_{T,d,\mathsf{AF}}(\beta)$ as a *soft repair* as it does not satisfy T but gets closer to it, and we can regard $\Pi^{*}_{T,d,\mathsf{AF}}(\beta)$ as *hard repair* as it does satisfy T. So hard repairs ensure conformity with T but at the loss of some of the original information in β, whereas soft repairs ensure no loss of the original information in β, but at the loss of some conformity with T.

The difference between reasoning with $\Pi_{T,d,\mathsf{AF}}(\beta)$ and $\Pi^{*}_{T,d,\mathsf{AF}}(\beta)$ is similar to the difference between revision and update in belief dynamics (Katsuno & Mendelzon, 1991). Let us recall a simple example from Katsuno & Mendelzon (1991) to illustrate the difference between revision and update.

Example 17. *There are two objects in a closed room, a magazine and a book, which can be either on the floor or on the table. Let m (b) denote the fact that the magazine (book) is on the floor and $\neg m$ ($\neg b$) that it is on the table. We know that one of the two objects is on the floor and the other on the table, but not which one ($m \Leftrightarrow \neg b$). Now we send a robot* *into the room to observe the situation and tell us whether the book is on the floor or on the table. The robot tells us b and accordingly we* revise *our beliefs to* $b \wedge \neg m$ *(as one object had to be on the table we can infer that it must be the magazine). In an alternative scenario, we send the robot into the room and order it to put the book on the floor, no matter where it was before. After the successful mission the robot informs us of b and we* update *our beliefs to b (our initial constraint that exactly one object is on the floor may not be true any more and we do not know the location of the magazine).*

Kern-Isberner & Rödder (2004) made a similar distinction to differentiate revision and update in a probabilistic setting, also under maximum entropy reasoning. More specifically, let P_0 be a uniform probability function, modelling a completely ignorant prior belief state, and let R be the set of initial probabilistic beliefs (formulas or rules quantified by probabilities). The initial epistemic state of the agent is then defined as $P_1 = P_0 * R$, which amounts to completing the incomplete beliefs in R by selecting the probability function P_1 with maximum entropy (in the case that P_0 is a uniform distribution). Revision by a set S of new beliefs is then defined via $P_2 = P_0 * (R \cup S)$, under the assumption that S is consistent with R (as both refer to the same world and model *uncertain* information to begin with). So revision amounts to completing the initial beliefs with the help of the new information S. Note that the initial beliefs R are not given up, but extended by the new information S. Furthermore, Kern-Isberner & Rödder define update by new information S via $P'_2 = (P_0 * R) * S$, which amounts to completing the beliefs in S in such a way that the resulting probability function P'_2 is as closest as possible to the initial epistemic state $P_1 = P_0 * R$ (using, essentially, also d_{KL} for measuring distances). In particular, it may be the case that P'_2 is no longer compatible with the initial beliefs R.

Let us come back to our scenario of probabilistic abstract argumentation. The partial probability assignment β can be regarded as our initial beliefs. Reasoning with $\Pi_{T,d,\mathsf{AF}}(\beta)$ then resembles *revising* those beliefs with the topological information of the argument graph, while retaining the original beliefs. Furthermore, reasoning with $\Pi^*_{T,d,\mathsf{AF}}(\beta)$ resembles *updating* those beliefs with the topological information of the argument graph. This amounts to interpreting the observation of the argumentation graph as an action, the arguments, which might be unrelated before, are now put into the context of an argumentation framework. Let us consider an example to illustrate this.

Example 18. *In analogy to Example 17 consider a room with three objects* a, b, c *which can be either on the floor or on the table. Due to previous observations delivered to us by a robot we assess* $\beta(a) = 0.7$ *and* $\beta(c) = 0.1$*, meaning that a is likely on the floor and c is likely on the table. We send the robot again into the room to provide more information and it gives us the following two observations: if a is on the floor then b is not on the floor, if b is on the floor then c is not on the floor. These observations can be modelled as the argumentation framework depicted in Figure 7. We* revise *our beliefs accordingly and derive that the probability of b being on the floor is* 0.6[4]*, assuming that our previous information is valid. In another scenario, we send the robot into the room and order it to arrange through a minimal number of moves a scenario, where both statements "if a is on the floor then b is*

4. We use the Euclidean distance d_2 and set $T = \{\mathsf{COH}, \mathsf{SOPT}\}$; it can be seen that for the attack (a, b) the value 0.3 would be derived for b and for the attack (b, c) the value 0.9 would be derived for b; roughly, the Euclidean distance is minimal for the value $(0.9 + 0.3)/2 = 0.6$.

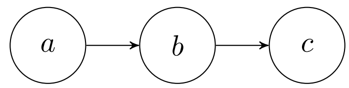

Figure 7: The argumentation framework from Example 18

not on the floor’ and “if b is on the floor then c is not on the floor” are true. Accordingly, we update *our beliefs to this new information. In particular, it is now impossible that a is on the floor and c is on the table (which was deemed quite likely in our initial beliefs). It can be seen that the assignment of* 0.4 *to both a and c, and* 0.6 *to b minimizes the distance to our previous beliefs, while satisfying both conditions* COH *and* SOPT.

So both distance-based consolidation or the alternative presented in this subsection are well-behaved and viable proposals. Which we choose to use depends on the priorities of the application. We leave a deeper investigation of the relationships to belief dynamics for future work.

8. Related Works

In the following, we discuss some works addressing similar topics as our work. In particular, we review other approaches to probabilistic (abstract) argumentation in Section 8.1, similar frameworks for quantitative uncertainty in Section 8.2, and ranking-based semantics for argumentation in 8.3.

8.1 Probabilistic Argumentation

The two main approaches to probabilistic (abstract) argumentation are the constellations and the epistemic approaches (Hunter, 2013).

- In the constellations approach, the uncertainty is in the topology of the graph (Dung & Thang, 2010; Li, Oren, & Norman, 2011; Hunter, 2012; Fazzinga, Flesca, & Parisi, 2013; Li, Oren, & Norman, 2013; Hunter & Thimm, 2014a; Dondio, 2014; Polberg & Doder, 2014; Doder & Woltran, 2014; Fazzinga, Flesca, & Parisi, 2015; Liao & Huang, 2015; Hadoux, Beynier, Maudet, Weng, & Hunter, 2015; Sun & Liao, 2016; Fazzinga et al., 2016). As an example, this approach is useful when one agent is not sure what arguments and attacks another agent is aware of, and so this can be captured by a probability distribution over the space of possible argument graphs.
- In the epistemic approach, the topology of the argument graph is fixed, but there is uncertainty about whether an argument is believed (Thimm, 2012; Hunter, 2013; Baroni, Giacomin, & Vicig, 2014; Hunter, 2014b, 2014a; Hunter & Thimm, 2014d, 2014c, 2014b; Hunter, 2015; Gabbay & Rodrigues, 2015; Hunter, 2016a, 2016b). A core idea of the epistemic approach is that the more likely an agent is to believe in an argument, the less likely it is to believe in an argument attacking it.

This paper provides a comprehensive account of the epistemic approach and we additionally considered the case of incomplete and inconsistent probability distributions. These problems were first raised by Hunter (2013) and Hunter & Thimm (2014d), but no systematic solutions to the problems were presented. In contrast in this paper, we have provided solutions based on well-justified notions of distance between probability distributions.

In quantifying disagreement between argument graphs, the distance between labellings has been considered in terms of the weighted sum of the number of labellings that differ (Booth, Caminada, Podlaszewski, & Rahwan, 2012). Various kinds of distance have also been considered in methods for epistemic enforcement in abstract argumentation (Baumann & Brewka, 2010; Baumann, 2012; Coste-Marquis, Konieczny, Maily, & Marquis, 2014c), for revising argument graphs (Coste-Marquis, Konieczny, & Maily, 2014a, 2014b), and for merging argument graphs (Coste-Marquis, Devred, Konieczny, Lagasquie-Schiex, & Marquis, 2007; Delobelle, Konieczny, & Vesic, 2015). There are related proposals for belief revision in argumentation (Cayrol, de Saint-Cyr, & Lagasquie-Schiex, 2010; Gabbay & Rodrigues, 2012; Bisquert, Cayrol, de Saint-Cyr, & Lagasquie-Schiex, 2013; Diller, Haret, Linsbichler, Rümmele, & Woltran, 2015) but they do not use distance measures.

Dung and Thang (2010) provided the first proposal to extend abstract argumentation with a probability distribution over sets of arguments which they use with a version of assumption-based argumentation in which a subset of the rules are probabilistic rules. Another approach to augmenting abstract argumentation with probabilities has used equations based on the structure of the graph to constrain the probability assignments, and these can be solved to calculate probabilities (Gabbay & Rodrigues, 2015). In another rule-based system for argumentation, the belief in the premises of an argument is used to calculate the belief in the argument (Riveret, Rotolo, Sartor, Prakken, & Roth, 2007). However, the proposal does not investigate further the nature of this assignment, for example with respect to abstract argumentation, but rather its use in dialogue is explored. In a logic-based approach, Verheij combines qualitative reasoning in terms of reasons and defeaters, with quantitative reasoning using argument strength which is modelled as the conditional probability of the conclusions given the premises (Verheij, 2014).

The epistemic approach to probabilistic argumentation has been used for logical arguments (Hunter, 2013). Each argument is a pair $\langle \Phi, \alpha \rangle$ where Φ entails α using classical logic. Uncertainty over the arguments is represented by a probability distribution over the models of the logical language. Then the probability of an argument is the probability of the premises being satisfied which is calculated as the sum of the probability of the models satisfying the premises. This can then be treated as an instantiation of the framework presented in this paper.

Another approach for probabilistic argumentation for logical arguments is the ABEL framework where reasoning with propositional information is augmented with probabilistic information so that individual arguments are qualified by a probability value (Haenni, 1998; Haenni, Kohlas, & Lehmann, 2000). The emphasis is on generating pros and cons for diagnosis. However, there is no consideration in ABEL of how this probabilistic information relates to Dung's proposals, or how it could be used to decide which arguments are acceptable according to Dung's dialectical semantics.

Probabilistic reasoning with logical statements has also been considered by Pollock (1995). However, the approach taken is to assign probabilities to formulae without considering the meaning of this in terms of models. Various issues arising from an assignment based on the frequency that a consequent holds when the antecedent holds are considered, as well as how such an assignment could be used for statistical syllogism. The emphasis of the work is therefore different as it does not consider what would be acceptable probability assignments for a language, and it does not consider how a probabilistic perspective relates to abstract argumentation.

Pollock (2001) has also expounded a non-probabilistic account (i.e. not conforming to probability calculus) of degrees of justification in argument-based defeasible reasoning. This is in part intended to account for attacking arguments that are too weak to defeat an inference but may nevertheless diminish the degree of justification of its conclusion, and for the way attacking arguments can be aggregated so that together they are stronger enough to defeat an inference.

There are other approaches to bringing probability theory into systems for dialogical argumentation. A probabilistic model of the opponent has been used in a dialogue strategy allowing the selection of moves for an agent based on what it believes the other agent is aware of and the moves it might take (Rienstra, Thimm, & Oren, 2013). In another approach to probabilistic opponent modelling, the history of previous dialogues is used to predict the arguments that an opponent might put forward (Hadjinikolis, Siantos, Modgil, Black, & McBurney, 2013). For modelling the possible dialogues that might be generated by a pair of agents, a probabilistic finite state machine can represent the possible moves that each agent can make in each state of the dialogue assuming a set of arguments that each agent is aware of (Hunter, 2014b). This has been generalised to POMDPs when there is uncertainty about what an opponent is aware of (Hadoux et al., 2015).

Some research has investigated relationships between Bayesian networks and argumentation. Bayesian networks can be used to model argumentative reasoning with arguments and counterarguments (Vreeswijk, 2004). In a similar vein, Bayesian networks can be used to capture aspects of argumentation in the Carneades model where the propagation of argument applicability and statement acceptability can be expressed through conditional probability tables (Grabmair, Gordon, & Walton, 2010). Going the other way, arguments can be generated from a Bayesian network, and this can be used to explain the Bayesian network (Timmer, Meyer, Prakken, Renooij, & Verheij, 2015), and argumentation can be used to combine multiple Bayesian networks (Nielsen & Parsons, 2007).

8.2 Other Quantitative Approaches for Argumentation

There are works incorporating other frameworks of quantitative uncertainty into argumentation. For example, Janssen, Cock, & Vermeir (2008) extend abstract argumentation by allowing the attack relation to be a *fuzzy relation*. In another fuzzy approach to argumentation (da Costa Pereira, Tettamanzi, & Villata, 2011), each argument is assigned a fuzzy value (i.e. a value in the unit interval) R to denote the reliability of the source of the argument and a fuzzy label L (i.e. a value in the unit interval) to denote the degree of acceptability of the argument. Two constraints are assumed on the fuzzy labelling for an argument $\mathcal{A}$ with the latter constraint being a fuzzy reformulation of Caminada's definition for labelling.

There is also a correspondence between our definition of the probability of an argument and the notion of plausibility in Dempster-Shafer theory (Shafer, 1976). Dempster-Shafer theory assumes a frame of discernment Ω that is a set of atoms (e.g. a set of hypotheses for a diagnosis). Each subset of Ω is regarded as a proposition. A mass distribution m is a function from 2^Ω into $[0,1]$ such that $m(\emptyset) = 0$ and $\sum_{A\subseteq\Omega} m(A) = 1$. To obtain the total belief in a set $A \subseteq \Omega$ (i.e. the extent to which all available evidence supports A) we need to sum all the mass for all subsets of A (i.e. propositions that imply A). Hence, a belief function $\mathsf{Bel} : 2^\Omega \to [0,1]$ is defined via $\mathsf{Bel}(A) = \sum_{B\subseteq A} m(B)$. Note the remaining evidence need not refute A (i.e., support the complement $A^c = \Omega \setminus A$) since the above does not imply $Bel(A) + Bel(A^c) = 1$. So some of the remaining evidence may be assigned to propositions not disjoint from A, and hence could plausibly be transferred directly to A in the light of further evidence. This then leads to the following notion of the plausibility of a proposition. A plausibility function $Pl : 2^\Omega \to [0,1]$ is then defined via $\mathsf{Pl}(A) = 1 - \mathsf{Bel}(A^c) = \sum_{B\cap A\neq\emptyset} m(B)$. So the plausibility of a proposition A is the sum of the mass assigned to propositions that are consistent with A. The constraint $B \cap A \neq \emptyset$ is weaker than $B \subseteq A$, so $\mathsf{Bel}(A) \leq \mathsf{Pl}(A)$.

Let $\mathsf{AF} = (\mathsf{Arg}, \to)$ be an abstract argumentation framework. Now if we consider $\Omega = \mathsf{Arg}$, then we have the following for each $\mathcal{A} \in \mathsf{Arg}$, where $\mathsf{Pl}(\{\mathcal{A}\})$ is known as the contour function.

$$\mathsf{Bel}(\{\mathcal{A}\}) = P(\{\mathcal{A}\}) \qquad\qquad \mathsf{Pl}(\{\mathcal{A}\}) = \sum_{X\subseteq\mathsf{Arg} \text{ s.t. } \mathcal{A}\in X} P(X) = P(\mathcal{A})$$

But a better alternative is to regard the power set of Arg as the frame of discernment since we regard each element of the power set to be disjoint from the other elements of the power set. Then for an argument $\mathcal{A} \in \mathsf{Args}$, we can define a proposition $\alpha_\mathcal{A}$ to be the maximal subset of the frame of discernment such that each element in that subset entails $\mathcal{A}$. So $\alpha_\mathcal{A}$ is the proposition denoting that argument $\mathcal{A}$ is accepted.

For example, suppose we have $\mathsf{Arg} = \{\mathcal{A}, \mathcal{B}\}$, and suppose we represent each subset of Arg by a two digit binary number where 11 denotes $\{\mathcal{A}, \mathcal{B}\}$, 10 denotes $\{\mathcal{A}\}$, 01 denotes $\{\mathcal{B}\}$, and 00 denotes $\emptyset$. Then the frame of discernment is $\Omega = \{11, 10, 01, 00\}$ and the maximal subset of Ω that entails $\mathcal{A}$ is $\{11, 10\}$. So we define $\alpha_\mathcal{A}$ to be $\{11, 10\}$. Assuming a mass distribution over Ω, the belief in the proposition $\alpha_\mathcal{A}$ is obtained as usual (using $\mathsf{Bel}(\alpha_\mathcal{A}) = \Sigma_{B\subseteq\alpha_\mathcal{A}} m(B)$) as follows.

$$\mathsf{Bel}(\alpha_\mathcal{A}) = m(\{11\}) + m(\{10\}) + m(\{11, 10\})$$

And the plausibility in proposition $\alpha_\mathcal{A}$ is obtained as usual (using $\mathsf{Pl}(\alpha_\mathcal{A}) = \Sigma_{B\cap\alpha_\mathcal{A}\neq\emptyset}\, m(B)$)) to give $\mathsf{Pl}(\alpha_\mathcal{A})$ as the following summation.

$$\begin{aligned}\mathsf{Pl}(\alpha_\mathcal{A}) = {} & m(\{11\}) + m(\{11, 01\}) + m(\{11, 00\}) + m(\{11, 01, 00\}) \\ & + m(\{10\}) + m(\{10, 01\}) + m(\{10, 00\}) + m(\{10, 01, 00\}) \\ & + m(\{11, 10\}) + m(\{11, 10, 01\}) + m(\{11, 10, 00\}) + m(\{11, 10, 01, 00\})\end{aligned}$$

Then if we use our probability function P over the subsets of Arg that we have used in the rest of this paper, we can see that it defines the mass distribution over singleton sets. For the above example, it means that we would have $m(\{11\}) = P(11)$, $m(\{10\}) = P(10)$, $m(\{01\}) = P(01)$, and $m(\{00\}) = P(00)$. Furthermore, for all non singleton sets the assignment given by our probability function is zero (e.g. $m(\{11, 10\}) = 0$). So $m(\{11\}) + m(\{10\}) + m(\{01\}) + m(\{00\}) = 1$.

These encodings in Dempster-Shafer theory of our epistemic approach to probabilistic abstract argumentation point to further ways that the idea of epistemic extensions can be developed. As the above example shows, we could consider the benefits of having a mass distribution over non-singleton sets. These benefits could include the suspension of the excluded middle which may be useful for representing ignorance. We could also consider the use of Dempster's rule of combination for combining mass from multiple sources.

8.3 Ranking-Based Argumentation

Recently, there has been some attention to *qualitative* approaches of ranking arguments, see Bonzon, Delobelle, Konieczny, & Maudet (2016b) for a recent survey and the works of Amgoud & Ben-Naim (2013, 2015, 2016), Amgoud, Ben-Naim, Doder, & Vesic (2016), Grossi & Modgil (2015), Matt & Toni (2008), Cayrol & Lagasquie-Schiex (2005), Thimm & Kern-Isberner (2014), and Bonzon, Delobelle, Konieczny, & Maudet (2016a) for some concrete approaches. These approaches aim at deriving a preference relation $\succeq_{\mathsf{AF}}$ among arguments of a framework AF by exploiting the topology of the argumentation graph in a particular way. Here, $\mathcal{A} \succeq_{\mathsf{AF}} \mathcal{B}$ means that $\mathcal{A}$ is as least as preferred as $\mathcal{B}$. Many of these approaches rely on numerical evaluations, e.g., the score $S_{\mathsf{AF}}(\mathcal{A})$ of an argument $\mathcal{A}$ in the approach of Amgoud (2016) is given by the unique solution to the system of equations given via

$$S_{\mathsf{AF}}(\mathcal{A}) = 1 + \sqrt[\alpha]{\sum_{\mathcal{B}\to\mathcal{A}} \frac{1}{S(\mathcal{B})^{\alpha}}}$$

with some fixed parameter $\alpha \in (0,\infty)$ for all arguments $\mathcal{A} \in \mathsf{Arg}$. The intuition implemented in the above equation is that the lower (stronger) the score of the attackers of $\mathcal{A}$ the larger (weaker) the score of $\mathcal{A}$. The obtained values do not usually have a specific meaning such as probabilities or fuzzy values but are used to specify a ranking $\succeq_{\mathsf{AF}}$ via $\mathcal{A} \succeq_{\mathsf{AF}} \mathcal{B}$ iff $S_{\mathsf{AF}}(\mathcal{A}) \leq S_{\mathsf{AF}}(\mathcal{B})$ (let $\succ_{\mathsf{AF}}$ and $\cong_{\mathsf{AF}}$ be defined accordingly). In general, approaches to ranking-based argumentation determine the strength of an argument by aggregating the strengths of its attackers, and possibly its defenders and their attackers, and so on.

Research in ranking-based semantics is also driven by postulates, see the work of Bonzon et al. (2016b) for an overview. We recall some of these now and consider them in our context of probabilistic abstract argumentation. As we will see, our approach is, in general compatible with the general view on ranking-based semantics but is based on different foundations. Let $\succeq_{\mathsf{AF}}$ be a preorder on the set Arg of an argumentation framework $\mathsf{AF} = (\mathsf{Arg},\to)$ and $\mathcal{A},\mathcal{B} \in \mathsf{Arg}$.

Abstraction If $\mathsf{AF} = (\mathsf{Arg},\to)$ and $\mathsf{AF}' = (\mathsf{Arg}',\to')$ are isomorphic and $\gamma : \mathsf{Arg} \to \mathsf{Arg}'$ be an isomorphism then $\mathcal{A} \succeq_{\mathsf{AF}} \mathcal{B}$ iff $\gamma(\mathcal{A}) \succeq_{\mathsf{AF}'} \gamma(\mathcal{B})$.

The above postulate demands that only the topology of a graph and not the names of arguments should influence the ranking. Our approach satisfies this general principle as well. However, we need to rephrase this property slightly in order to be applicable. For argumentation frameworks $\mathsf{AF} = (\mathsf{Arg},\to)$ and $\mathsf{AF}' = (\mathsf{Arg}',\to')$, a bijective function $\gamma : \mathsf{Arg} \to \mathsf{Arg}'$, and a probability function $P \in \mathcal{P}(\mathsf{AF})$ define $\gamma(P) \in \mathcal{P}(\mathsf{AF}')$ via $\gamma(P)(\mathcal{A}) = P(\gamma^{-1}(\mathcal{A}))$ for all $\mathcal{A} \in \mathsf{Arg}'$. For a set $\mathcal{P} \subseteq \mathcal{P}(\mathsf{AF})$ abbreviate $\gamma(\mathcal{P}) = \{\gamma(P) \mid P \in \mathcal{P}\}$.

Proposition 25. *Let* $\mathsf{AF} = (\mathsf{Arg}, \rightarrow)$ *and* $\mathsf{AF}' = (\mathsf{Arg}', \rightarrow')$ *be isomorphic and* $\gamma : \mathsf{Arg} \rightarrow \mathsf{Arg}'$ *be an isomorphism. For every* $T \subseteq \{$*COH,SFOU, FOU, SOPT, OPT, JUS, TER, RAT, NEU, INV, MAX, MIN*$\}$, $\gamma(\mathcal{P}_T(\mathsf{AF})) = \mathcal{P}_T(\mathsf{AF}')$.

Proof. We only consider $T = \{\mathsf{COH}\}$, all other properties and combinations thereof work analogously. Let $P \in \mathcal{P}_T(\mathsf{AF})$. Recall that for $\mathcal{A}, \mathcal{B} \in \mathsf{Arg}$ we have $\mathcal{A} \rightarrow \mathcal{B}$ iff $\gamma(\mathcal{A}) \rightarrow' \gamma(\mathcal{B})$. Then for all $\mathcal{A}, \mathcal{B} \in \mathsf{Arg}'$ with $\mathcal{A} \rightarrow \mathcal{B}$,

$$\gamma(P)(\mathcal{A}) = P(\gamma^{-1}(\mathcal{A})) \leq 1 - P(\gamma^{-1}(\mathcal{B})) = 1 - \gamma(P)(\mathcal{B})$$

Note that the inequality holds as P is coherent and $\gamma^{-1}(\mathcal{A}) \rightarrow \gamma^{-1}(\mathcal{B})$. So $\gamma(P) \in \mathcal{P}_T(\mathsf{AF}')$ and therefore $\gamma(\mathcal{P}_T(\mathsf{AF})) \subseteq \mathcal{P}_T(\mathsf{AF}')$. The other direction is analogous. $\square$

Another property for ranking-based semantics is the following. Recall that for an argumentation framework AF we denote by $CC(\mathsf{AF})$ the set of simply connected components of AF.

Independence For all $\mathsf{AF}' = (\mathsf{Arg}', \rightarrow') \in CC(\mathsf{AF})$, for all $\mathcal{A}, \mathcal{B} \in \mathsf{Arg}'$, $\mathcal{A} \succ_{\mathsf{AF}'} \mathcal{B}$ implies $\mathcal{A} \succ_{\mathsf{AF}} \mathcal{B}$.

Independence demands that the ranking between two arguments should not be influenced by arguments disconnected from these two. In our context we can phrase this property as follows. Let $\mathsf{AF} = (\mathsf{Arg}, \rightarrow)$ and $\mathsf{AF}' = (\mathsf{Arg}', \rightarrow')$ with $\mathsf{Arg}' \subseteq \mathsf{Arg}$ and $\rightarrow' = \rightarrow \cap (\mathsf{Arg}' \times \mathsf{Arg}')$. For a probability function $P \in \mathcal{P}(\mathsf{AF})$ define $P|_{\mathsf{AF}'} \in \mathcal{P}(\mathsf{AF}')$ via

$$P|_{\mathsf{AF}'}(E) = \sum_{E' \subseteq \mathsf{Arg} \setminus \mathsf{Arg}'} P(E \cup E')$$

for all $E \subseteq \mathsf{Arg}$. Observe that $P|_{\mathsf{AF}'}(\mathcal{A}) = P(\mathcal{A})$ for all $\mathcal{A} \in \mathsf{Arg}'$, so $P|_{\mathsf{AF}'}$ is the projection of P on AF'. For a set $\mathcal{P} \subseteq \mathcal{P}(\mathsf{AF})$ let $\mathcal{P}|_{\mathsf{AF}'} = \{P|_{\mathsf{AF}'} \mid P \in \mathcal{P}\}$.

Proposition 26. *For all* $\mathsf{AF}' = (\mathsf{Arg}', \rightarrow') \in CC(\mathsf{AF})$*, for all* $T \subseteq \{$*COH,SFOU, FOU, SOPT, OPT, JUS, TER, RAT, NEU, INV, MAX, MIN*$\}$, $\mathcal{P}_T(\mathsf{AF})|_{\mathsf{AF}'} = \mathcal{P}_T(\mathsf{AF}')$.

Proof. Again, we will only consider $T = \{\mathsf{COH}\}$, all other properties and combinations thereof work analogously. Let $P \in \mathcal{P}_T(\mathsf{AF})$ and $\mathcal{A}, \mathcal{B} \in \mathsf{Arg}'$ with $\mathcal{A} \rightarrow \mathcal{B}$, then

$$P|_{\mathsf{AF}'}(\mathcal{A}) = P(\mathcal{A}) \leq 1 - P(\mathcal{B}) = 1 - P|_{\mathsf{AF}'}(\mathcal{B})$$

So $P|_{\mathsf{AF}'} \in \mathcal{P}_T(\mathsf{AF}')$ and $\mathcal{P}_T(\mathsf{AF})|_{\mathsf{AF}'} \subseteq \mathcal{P}_T(\mathsf{AF}')$. For the other direction note that every coherent probability function on AF' can be extended to a coherent probability function on AF by assigning probability 0.5 all arguments in $\mathsf{Arg} \setminus \mathsf{Arg}'$. $\square$

The above result implies that if $P'(\mathcal{A}) > P'(\mathcal{B})$ for all $P' \in \mathcal{P}_T(\mathsf{AF}')$ then $P(\mathcal{A}) > P(\mathcal{B})$ for all $P \in \mathcal{P}_T(\mathsf{AF})$ as well. Some postulates for ranking-based semantics have direct counterparts in our scenario. For example, consider

Void Precedence If $\mathsf{Att}_{\mathsf{AF}}(\mathcal{A}) = \emptyset$ and $\mathsf{Att}_{\mathsf{AF}}(\mathcal{B}) \neq \emptyset$ then $\mathcal{A} \succ_{\mathsf{AF}} \mathcal{B}$.

This postulate corresponds in its non-strict version to our FOU property.

Proposition 27. *If* $\mathsf{Att}_{\mathsf{AF}}(\mathcal{A}) = \emptyset$ *and* $\mathsf{Att}_{\mathsf{AF}}(\mathcal{B}) \neq \emptyset$ *and* $P \in \mathcal{P}_{\mathsf{FOU}}(\mathsf{AF})$ *then* $P(\mathcal{A}) \geq P(\mathcal{B})$*.*

Proof. From $\mathsf{Att}_{\mathsf{AF}}(\mathcal{A}) = \emptyset$ and $P \in \mathcal{P}_{\mathsf{FOU}}(\mathsf{AF})$ it follows $P(\mathcal{A}) = 1$. □

Note, however, that we do not obtain $P(\mathcal{A}) > P(\mathcal{B})$ in general. However, our approach obviously satisfies *Non-attacked Equivalence* which demands that all non-attacked arguments have the same rank.

Non-attacked Equivalence If $\mathsf{Att}_{\mathsf{AF}}(\mathcal{A}) = \emptyset$ and $\mathsf{Att}_{\mathsf{AF}}(\mathcal{B}) = \emptyset$ then $\mathcal{A} \cong_{\mathsf{AF}} \mathcal{B}$.

Proposition 28. *If* $\mathsf{Att}_{\mathsf{AF}}(\mathcal{A}) = \emptyset$ *and* $\mathsf{Att}_{\mathsf{AF}}(\mathcal{B}) = \emptyset$ *and* $P \in \mathcal{P}_{\mathsf{FOU}}(\mathsf{AF})$ *then* $P(\mathcal{A}) = P(\mathcal{B})$*.*

Another postulate is concerned with *attack branches*. An attack branch for $\mathcal{A}$ is a sequence of arguments $\mathcal{B}_1, \ldots, \mathcal{B}_n$ with $\mathsf{Att}_{\mathsf{AF}}(\mathcal{B}_1) = \emptyset$, $\mathcal{B}_i \rightarrow \mathcal{B}_{i+1}$ for $i = 1, \ldots, n-1$, $\mathcal{B}_n \rightarrow \mathcal{A}$, and n is odd.

Attack vs. Full Defense If $\mathcal{A}$ has no attack branch, $\mathsf{Att}_{\mathsf{AF}}(\mathcal{C}) = \emptyset$, and $\mathsf{Att}_{\mathsf{AF}}(\mathcal{B}) = \{\mathcal{C}\}$, then $\mathcal{A} \succ_{\mathsf{AF}} \mathcal{B}$.

This property is satisfied in its non-strict version as well, if we consider at least the properties FOU and COH.

Proposition 29. *If* $\mathcal{A}$ *has no attack branch,* $\mathsf{Att}_{\mathsf{AF}}(\mathcal{C}) = \emptyset$*, and* $\mathsf{Att}_{\mathsf{AF}}(\mathcal{B}) = \{\mathcal{C}\}$*, then for all* $P \in \mathcal{P}_{\{\mathsf{FOU},\mathsf{COH}\}}(\mathsf{AF})$*,* $P(\mathcal{A}) \geq P(\mathcal{B})$*.*

Proof. For $P \in \mathcal{P}_{\{\mathsf{FOU},\mathsf{COH}\}}(\mathsf{AF})$ we have $P(\mathcal{C}) = 1$ due to FOU and then $P(\mathcal{B}) = 0$ due to COH. □

Bonzon et al. (2016b) discuss a series of further postulates which are, in general, not satisfied by epistemic probabilistic abstract argumentation. The main difference between these two families of approach is that ranking-based approaches aim at incorporating some form of *accrual* (Prakken, 2005) into the evaluation process while epistemic probabilistic abstract argumentation aims at being as close as possible to classical abstract argumentation in this regard. To be more precise, one postulate that is satisfied by all ranking-based approaches considered by Bonzon et al. (2016b) is "Addition of Attack Branch", which states that adding an attack branch to an argument decreases its ranking. For example, an argument attacked by n unattacked arguments should be ranked higher than an argument attacked by $n+1$ unattacked arguments. Classical abstract argumentation (Dung, 1995) does not distinguish these two cases, both arguments will be labeled "out" in any reasonable labelling. Furthermore, any probability function P which satisfies at least FOU and COH will assign probability zero to both arguments as well. So, although similar in spirit, these two families of approaches rely on different foundational assumptions.

9. Summary

The epistemic approach provides a finer grained assessment of an argument graph than given by the basic notions of extensions. With labellings, arguments are labelled as in, out, or undec, whereas with the epistemic approach an argument can take any value in $[0, 1]$. By adopting constraints on the probability distribution, we have shown how the epistemic approach subsumes Dung's approach. However, we have also argued that there is a need for a view where we adopt weaker constraints on the probability distribution. For instance, an important aspect of the epistemic approach is the representation of disbelief in arguments even when they are unattacked. It is not always possible or practical to identify a counterargument to reject in argumentation, and often it is quite natural to directly represent the disbelief in an argument without consideration of the counterargument.

The epistemic approach is also useful for modelling the belief that an opponent might have in the arguments that could be presented, which is useful for example when deciding on the best arguments to present in order to persuade that opponent. Strategies in dialogical argumentation are an important research issue (Thimm, 2014). By harnessing a model of the beliefs of opponent, better choices can be made by an agent (see for example Hunter, 2015).

We also considered incomplete probability distributions and probability distributions that are inconsistent with a set of constraints. These issues commonly arise when considering multiple agents. For instance, when using a probability distribution to represent the beliefs of an opponent, the opponent may have made explicit its beliefs in specific arguments (perhaps by positing them, or by answering queries regarding them). Normally, what is known about the beliefs of the opponent will be incomplete. To give an example of dealing with inconsistency, we can use the probability distribution to represent the feedback obtained from an audience of a television debate. Here, the probability distribution might be inconsistent with the chosen constraints. If we assume that the audience does conform to the constraints, and that probability distribution fails to satisfy the constraints, then we can "repair" the probability distribution, using our approaches of "soft" and "hard repair".

Acknowledgements

We thank Sylwia Polberg for making us aware of an erroneous result in a previous version of this paper.

References

Amgoud, L., & Ben-Naim, J. (2013). Ranking-based semantics for argumentation frameworks. In Liu, W., Subrahmanian, V. S., & Wijsen, J. (Eds.), *Proceedings of the 7th International Conference on Scalable Uncertainty Management (SUM'13)*, Vol. 8078 of *Lecture Notes In Artificial Intelligence*, Washington, DC, USA.

Amgoud, L., & Ben-Naim, J. (2015). Argumentation-based ranking logics. In *Proceedings of the 14th International Conference on Autonomous Agents and Multiagent Systems (AAMAS'15)*.

Amgoud, L., & Ben-Naim, J. (2016). Axiomatic foundations of acceptability semantics. In *Proceedings of the 15th International Conference on Principles of Knowledge Representation and Reasoning (KR'16)*.

Amgoud, L., Ben-Naim, J., Doder, D., & Vesic, S. (2016). Ranking arguments with compensation-based semantics. In *Proceedings of the 15th International Conference on Principles of Knowledge Representation and Reasoning (KR'16)*.

Baroni, P., Caminada, M., & Giacomin, M. (2011). An introduction to argumentation semantics. *The Knowledge Engineering Review*, *26*(4), 365–410.

Baroni, P., Giacomin, M., & Vicig, P. (2014). On rationality conditions for epistemic probabilities in abstract argumentation. In *Proceedings of the Fifth International Conference on Computational Models of Argumentation (COMMA'14)*.

Baumann, R. (2012). What does it take to enforce an argument? Minimal change in abstract argumentation. In *Proceedings of the 20th European Conference on Artificial Intelligence (ECAI'12)*, pp. 127–132.

Baumann, R., & Brewka, G. (2010). Expanding argumentation frameworks: enforcing and monotoncity results. In *Proceedings of the Third International Conference on Computational Models of Argumentation (COMMA'10)*, pp. 75–86.

Besnard, P., & Hunter, A. (2008). *Elements of Argumentation*. The MIT Press.

Bex, F., & Renooij, S. (2016). From arguments to constraints on a Bayesian network. In *Proceedings of the Sixth International Conference on Computational Models of Argumentation (COMMA'16)*.

Bisquert, P., Cayrol, C., de Saint-Cyr, F. D., & Lagasquie-Schiex, M. (2013). Enforcement in argumentation is a kind of update. In *Proceedings of the 7th International Conference on Scalable Uncertainty Management (SUM'13)*, Vol. 8078 of *LNCS*, pp. 30–42. Springer.

Bonzon, E., Delobelle, J., Konieczny, S., & Maudet, N. (2016a). Argumentation ranking semantics based on propagation. In *Proceedings of the Sixth International Conference on Computational Models of Argumentation (COMMA'16)*.

Bonzon, E., Delobelle, J., Konieczny, S., & Maudet, N. (2016b). A comparative study of ranking-based semantics for abstract argumentation. In *Proceedings of the 30th AAAI Conference on Artificial Intelligence (AAAI'16)*.

Booth, R., Caminada, M., Podlaszewski, M., & Rahwan, I. (2012). Quantifying disagreement in argument-based reasoning. In Conitzer, Winikoff, Padgham, & van der Hoek (Eds.), *Proceedings of the 11th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2012)*.

Boyd, S. P., & Vandenberghe, L. (2004). *Convex Optimization*. Cambridge University Press.

Caminada, M. (2006). Semi-stable semantics. In *Proceedings of the First International Conference on Computational Models of Argumentation (COMMA'06)*, pp. 121–130.

Caminada, M., & Gabbay, D. (2009). A logical account of formal argumentation. *Studia Logica*, *93*, 109–145.

Cayrol, C., de Saint-Cyr, F. D., & Lagasquie-Schiex, M.-C. (2010). Change in abstract argumentation frameworks: adding an argument. *Journal of Artificial Intelligence Research*, *38*, 49–84.

Cayrol, C., & Lagasquie-Schiex, M.-C. (2005). Graduality in argumentation. *Journal of Artificial Intelligence Research*, *23*, 245–297.

Cintula, P., Hájek, P., & Noguera, C. (Eds.). (2011). *Handbook of Mathematical Fuzzy Logic*, Vol. 1. College Publications.

Coste-Marquis, S., Devred, C., Konieczny, S., Lagasquie-Schiex, M., & Marquis, P. (2007). On the merging of Dung's argumentation systems. *Artificial Intelligence*, *171*(10-15), 730–753.

Coste-Marquis, S., Konieczny, S., & Maily, J.-G. (2014a). On the revision of argumentation systems: minimal change of argument statuses. In *Proceedings of the 14th International Conference on Principles of Knowledge Representation and Reasoning (KR'14)*, pp. 72–81.

Coste-Marquis, S., Konieczny, S., & Maily, J.-G. (2014b). A translation-based approach from revision of argumentation frameworks. In *Proceedings of the 14th European Conference on Logics in Artificial Intelligence (JELIA'14)*, pp. 77–85.

Coste-Marquis, S., Konieczny, S., Maily, J.-G., & Marquis, P. (2014c). Extension enforcement in abstract argumentation as an optimization problem. In *Proceedings of the 24th International Joint Conference on Artificial Intelligence (IJCAI'15)*, pp. 2876–2882.

da Costa Pereira, C., Tettamanzi, A., & Villata, S. (2011). Changing one's mind: erase or rewind? Possibilistic belief revision with fuzzy argumentation based on trust. In *Proceedings of the Twenty-Second International Joint Conference on Artificial Intelligence (IJCAI'11)*.

De Bona, G., & Finger, M. (2015). Measuring inconsistency in probabilistic logic: rationality postulates and Dutch book interpretation. *Artificial Intelligence*, *227*, 140–164.

Delobelle, J., Konieczny, S., & Vesic, S. (2015). On the aggregation of argumentation frameworks. In *Proceedings of the 24th International Joint Conference on Artificial Intelligence (IJCAI'15)*, pp. 2911–2917.

Diller, M., Haret, A., Linsbichler, T., Rümmele, S., & Woltran, S. (2015). An extension-based approach to belief revision in abstract argumentation. In *Proceedings of the 24th International Joint Conference on Artificial Intelligence (IJCAI'15)*, pp. 2926–2932.

Doder, D., & Woltran, S. (2014). Probabilistic argumentation frameworks – a logical approach. In Straccia, U., & Calì, A. (Eds.), *Scalable Uncertainty Management*, Vol. 8720 of *Lecture Notes in Computer Science*, pp. 134–147. Springer International Publishing.

Dondio, P. (2014). Multi-valued and probabilistic argumentation frameworks. In *Proceedings of the Fifth International Conference on Computational Models of Argumentation (COMMA'14)*.

Dung, P. (1995). On the acceptability of arguments and its fundamental role in nonmonotonic reasoning, logic programming and n-person games. *Artificial Intelligence*, *77*(2), 321–358.

Dung, P., & Thang, P. (2010). Towards (probabilistic) argumentation for jury-based dispute resolution. In *Proceedings of the Third International Conference on Computational Models of Argumentation (COMMA'10)*, pp. 171–182, Amsterdam, The Netherlands, The Netherlands. IOS Press.

Fazzinga, B., Flesca, S., & Parisi, F. (2013). On the complexity of probabilistic abstract argumentation. In *Proceedings of the 23rd International Joint Conference on Artificial Intelligence (IJCAI'13)*.

Fazzinga, B., Flesca, S., & Parisi, F. (2015). On the complexity of probabilistic abstract argumentation frameworks. *ACM Transactions on Computational Logic*, *16*(3), 22:1–22:39.

Fazzinga, B., Flesca, S., & Furfaro, F. (2016). Computing extensions' probabilities in probabilistic abstract argumentation: beyond independence. In *Proceedings of the 22nd European Conference on Artificial Intelligence (ECAI'16)*.

Gabbay, D., & Rodrigues, O. (2012). A numerical approach to the merging of argumentation networks. In *Proceedings of the 13th International Workshop on Computational Logic in Multi-Agent Systems (CLIMA'12)*, Vol. 7486 of *LNCA*, pp. 195–212. Springer.

Gabbay, D., & Rodrigues, O. (2015). Probabilistic argumentation: an equational approach. *Logica Universalis*, *9*(3), 345–382.

Garcia, A., & Simari, G. R. (2004). Defeasible logic programming: an argumentative approach. *Theory and Practice of Logic Programming*, *4*(1–2), 95–138.

Gelfond, M., & Leone, N. (2002). Logic programming and knowledge representation - the a-prolog perspective. *Artificial Intelligence*, *138*(1-2), 3–38.

Grabmair, M., Gordon, T., & Walton, D. (2010). Probabilistic semantics for the carneades argument model using Bayesian networks. In *Proceedings of the Third International Conference on Computational Models of Argumentation (COMMA'10)*, pp. 255–266.

Grant, J., & Hunter, A. (2013). Distance-based measures of inconsistency. In *Proceedings of the 12th Europen Conference on Symbolic and Quantitative Approaches to Reasoning with Uncertainty (ECSQARU'13)*, Vol. 7958 of *Lecture Notes in Computer Science*, pp. 230–241. Springer.

Grossi, D., & Modgil, S. (2015). On the graded acceptability of arguments. In *Proceedings of the 24th International Joint Conference on Artificial Intelligence (IJCAI'15)*, pp. 868–874.

Hadjinikolis, C., Siantos, Y., Modgil, S., Black, E., & McBurney, P. (2013). Opponent modelling in persuasion dialogues.. In *Proceedings of the 22th International Joint Conference on Artificial Intelligence (IJCAI'13)*, pp. 164–170.

Hadoux, E., Beynier, A., Maudet, N., Weng, P., & Hunter, A. (2015). Optimization of probabilistic argumentation with Markov decision models. In *Proceedings of the 24th International Joint Conference on Artificial Intelligence (IJCAI'15)*.

Haenni, R. (1998). Modelling uncertainty with propositional assumptions-based systems. In *Applications of Uncertainty Formalisms*, LNCS, pp. 446–470. Springer.

Haenni, R., Kohlas, J., & Lehmann, N. (2000). Probabilistic argumentation systems. In Gabbay, D., & Smets, P. (Eds.), *Handbook of Defeasible Reasoning and Uncertainty Management Systems, Volume 5*, pp. 221–288. Kluwer.

Hansen, P., & Jaumard, B. (2000). Probabilistic satisfiability. In Kohlas, J., & Moral, S. (Eds.), *Handbook of Defeasible Reasoning and Uncertainty Management Systems: Algorithms for Uncertainty and Defeasible Reasoning*, pp. 321–367. Springer Netherlands, Dordrecht.

Hunter, A. (2012). Some foundations for probabilistic abstract argumentation. In *Proceedings of the Fourth International Conference on Computational Models of Argumentation (COMMA'12)*, pp. 117–128. IOS Press.

Hunter, A. (2013). A probabilistic approach to modelling uncertain logical arguments. *International Journal of Approximate Reasoning*, *54*(1), 47–81.

Hunter, A. (2014a). Probabilistic qualification of attack in abstract argumentation. *International Journal of Approximate Reasoning*, *55*, 607–638.

Hunter, A. (2014b). Probabilistic strategies in dialogical argumentation. In *Proceedings of the 8th International Conference on Scalable Uncertainty Management (SUM'14)*.

Hunter, A. (2015). Modelling the persuadee in asymmetric argumentation dialogues for persuasion. In *Proceedings of the 24th International Joint Conference on Artificial Intelligence (IJCAI'15)*.

Hunter, A. (2016a). Persuasion dialogues via restricted interfaces using probabilistic argumentation. In *Proceedings of the 10th International Conference on Scalable Uncertainty Management (SUM'16)*.

Hunter, A. (2016b). Two dimensional uncertainty in persuadee modelling in argumentation. In *Proceedings of the 22nd European Conference on Artificial Intelligence (ECAI'16)*.

Hunter, A., & Konieczny, S. (2010). On the measure of conflicts: Shapley inconsistency values. *Artificial Intelligence*, *174*(14), 1007–1026.

Hunter, A., & Thimm, M. (2014a). Probabilistic argument graphs for argumentation lotteries. In *Proceedings of the Fifth International Conference on Computational Models of Argumentation (COMMA'14)*.

Hunter, A., & Thimm, M. (2014b). Probabilistic argumentation with epistemic extensions. In *Proceedings of the International Workshop on Defeasible and Ampliative Reasoning (DARe'14)*.

Hunter, A., & Thimm, M. (2014c). Probabilistic argumentation with epistemic extensions and incomplete information. Tech. rep., ArXiv.

Hunter, A., & Thimm, M. (2014d). Probabilistic argumentation with incomplete information. In *Proceedings of the 21st European Conference on Artificial Intelligence (ECAI'14)*, pp. 1033–1034.

Hunter, A., & Thimm, M. (2016a). On partial information and contradictions in probabilistic abstract argumentation. In *Proceedings of the 15th International Conference on Principles of Knowledge Representation and Reasoning (KR'16)*.

Hunter, A., & Thimm, M. (2016b). Optimization of dialectical outcomes in dialogical argumentation. *International Journal of Approximate Reasoning*, *78*, 73–102.

Janssen, J., Cock, M. D., & Vermeir, D. (2008). Fuzzy argumentation frameworks. In *Proceedings of the 12th International Conference on Information Processing and Management of Uncertainty in Knowledge-Based Systems (IPMU'08)*, pp. 513–520. Springer.

Jaumard, B., Hansen, P., & Poggi, M. (1991). Column generation methods for probabilistic logic. *ORSA - Journal on Computing*, *3*(2), 135–148.

Katsuno, H., & Mendelzon, A. O. (1991). On the difference between updating a knowledge base and revising it. In *Proceedings of the Second International Conference on the Principles of Knowledge Representation and Reasoning (KR'91)*, pp. 387–394.

Kern-Isberner, G., & Rödder, W. (2004). Belief revision and information fusion on optimum entropy. *International Journal of Intelligent Systems*, *19*(9), 837–857.

Li, H., Oren, N., & Norman, T. J. (2011). Probabilistic argumentation frameworks. In *Proceedings of the First International Workshop on the Theory and Applications of Formal Argumentation (TAFA'11)*.

Li, H., Oren, N., & Norman, T. J. (2013). Relaxing independence assumptions in probabilistic argumentation. In *Proceedings of the Tenth International Workshop on Argumentation in Multi-Agent Systems (ArgMAS'13)*.

Liao, B., & Huang, H. (2015). Formulating semantics of probabilistic argumentation by characterizing subgraphs. In *Logic, Rationality, and Interaction: 5th International Workshop (LORI'15)*, Vol. 9394 of *Lecture Notes in Computer Science*, pp. 243–254. Springer.

Lukasiewicz, T. (2000). Probabilistic logic programming with conditional constraints. *ACM Transactions on Computational Logic*, *2*(3), 289–339.

Matt, P.-A., & Toni, F. (2008). A game-theoretic perspective on the notion of argument strength in abstract argumentation. In Hölldobler, S., Lutz, C., & Wansing, H. (Eds.), *Proceedings of the 11th European Conference on Logics in Artificial Intelligence (JELIA'08)*, pp. 285–297. Springer Berlin Heidelberg.

Modgil, S., & Prakken, H. (2014). The ASPIC+ framework for structured argumentation: a tutorial. *Argument and Computation*, *5*, 31–62.

Nielsen, S., & Parsons, S. (2007). An application of formal argumentation: fusing Bayesian networks in multi-agent systems. *Artificial Intelligence*, *171*(10-15), 754–775.

Nute, D. (1994). Defeasible logic. In Gabbay, D., Hogger, C. J., & Robinson, J. A. (Eds.), *Handbook of logic in artificial intelligence and logic programming (vol. 3)*, pp. 353–395. Oxford University Press, Inc., New York, NY, USA.

Paris, J. B. (1994). *The Uncertain Reasoner's Companion – A Mathematical Perspective*. Cambridge University Press.

Pearl, J. (1988). *Probabilistic reasoning in intelligent systems: networks of plausible inference*. Springer.

Polberg, S., & Doder, D. (2014). Probabilistic abstract dialectical frameworks. In *Proceedings of the 14th European Conference on Logics in Artificial Intelligence (JELIA'14)*.

Pollock, J. (1995). *Cognitive Carpentry. A Blueprint for How to Build a Person*. MIT Press.

Pollock, J. L. (2001). Defeasible reasoning with variable degrees of justification. *Artificial Intelligence*, *133*, 233–282.

Prakken, H. (2005). A study of accrual of arguments, with applications to evidential reasoning. In *Proceedings of the Tenth International Conference on Artificial Intelligence and Law (ICAIL'05)*, pp. 85–94.

Rahwan, I., & Simari, G. R. (Eds.). (2009). *Argumentation in Artificial Intelligence*. Springer-Verlag.

Reiter, R. (1980). A logic for default reasoning. *Artificial Intelligence*, *13*, 81–132.

Rienstra, T., Thimm, M., & Oren, N. (2013). Opponent models with uncertainty for strategic argumentation. In *Proceedings of the 23rd International Joint Conference on Artificial Intelligence (IJCAI'13)*.

Riveret, R., & Governatori, G. (2016). On learning attacks in probabilistic abstract argumentation. In *Proceedings of the 15th International Conference on Autonomous Agents and Multiagent Systems (AAMAS'16)*.

Riveret, R., Rotolo, A., Sartor, G., Prakken, H., & Roth, B. (2007). Success chances in argument games: a probabilistic approach to legal disputes. In *Proceedings of the 20th Annual Conference on Legal Knowledge and Information Systems (JURIX'07)*, pp. 99–108. IOS Press.

Shafer, G. (1976). *A Mathematical Theory of Evidence*. Princeton University Press.

Sun, X., & Liao, B. (2016). Probabilistic argumentation, a small step for uncertainty, a giant step for complexity. In *Multi-Agent Systems and Agreement Technologies, 13th European Conference, EUMAS 2015, and Third International Conference, AT 2015, Athens, Greece, December 17-18, 2015, Revised Selected Papers*. Springer.

Thimm, M. (2012). A probabilistic semantics for abstract argumentation. In *Proceedings of the 20th European Conference on Artificial Intelligence (ECAI'12)*.

Thimm, M. (2013). Inconsistency measures for probabilistic logics. *Artificial Intelligence*, *197*, 1–24.

Thimm, M. (2014). Strategic argumentation in multi-agent systems. *Künstliche Intelligenz, Special Issue on Multi-Agent Decision Making*, *28*(3), 159–168.

Thimm, M., & Gabbay, D. M. (Eds.). (2016). *Special Issue: Probabilistic and Quantitative Approaches to Computational Argumentation*. IfCoLog Journal of Logics and their Applications, Volume 3 Number 2.

Thimm, M., & Kern-Isberner, G. (2014). On controversiality of arguments and stratified labelings. In *Proceedings of the Fifth International Conference on Computational Models of Argumentation (COMMA'14)*.

Timmer, S. T., Meyer, J.-J. C., Prakken, H., Renooij, S., & Verheij, B. (2015). Explaining Bayesian networks using argumentation. In *Proceedings of the 13th European Conference on Symbolic and Quantitative Approaches to Reasoning with Uncertainty (ECSQARU'15)*.

Toni, F. (2014). A tutorial on assumption-based argumentation. *Argument & Computation*, *5*(1), 89–117.

Verheij, B. (2014). Arguments and their strength: revisiting Pollock's anti-probabilistic starting points. In *Proceedings of the Fifth International Conference on Computational Models of Argumentation (COMMA'14)*.

Vreeswijk, G. (2004). Argumentation in Bayesian belief networks. In *Proceedings of the First International Workshop on Argumentation in Multi-Agent Systems (ArgMAS'04)*, pp. 111–129.

Wu, Y., & Caminada, M. (2010). A labelling-based justification status of arguments. *Studies in Logic*, *3*(4), 12–29.