

# Probabilistic Argumentation with Epistemic Extensions

Anthony Hunter<sup>1</sup> and Matthias Thimm<sup>2</sup>

<sup>1</sup> Department of Computer Science, University College London, United Kingdom

<sup>2</sup> Institute for Web Science and Technologies (WeST), 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 a probability assignment to each argument. There are various interpretations that can be ascribed to this assignment. In this paper, we regard the assignment as denoting the belief that an agent has that an argument is justifiable, i. e., that both the premises of the argument and the derivation of the claim of the argument from its premises are valid. This leads to the notion of an epistemic extension which is the subset of the arguments in the graph that are believed to some degree (which we defined as the arguments that have a probability assignment greater than 0.5). We consider various constraints on the probability assignment. Some constraints correspond to standard notions of extensions, such as grounded or stable extensions, and some constraints give us new kinds of extensions.

## 1 Introduction

Abstract argumentation, as proposed by Dung [1], provides a simple and appealing representation in the form of a directed graph. Each node denotes an argument and each arc denotes one argument attacking another. Abstract argumentation also provides a set of options for determining which arguments can be accepted together (i. e. an extension), and which arguments can be rejected. Recently there has been interest in augmenting abstract argumentation with a probabilistic assignment to each argument [2–5]. Once we introduce a probability assignment to each argument, we have extra information about the argumentation, and this means we can refine the evaluation of an argument graph. In most of these proposals (i. e. [2–4]), the emphasis is on what should the structure of the graph be. So the probability of the argument denotes the degree to which the argument should be in the graph.

In this paper, we take a different approach. We regard the assignment as denoting the belief that an agent has that an argument is justifiable, i. e., that both the premises of the argument and the derivation of the claim of the argument from its premises are valid. 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 nor 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 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 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 now model how audiences judge argumentation. Consider for example how a member of the audience of a debate 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. More generally, if we want to make computational models of argument where we can capture persuasion, we need to take account of the belief that participants or audiences have in the individual arguments that are posited. We see the non-standard view as being particularly important in addressing this need.

This paper builds on previous work [5, 6] but extends it in several directions. In particular, the contributions of this paper are as follows:

1. We investigate the notion of epistemic extensions and show their usefulness with respect to classical extensions (Section 3).
2. We adopt and significantly extend properties for *standard epistemic extensions* from [5, 6] 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 apply the notion of epistemic extensions to the problem of completing partial probability assignments and show the feasibility of this approach (Section 6).

Proofs of technical results can be found in an extended technical report [7].

## 2 Preliminaries

An *abstract argumentation framework* AF is a tuple  $AF = (\text{Arg}, \rightarrow)$  where Arg is a set of arguments and  $\rightarrow$  is a relation  $\rightarrow \subseteq \text{Arg} \times \text{Arg}$ . For two arguments  $\mathcal{A}, \mathcal{B} \in \text{Arg}$  the relation  $\mathcal{A} \rightarrow \mathcal{B}$  means that argument  $\mathcal{A}$  attacks argument  $\mathcal{B}$ . For  $\mathcal{A} \in \text{Arg}$  define  $\text{Att}_{\text{AF}}(\mathcal{A}) = \{\mathcal{B} \mid \mathcal{B} \rightarrow \mathcal{A}\}$ .

Semantics are given to abstract argumentation frameworks by means of extensions [1] or labellings [8]. In this work, we use the latter. A labelling  $L$  is a function  $L : \text{Arg} \rightarrow \{\text{in}, \text{out}, \text{undec}\}$  that assigns to each argument  $\mathcal{A} \in \text{Arg}$  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  $\text{in}(L) = \{\mathcal{A} \mid L(\mathcal{A}) = \text{in}\}$  and  $\text{out}(L)$  resp.  $\text{undec}(L)$  be defined analogously. The set  $\text{in}(L)$  for a labelling  $L$  is also called *extension*. A labelling  $L$  is called *conflict-free* if for no  $\mathcal{A}, \mathcal{B} \in \text{in}(L)$  we have that  $\mathcal{A} \rightarrow \mathcal{B}$ .

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

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

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

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

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 can be phrased by imposing further constraints. Let  $\text{AF} = (\text{Arg}, \rightarrow)$  be an abstract argumentation framework and  $L : \text{Arg} \rightarrow \{\text{in}, \text{out}, \text{undec}\}$  a complete labelling. Then

- $L$  is *grounded* if and only if  $\text{in}(L)$  is minimal,
- $L$  is *preferred* if and only if  $\text{in}(L)$  is maximal,
- $L$  is *stable* if and only if  $\text{undec}(L) = \emptyset$ , and
- $L$  is *semi-stable* if and only if  $\text{undec}(L)$  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 [1].

### 3 Epistemic extensions

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  $\mathcal{X}$ . A *probability function*  $P$  on some finite set  $\mathcal{X}$  is a function  $P : 2^{\mathcal{X}} \rightarrow [0, 1]$  with  $\sum_{X \subseteq \mathcal{X}} P(X) = 1$ . Let  $P : 2^{\text{Arg}} \rightarrow [0, 1]$  be a probability function on  $\text{Arg}$ . We abbreviate  $P(\mathcal{A}) = \sum_{\mathcal{A} \in E \subseteq \text{Arg}} P(E)$ . This means that the probability of an argument is the sum of the probabilities of all sets of arguments that contain that argument. The following definition is a generalization of the notion of epistemic extensions given in [6]. For an argument  $\mathcal{A}$ , it is labelled *in* 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  $P(\mathcal{A}) = 0.5$ ). More specifically, let  $\text{AF} = (\text{Arg}, \rightarrow)$  be an abstract argumentation framework and  $P : 2^{\text{Arg}} \rightarrow [0, 1]$  a probability function on  $\text{Arg}$ . The labelling  $L_P : \text{Arg} \rightarrow \{\text{in}, \text{out}, \text{undec}\}$  defined via the following constraints is called the *epistemic labelling* of  $P$ :

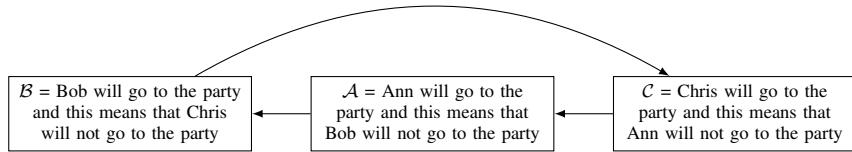
- $L_P(\mathcal{A}) = \text{in}$  iff  $P(\mathcal{A}) > 0.5$
- $L_P(\mathcal{A}) = \text{out}$  iff  $P(\mathcal{A}) < 0.5$

$$- L_P(\mathcal{A}) = \text{undec} \quad \text{iff} \quad P(\mathcal{A}) = 0.5$$

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  $X = \text{in}(L_P)$ . Furthermore, we say that a labelling  $L$  and a probability function  $P$  are *congruent*, denoted by  $L \sim P$ , if for all  $\mathcal{A} \in \text{Arg}$  we have  $L(\mathcal{A}) = \text{in} \Leftrightarrow P(\mathcal{A}) = 1$ ,  $L(\mathcal{A}) = \text{out} \Leftrightarrow P(\mathcal{A}) = 0$ , and  $L(\mathcal{A}) = \text{undec} \Leftrightarrow P(\mathcal{A}) = 0.5$ . 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 (see Section 5).

To further illustrate the epistemic extensions, consider the graph given in Figure 1. 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).



**Fig. 1.** Example of three arguments in a simple cycle.

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  $\text{AF} = (\text{Arg}, \rightarrow)$  be an abstract argumentation framework and  $P : 2^{\text{Arg}} \rightarrow [0, 1]$ . Consider the following properties (note that COH is from [6] and JUS is from [5]):

**COH**  $P$  is *coherent* wrt. AF if for every  $\mathcal{A}, \mathcal{B} \in \text{Arg}$ , 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  $\mathcal{A} \in \text{Arg}$  with  $\text{Att}_{\text{AF}}(\mathcal{A}) = \emptyset$ .

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

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

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

**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  $\mathcal{A} \in \text{Arg}$ .

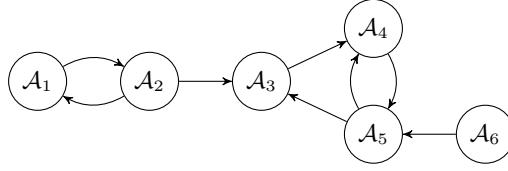
COH ensures that if argument  $\mathcal{A}$  attacks argument  $\mathcal{B}$ , then the degree to which  $\mathcal{A}$  is believed caps the degree to which  $\mathcal{B}$  can be believed; SFOU ensures that if an argument is not attacked, then the argument is not disbelieved (i.e.  $P(\mathcal{A}) \geq 0.5$ ); 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; 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 1.* Consider  $\text{AF} = (\text{Arg}, \rightarrow)$  depicted in Fig. 2 and the probability functions depicted in Table 1 (note that the probability functions there are only partially defined by giving the probabilities of arguments). The following observations can be made: 1.)  $P_1$  is semi-founded, founded, but neither coherent, optimistic, semi-optimistic, ternary, nor justifiable, 2.)  $P_2$  is coherent and semi-optimistic, but neither semi-founded, founded, optimistic, ternary, nor justifiable, 3.)  $P_3$  is coherent, semi-optimistic, semi-founded, founded, optimistic, and justifiable, but not ternary, 4.)  $P_4$  is semi-founded, founded, optimistic, and semi-optimistic, but neither coherent, justifiable, nor ternary, and 5.)  $P_5$  is coherent, semi-founded, semi-optimistic, and ternary but neither optimistic, justifiable, nor founded.

	$\mathcal{A}_1$	$\mathcal{A}_2$	$\mathcal{A}_3$	$\mathcal{A}_4$	$\mathcal{A}_5$	$\mathcal{A}_6$
$P_1$	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 1

Let  $\mathcal{P}$  be the set of all probability functions,  $\mathcal{P}(\text{AF})$  be the set of all probability functions on  $\text{Arg}$ , and  $\mathcal{P}_t(\text{AF})$  be the set of all  $t$ -probability functions with  $t \in \{\text{COH}, \text{SFOU}, \text{FOU}, \text{OPT}, \text{SOPT}, \text{JUS}, \text{TER}\}$ . We obtain the following relationships between the different classes of probability functions.



**Fig. 2.** A simple argumentation framework

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

1.  $\emptyset \subsetneq \mathcal{P}_{\text{JUS}}(\text{AF}) \subsetneq \mathcal{P}_{\text{COH}}(\text{AF}) \subsetneq \mathcal{P}(\text{AF})$
2.  $\mathcal{P}_{\text{OPT}}(\text{AF}) = \mathcal{P}_{\text{SOPT}}(\text{AF}) \cap \mathcal{P}_{\text{FOU}}(\text{AF})$ .
3.  $\mathcal{P}_{\text{FOU}}(\text{AF}) \subsetneq \mathcal{P}_{\text{SFOU}}(\text{AF})$ .
4.  $\emptyset \subsetneq \mathcal{P}_{\text{TER}}(\text{AF}) \subsetneq \mathcal{P}(\text{AF})$ .

For the proof of item 1.) of the above proposition see [5] and [6]. 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  $P \in \mathcal{P}_{\text{SFOU}}(\text{AF})$ .

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

1.  $P \in \mathcal{P}_{\text{TER}}(\text{AF})$ ;
2. if  $P(\mathcal{A}) = 1$  then  $P(\mathcal{B}) = 0$  for all  $\mathcal{B} \in \text{Arg}$  with  $\mathcal{B} \rightarrow \mathcal{A}$ ;
3. if  $P(\mathcal{B}) = 0$  for all  $\mathcal{B}$  with  $\mathcal{B} \rightarrow \mathcal{A}$  then  $P(\mathcal{A}) = 1$ ;
4. if  $P(\mathcal{A}) = 0$  then there is  $\mathcal{B} \in \text{Arg}$  with  $P(\mathcal{B}) = 1$  and  $\mathcal{B} \rightarrow \mathcal{A}$ ;
5. if  $P(\mathcal{B}) = 1$  for some  $\mathcal{B}$  with  $\mathcal{B} \rightarrow \mathcal{A}$  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  $P \sim L$ .

Completeness of probability functions can be characterized by the aforementioned properties as follows.

**Proposition 3.** For an argument graph  $\text{AF}$ , and for  $P \in \mathcal{P}(\text{AF})$ ,  $P$  is a complete probability function iff  $P \in \mathcal{P}_{\text{COH}}(\text{AF}) \cap \mathcal{P}_{\text{FOU}}(\text{AF}) \cap \mathcal{P}_{\text{TER}}(\text{AF})$ .

In the same way that Caminada and Gabbay [9] 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.

**Proposition 4.** *Let  $\text{AF} = (\text{Arg}, \rightarrow)$  be an abstract argumentation framework and  $P \in \mathcal{P}(\text{AF})$ . If  $P$  is a complete probability function for  $\text{AF}$  and the restriction specified in Table 2 holds for  $P$ , then the corresponding type of epistemic labelling is obtained.*

Restriction on 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

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

**Proposition 5.** *(From [5]) Let  $\text{AF} = (\text{Arg}, \rightarrow)$  be an argumentation framework.*

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

So Proposition 4 and Proposition 5 provide two ways to identify probability functions that capture specific types of labellings. Each of these results show that standard notions of classical semantics (e. g. admissibility) 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.

**Proposition 6.** *For each complete labelling  $L$ , if there is  $\mathcal{A}$  such that  $L(\mathcal{A}) \neq \text{undec}$ , then there are infinitely many probability functions  $P$  such that  $L_P = L$ .*

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

---

<sup>3</sup> Define the *entropy*  $H(P)$  of  $P$  as  $H(P) = - \sum_{E \subseteq \text{Arg}} P(E) \log P(E)$

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  $\text{AF} = (\text{Arg}, \rightarrow)$  be an abstract argumentation framework and  $P : 2^{\text{Arg}} \rightarrow [0, 1]$ .

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

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

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

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

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

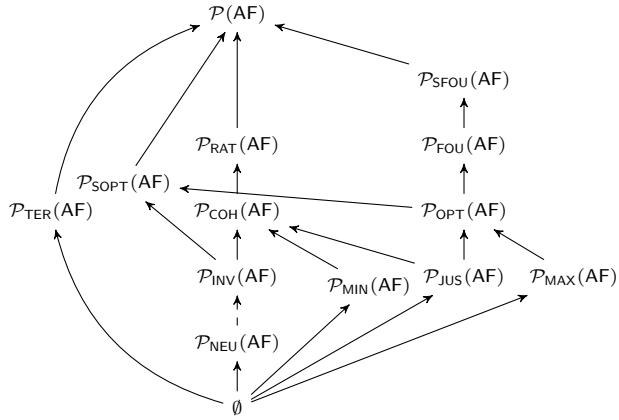
We explain these constraints as follows: **RAT** ensures that if argument  $\mathcal{A}$  attacks argument  $\mathcal{B}$ , and  $\mathcal{A}$  is believed (i. e.  $P(\mathcal{A}) > 0.5$ ), then  $\mathcal{B}$  is not believed (i. e.  $P(\mathcal{B}) \leq 0.5$ ); **NEU** ensures that all arguments are 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 2.* We continue Example 1, the framework  $\text{AF} = (\text{Arg}, \rightarrow)$  depicted in Fig. 2, 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, involuntary, maximal, nor minimal, 2.)  $P_1$  and  $P_4$  are neither rational, neutral, involuntary, maximal, nor minimal, and 3.)  $P_5$  is rational, neutral, and involuntary but neither maximal nor minimal.

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

**Proposition 7.** *Let  $\text{AF} = (\text{Arg}, \rightarrow)$  be an abstract argumentation framework.*

1.  $\emptyset \subsetneq \mathcal{P}_{\text{JUS}}(\text{AF}) \subsetneq \mathcal{P}_{\text{COH}}(\text{AF}) \subsetneq \mathcal{P}_{\text{RAT}}(\text{AF}) \subsetneq \mathcal{P}(\text{AF})$
2.  $\emptyset \subsetneq \mathcal{P}_{\text{NEU}}(\text{AF}) \subseteq \mathcal{P}_{\text{INV}}(\text{AF}) \subsetneq \mathcal{P}_{\text{COH}}(\text{AF})$
3.  $\emptyset \subsetneq \mathcal{P}_{\text{INV}}(\text{AF}) \subsetneq \mathcal{P}_{\text{SOPT}}(\text{AF})$
4.  $\emptyset \subsetneq \mathcal{P}_{\text{MIN}}(\text{AF}) \subsetneq \mathcal{P}_{\text{COH}}(\text{AF})$



**Fig. 3.** Classes of probability functions (a normal arrow  $\rightarrow$  indicates a strict subset relation, a dashed arrow  $\dashrightarrow$  indicates a subset relation)

$$5. \emptyset \subsetneq P_{\text{MAX}}(\text{AF}) \subsetneq P_{\text{OPT}}(\text{AF})$$

Together with Examples 1 and 2 we obtain the strict classification of classes of probability functions as depicted in Figure 3.

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.

**Proposition 8.** *If  $L$  is an admissible labelling, then there is a  $P \in \mathcal{P}_{\text{RAT}}(\text{AF})$  such that  $L \sim P$ .*

Furthermore, the epistemic labelling corresponding to each probability function that satisfies the RAT property is conflict-free.

**Proposition 9.** *Let  $\text{AF} = (\text{Arg}, \rightarrow)$  be an abstract argumentation framework. For each  $P \in \mathcal{P}_{\text{RAT}}(\text{AF})$ ,  $\text{in}(L_P)$  is a conflict-free set of arguments in  $\text{AF}$ .*

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

**Proposition 10.** *Let  $\text{AF} = (\text{Arg}, \rightarrow)$  be an abstract argumentation framework. If  $\text{AF}$  contains an odd cycle (i. e. there is a sequence of attacks  $A_1 \rightarrow A_2 \rightarrow \dots \rightarrow A_k$  where  $A_1 = A_k$  and  $k$  is an even number), and  $P \in \mathcal{P}_{\text{INV}}(\text{AF})$  then  $P \in \mathcal{P}_{\text{NEU}}(\text{AF})$ .*

Even when the graph is acyclic, it may be the case that there is no involuntary 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 involuntary 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 involuntary and not neutral. But even here there are constraints as captured in the next result.

**Proposition 11.** *If  $P \in \mathcal{P}_{\text{INV}}(\text{AF})$ , then for all  $\mathcal{B}_i, \mathcal{B}_j \in \text{Att}_{\text{AF}}(\mathcal{A})$  we have  $P(\mathcal{B}_i) = P(\mathcal{B}_j)$ .*

When  $P \in \mathcal{P}_{\text{MAX}}(\text{AF})$ , 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. However, there is compatibility between a probability function being maximal and a probability function being founded since each  $P \in \mathcal{P}_{\text{MAX}}(\text{AF})$  is in  $\mathcal{P}_{\text{FOU}}(\text{AF})$ .

**Proposition 12.** *Let  $\text{AF} = (\text{Arg}, \rightarrow)$  be an abstract argumentation framework. If there are  $\mathcal{A}, \mathcal{B} \in \text{Arg}$  such that  $\mathcal{A} \rightarrow \mathcal{B}$ , then  $\mathcal{P}_{\text{RAT}}(\text{AF}) \cap \mathcal{P}_{\text{MAX}}(\text{AF}) = \emptyset$ .*

In conclusion, we have identified epistemic extensions that are obtained from rational probability functions as being an appealing 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 whilst taking the structure of the argument graph into account. We have also considered alternatives such as the involuntary probability functions but these are over-constrained.

## 6 Application: Partial probability functions

Assigning a probability value to an argument can be useful for a variety of purposes such as representing the belief that the premises of the argument are valid, or the belief that the claim is valid given that the premises are valid, or the belief that both the premises and claim are valid, or the belief that the argument should appear in the argument graph, etc. However, given an argument graph, it may be difficult for a user to assign a value to every argument. The user might have knowledge in order to identify a value for some arguments, but the user may be unable or unwilling to make assignments to the remaining arguments. This means that the user can only provide a partial assignment. If this is the case, then it would be desirable to have techniques to handle this incomplete information. Ideally, we would like to identify an appropriate assignment for all the arguments based on the assignment to the subset of arguments.

More specifically, a partial function  $\pi : \text{Arg} \rightarrow [0, 1]$  on  $\text{Arg}$  is called a *partial probability assignment*. A probability function  $P \in \mathcal{P}(\text{Arg})$  is  $\pi$ -compliant if for every  $\mathcal{A} \in \text{dom } \pi$  we have  $\pi(\mathcal{A}) = P(\mathcal{A})$ . Let  $\mathcal{P}^\pi(\text{AF}) \subseteq \mathcal{P}(\text{AF})$  be the set of all  $\pi$ -compliant probability functions. The question that arises is that given an abstract argumentation framework  $\text{AF} = (\text{Arg}, \rightarrow)$  and a partial probability assignment  $\pi$ , how do we determine  $P \in \mathcal{P}(\text{Arg})$  that is most compatible with both  $\text{AF}$  and  $\pi$ , i. e., which  $P \in \mathcal{P}(\text{Arg})$  do we select as a meaningful representative? This question has also been addressed in similar ways for partial probabilistic information without argumentation, cf. e. g. [10]. There, the principle of *maximum entropy* has been used to complete incomplete probabilistic information in probabilistic logics. As a first step, we investigate the properties of the sets of probability functions which are defined by our different axioms. An important requirement for applying maximum entropy approaches is that the probability function with maximum entropy is uniquely determined. A sufficient property to ensure this, is that the set under consideration is both *convex* and *closed*.<sup>4</sup> More generally,

---

<sup>4</sup> A set  $X$  is *convex* if for  $x_1, x_2 \in X$  it also holds that  $\delta x_1 + (1 - \delta)x_2 \in X$  for every  $\delta \in [0, 1]$ ; a set  $X$  is closed if for every converging sequence  $x_1, x_2, \dots$  with  $x_i \in X$  ( $i \in \mathbb{N}$ ) we have that  $\lim_{i \rightarrow \infty} x_i \in X$

maximizing any strictly convex function over a convex set has a unique solution and also most interesting distance measures fall into this category. So if we are to identify probability functions that complete the missing assignments, knowing that for specific sets of probability functions (i. e. those that satisfy specific axioms) that they are convex and closed, means that we may find an appropriate probability function.

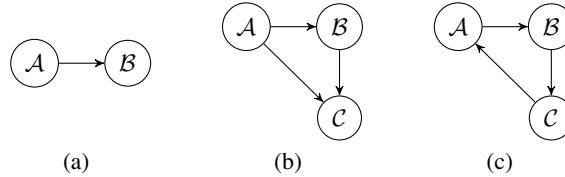
**Proposition 13.** *Let  $\text{AF} = (\text{Arg}, \rightarrow)$  be an abstract argumentation framework.*

1. *The sets  $\mathcal{P}(\text{AF})$ ,  $\mathcal{P}_{\text{COH}}(\text{AF})$ ,  $\mathcal{P}_{\text{TER}}(\text{AF})$ ,  $\mathcal{P}_{\text{NEU}}(\text{AF})$ ,  $\mathcal{P}_{\text{INV}}(\text{AF})$ ,  $\mathcal{P}_{\text{SFOU}}(\text{AF})$ ,  $\mathcal{P}_{\text{FOU}}(\text{AF})$ ,  $\mathcal{P}_{\text{OPT}}(\text{AF})$ ,  $\mathcal{P}_{\text{SOPT}}(\text{AF})$ ,  $\mathcal{P}_{\text{JUS}}(\text{AF})$ ,  $\mathcal{P}_{\text{MIN}}(\text{AF})$ ,  $\mathcal{P}_{\text{MAX}}(\text{AF})$  are convex and closed.*
2. *The set  $\mathcal{P}_{\text{RAT}}(\text{AF})$  is not convex but closed.*

The above proposition suggests that most of our properties are suitable for convex optimization problems such as maximum entropy estimation. Moreover, the same is true for the notion of partial probability assignments.

**Proposition 14.** *For every partial probability assignment  $\pi$  the set  $\mathcal{P}^\pi(\text{AF})$  is convex and closed.*

Let  $t$  be any one of our properties which lead to a convex and closed set of probability functions (or any combination of those). If it is the case that there is at least one  $\pi$ -compliant  $P$  in  $\mathcal{P}_t(\text{AF})$  then (thanks to the convexity properties) we have that the intersection of  $\mathcal{P}^\pi(\text{AF})$  and  $\mathcal{P}_t(\text{AF})$  is convex and closed as well, cf. [10]. In that case, we can select the probability function with maximal entropy within this intersection (which is uniquely defined). As for the rationale of this decision, several results from probability reasoning, as for example discussed in [10], could be harnessed. We continue with



**Fig. 4.** Argumentation frameworks for Example 3

some examples to illustrate the definitions and to investigate some of our concerns in dealing with partial assignments.

*Example 3.* For the argumentation framework depicted in Figure 4(a) consider  $\pi_1$  with  $\pi_1(\mathcal{A}) = 1$ . Obviously, the most reasonable choice for a  $\pi_1$ -compliant  $P \in \mathcal{P}(\text{AF})$  would be  $P(\mathcal{A}) = 1$  and  $P(\mathcal{B}) = 0$  (by obeying the property of involution). Furthermore, for  $\pi_2(\mathcal{B}) = 0.3$  we would have  $P(\mathcal{A}) = 0.7$  and  $P(\mathcal{B}) = 0.3$  following the same rationale.

For the argumentation framework depicted in Figure 4(b) consider  $\pi_3$  with  $\pi_3(\mathcal{C}) = 0.4$ . A possible choice for  $P$  would be  $P(\mathcal{A}) = 0.6$ ,  $P(\mathcal{B}) = 0.4$ , and  $P(\mathcal{C}) = 0.4$  (having thus a maximally committed function that is coherent). But note that the set  $\mathcal{P}^\pi(\text{AF}) \cap \mathcal{P}_{\text{COH}}(\text{AF})$  does contain more than this single probability function. Furthermore, for  $\pi_4$  with  $\pi_4(\mathcal{B}) = 0.7$  and  $\pi_4(\mathcal{C}) = 0.6$  one would only guess  $P(\mathcal{A}) \leq 0.3$  but

due to the “inconsistency” of  $\pi_4$  (violating the coherence condition), what is the best choice?

For the argumentation framework depicted in Figure 4(c) consider  $\pi_5$  with  $\pi_5(\mathcal{A}) = 0.4$  and the following four selections  $P_1, P_2, P_3, P_4 \in \mathcal{P}(\text{AF})$ :

$$\begin{array}{llll} P_1(\mathcal{A}) = 0.4 & P_2(\mathcal{A}) = 0.4 & P_3(\mathcal{A}) = 0.4 & P_4(\mathcal{A}) = 0.4 \\ P_1(\mathcal{B}) = 0.6 & P_2(\mathcal{B}) = 0.4 & P_3(\mathcal{B}) = 0.5 & P_4(\mathcal{B}) = 0.2 \\ P_1(\mathcal{C}) = 0.4 & P_2(\mathcal{C}) = 0.6 & P_3(\mathcal{C}) = 0.5 & P_4(\mathcal{C}) = 0.3 \end{array}$$

All of the above probability functions are  $\pi_5$ -compliant and coherent. Function  $P_4$  is not maximally committed and as such is perhaps not a good choice. Both  $P_1$  and  $P_2$  are “extreme points of view” and model some kind of probabilistic stable semantics. The function  $P_3$  is as unbiased as possible but still “reasonable” as it models probabilistic grounded semantics. Note that  $P_3$  is also the probability function with maximal entropy in  $\mathcal{P}^\pi(\text{AF}) \cap \mathcal{P}_{\text{COH}}(\text{AF})$ .

Given  $\mathcal{P}^\pi(\text{AF})$  and  $\mathcal{P}_t(\text{AF})$ , we can either select  $P \in \mathcal{P}^\pi(\text{AF})$  that is “as close as possible to”  $\mathcal{P}_t(\text{AF})$  or  $P \in \mathcal{P}_t(\text{AF})$  that is “as close as possible to”  $\mathcal{P}^\pi(\text{AF})$ . In future work, we will investigate definitions for “as close as possible to”, and we will explore the pros and cons of each of these alternatives for selecting  $P$ .

## 7 Discussion

In this paper, we have investigated the use of a probability function to represent belief in an argument. We use this to identify an epistemic labelling, and thereby an epistemic extension. We have considered various constraints on the probability function leading to two views on using the probability functions, namely the standard view, and the non-standard view. We applied our classification of properties of probabilistic argumentation to the problem of completing partial probability assignments. A first investigation leads us to believe that maximizing entropy within probability functions of a specific type gives appropriate results for this problem. In future work, we will investigate this issue in more depth.

The work in this paper contrasts with other research on introducing a probability assignment to arguments such as [3]. There, a probability distribution over the subgraphs of the argument graph is introduced, and this can then be used to give a probability assignment for a set of arguments being an admissible set or extension of the argument graph. They assume independence between arguments which in general is not appropriate. To address this shortcoming, the set of spanning subgraphs can be used as a sample space, thereby obviating the need for an independence assumption between arguments [4]. This probability distribution over subgraphs is used to give a probability assignment to extensions. For a semantics  $X$  (such as grounded, preferred, or stable), the probability that a set of arguments  $\Gamma$  is an extension according to semantics  $X$ , denoted  $P_X(\Gamma)$ , is the sum of the probability assigned to the subgraphs for which  $\Gamma$  is an extension according to semantics  $X$ . The uncertainty that is being handled in [3, 4] is about the structure of the graph, and it is therefore a different kind of uncertainty model to the one being addressed by this paper. Similarly, the work [11] extends abstract argumentation by allowing the attack relation to be a *fuzzy relation*.

## References

1. Dung, P.M.: On the Acceptability of Arguments and its Fundamental Role in Nonmonotonic Reasoning, Logic Programming and n-Person Games. *Artificial Intelligence* **77**(2) (1995) 321–358
2. Dung, P., Thang, P.: Towards (probabilistic) argumentation for jury-based dispute resolution. In: Computational Models of Argument (COMMA'10), IOS Press (2010) 171–182
3. Li, H., Oren, N., Norman, T.: Probabilistic argumentation frameworks. In: Proceedings of the First International Workshop on the Theory and Applications of Formal Argumentation (TAFA'11). (2011)
4. Hunter, A.: Some foundations for probabilistic abstract argumentation. In: Computational Models of Argument (COMMA'12), IOS Press (2012) 117–128
5. Thimm, M.: A probabilistic semantics for abstract argumentation. In: Proceedings of the 20th European Conference on Artificial Intelligence (ECAI'12). (2012)
6. Hunter, A.: A probabilistic approach to modelling uncertain logical arguments. *International Journal of Approximate Reasoning* (2013)
7. Hunter, A., Thimm, M.: Probabilistic argumentation with epistemic extensions and incomplete information. Technical report, ArXiv (May 2014) <http://arxiv.org/abs/1405.3376>.
8. Wu, Y., Caminada, M.: A labelling-based justification status of arguments. *Studies in Logic* **3**(4) (2010) 12–29
9. Caminada, M., Gabbay, D.: A logical account of formal argumentation. *Studia Logica* **93** (2009) 109–145
10. Paris, J.B.: The Uncertain Reasoner's Companion – A Mathematical Perspective. Cambridge University Press (1994)
11. Janssen, J., Cock, M.D., Vermeir, D.: Fuzzy argumentation frameworks. In: Proc. of the 12th Int. Conf. on Information Processing and Management of Uncertainty in Knowledge-Based Systems (IPMU'08), Springer (2008) 513–520