On the Complexity of Probabilistic Abstract Argumentation*

Bettina Fazzinga, Sergio Flesca, Francesco Parisi
DIMES - University of Calabria
Rende (CS), Italy
{bfazzinga,flesca,fparisi}@dimes.unical.it

Abstract

Probabilistic abstract argumentation combines Dung's abstract argumentation framework with probability theory in order to model uncertainty in argumentation. In this setting, we address the fundamental problem of computing the probability that a set of arguments is an *extension* according to a given semantics. We focus on the most popular semantics (i.e., *admissible*, *stable*, *complete*, *grounded*, *preferred*, *ideal*), and show the following dichotomy result: computing the probability that a set of arguments is an extension is either *PTIME* or *FP*$^{\#P}$-complete depending on the semantics adopted. Our *PTIME* results are particularly interesting, as they hold for some semantics for which no polynomial-time technique was known so far.

1 Introduction

Argumentation allows disputes to be modeled, which arise between two or more parties, each of them providing arguments to assert her reasons. Although argumentation is strongly related to philosophy and law, it has gained remarkable interest in AI as a reasoning model for representing dialogues, making decisions, and handling inconsistency and uncertainty [Bench-Capon and Dunne, 2007; Besnard and Hunter, 2008; Rahwan and Simari, 2009].

A simple but powerful argumentation framework is that proposed in the seminal paper [Dung, 1995], where the *abstract argumentation framework* (AAF) was introduced. An AAF is a pair $\langle A, D\rangle$ consisting of a set A of *arguments*, and of a binary relation D over A, called *defeat* (or, equivalently, *attack*) relation. Basically, an argument is an abstract entity that may attack and/or be attacked by other arguments. For instance, consider the following scenario (inspired by an example in [Hunter, 2013]), where we are interested in deciding whether to organize or not a BBQ party in our garden on Saturday. Assume that our decision should be taken considering the argument a, which is *"Our friends will have great fun at the party"*, and the argument b, which is *"Saturday will rain"* (according to the BBC weather forecasting service). This scenario can be modeled by the AAF $\mathcal{A}$, whose set of arguments is $\{a, b\}$, and whose defeat relation consists of the defeat $\delta = (b, a)$, meaning that the fun at the party is jeopardized if it rains.

Several semantics for AAFs, such as *admissible*, *stable*, *preferred*, *complete*, *grounded*, and *ideal*, have been proposed [Dung, 1995; Dung *et al.*, 2007; Baroni and Giacomin, 2009] to identify "reasonable" sets of arguments, called *extensions*. Basically, each of these semantics corresponds to some properties which "certify" whether a set of arguments can be profitably used to support a point of view in a discussion. For instance, a set S of arguments is an extension according to the admissible semantics if it has two properties: it is conflict-free (that is, there is no defeat between arguments in S), and every argument (outside S) attacking an argument in S is counterattacked by an argument in S. Intuitively enough, the fact that a set is an extension according to the admissible semantics means that, using the arguments in S, you do not contradict yourself, and you can rebut to anyone who uses any of the arguments outside S to contradict yours. The other semantics correspond to other ways of determining whether a set of arguments would be a "good point" in a dispute, and will be described in the core of the paper. For any semantics *sem*, the fundamental problem of verifying whether a set S of arguments is an extension according to *sem* is denoted as $\textsc{Ver}^{sem}(S)$, and its complexity, for the above-mentioned semantics, has been addressed in [Dunne and Wooldridge, 2009; Dunne, 2009].

As a matter of fact, in the real world, arguments and defeats are often uncertain, thus, several proposals have been made to model uncertainty in AAFs, by considering weights, preferences, or probabilities associated with arguments and/or defeats. In this regard, [Dung and Thang, 2010; Li *et al.*, 2011; Thimm, 2012; Rienstra, 2012] have recently extended the original Dung framework in order to achieve probabilistic abstract argumentation frameworks (PrAFs), where uncertainty of arguments and defeats is modeled by exploiting the probability theory. In particular, [Li *et al.*, 2011] proposed a PrAF where both arguments and defeats are associated with probability values. For instance, a PrAF $\mathcal{F}_{\mathcal{A}}$ can be obtained from the AAF $\mathcal{A}$ by considering the arguments a, b, and the defeat δ as probabilistic events, having probabilities $Pr(a) = .9$, $Pr(b) = .7$, and $Pr(\delta) = .9$. Basically, this means that there

*The first and the second author were partially supported by the project PON01_01286 - eJRM.

is some uncertainty about the fact that our friends will have fun at the party, about the truthfulness of the BBC weather forecasting service, and about the fact that the bad weather forecast actually entails that the party will be disliked by our friends.

The issue of how to assign probabilities to arguments and defeats in abstract argumentation, with particular reference to the PrAF proposed in [Li *et al.*, 2011], has been deeply investigated in [Hunter, 2012; 2013], where the *justification* and the *premise* perspectives have been introduced. In this paper, we do not address this issue, but, assuming that the probabilities of arguments and defeats are given, we tackle the probabilistic counterpart of the problem $\text{VER}^{sem}(S)$, that is, the problem $\text{PROB}^{sem}_{\mathcal{F}}(S)$ of *computing the probability* $Pr^{sem}_{\mathcal{F}}(S)$ *that a set* S *of arguments is an extension according to a given semantics sem*. To this end, we consider the PrAF proposed in [Li *et al.*, 2011], which is based on the notion of *possible world*. Basically, given a PrAF $\mathcal{F}$, a possible world represents a (deterministic) scenario consisting of some subset of the arguments and defeats in $\mathcal{F}$. Hence, a possible world can be viewed as an AAF containing exactly the arguments and the defeats occurring in the represented scenario. For instance, considering the above-introduced PrAF $\mathcal{F}_A$, the possible world $\langle\{a\},\emptyset\rangle$ is the AAF representing the scenario where only a occurs, while the possible world $\langle\{a,b\},\{\delta\}\rangle$ is the AAF representing the scenario where all the arguments and defeats occur.

In [Li *et al.*, 2011] it was shown that a PrAF admits a unique probability distribution over the set of possible worlds, which assigns a probability value to each possible world coherently with the probabilities of arguments and defeats. This follows from the assumption that arguments are viewed as pairwise independent probabilistic events, while each defeat is viewed as a probabilistic event conditioned by the occurrence of the arguments it relates, but independent from any other event. Once shown that a PrAF admits a unique probability distribution over the set of possible worlds, the probability $Pr^{sem}_{\mathcal{F}}(S)$ is naturally defined as the sum of the probabilities of the possible worlds where the set S of arguments is an extension according to the semantics *sem*. Unfortunately, as pointed out in [Li *et al.*, 2011], computing $Pr^{sem}_{\mathcal{F}}(S)$ by directly exploiting its definition would result in an exponential time algorithm. However, this does not mean that $\text{PROB}^{sem}_{\mathcal{F}}(S)$ is intractable, since it does not rule out the possibility that polynomial time algorithms for computing $Pr^{sem}_{\mathcal{F}}(S)$ exist. Hence, the point becomes: *what actually is the computational cost of computing* $Pr^{sem}_{\mathcal{F}}(S)$*?* To the best of our knowledge this question has not been answered so far.

Main contributions. In this paper, we characterize the computational complexity of the problem $\text{PROB}^{sem}_{\mathcal{F}}(S)$ of computing $Pr^{sem}_{\mathcal{F}}(S)$, where *sem* is one of the following semantics: *admissible*, *stable*, *complete*, *grounded*, *preferred*, *ideal*. The complexity of $\text{PROB}^{sem}_{\mathcal{F}}(S)$ is reported in (the rightmost column of) Table 1, which, for the sake of completeness, also reports the results concerning $\text{VER}^{sem}(S)$ [Dunne and Wooldridge, 2009; Dunne, 2009].

sem	$\text{VER}^{sem}(S)$	$\text{PROB}^{sem}_{\mathcal{F}}(S)$
admissible	*PTIME*	*PTIME*
stable	*PTIME*	*PTIME*
complete	*PTIME*	$FP^{\#P}$-complete
grounded	*PTIME*	$FP^{\#P}$-complete
preferred	*coNP*-complete	$FP^{\#P}$-complete
ideal	*coNP*-complete	$FP^{\#P}$-complete

Table 1: Complexity of $\text{VER}^{sem}(S)$ and $\text{PROB}^{sem}_{\mathcal{F}}(S)$.

Our results are interesting from two standpoints. First, comparing the complexity of $\text{PROB}^{sem}_{\mathcal{F}}(S)$ with that of its deterministic counterpart $\text{VER}^{sem}(S)$ shows that (*i*) for some semantics (that is, complete and grounded) $\text{VER}^{sem}(S)$ is tractable while $\text{PROB}^{sem}_{\mathcal{F}}(S)$ is not; (*ii*) for other semantics (that is, admissible and stable) the two problems are both tractable; and, finally, there are semantics (that is, preferred and ideal) for which these problems are both intractable.

Second, our complexity analysis allows us to understand for which semantics computing $Pr^{sem}_{\mathcal{F}}(S)$ is feasible in practice or not. Indeed, [Li *et al.*, 2011] claimed that computing the exact value of probability $Pr^{sem}_{\mathcal{F}}(S)$ requires exponential time, and then proposed a Monte-Carlo simulation approach to approximate $Pr^{sem}_{\mathcal{F}}(S)$. The claim of [Li *et al.*, 2011] is based on the implicit assumption that, in order to compute $Pr^{sem}_{\mathcal{F}}(S)$, it is necessary to look at all the possible worlds of $\mathcal{F}$, which are exponential in number. However, as far as the admissible and stable semantics are concerned, our results show that the exact value of $Pr^{sem}_{\mathcal{F}}(S)$ can be determined in polynomial time, without enumerating the possible worlds. Hence, using approximate techniques for estimating $Pr^{sem}_{\mathcal{F}}(S)$ is not needed for these semantics. In contrast, the fact that $\text{PROB}^{sem}_{\mathcal{F}}(S)$ is $FP^{\#P}$-complete for the other semantics (complete, grounded, preferred, ideal) backs the use of the approach of [Li *et al.*, 2011] in these cases.

2 Preliminaries

In this section, we briefly recall some basic notions about computational complexity, and concisely overview Dung's abstract argumentation framework, and its probabilistic extension introduced in [Li *et al.*, 2011].

2.1 Complexity

The computational complexity of the problem addressed in this paper is related to the complexity classes of counting problems. A counting problem f is a function from strings over a finite alphabet into integers. $\#P$ is the complexity class of the functions f such that f counts the number of accepting paths of a nondeterministic polynomial-time Turing machine [Valiant, 1979]. Although the problem addressed in the paper is closely related to $\#P$, strictly speaking, it cannot belong to it, since the outputs of our problem are not integers. In fact, we deal with the class $FP^{\#P}$, that is, the class of functions computable by a polynomial-time Turing machine with a $\#P$ oracle.

In the following, we will show that, depending on the adopted semantics, the problem $\textsc{Prob}^{sem}_{\mathcal{F}}(S)$ of computing $Pr^{sem}_{\mathcal{F}}(S)$ belongs to either $FP^{\#P}$ or *PTIME*. We also show lower bounds on the complexity of $\textsc{Prob}^{sem}_{\mathcal{F}}(S)$. In this regard, note that a function is $FP^{\#P}$-hard iff it is $\#P$-hard, and thus to prove that a problem is $FP^{\#P}$-hard it suffices to reduce a $\#P$-hard problem to it.

2.2 Abstract Argumentation

An *abstract argumentation framework* [Dung, 1995] (*AAF*) is a pair $\langle A, D\rangle$, where A is a finite set, whose elements are referred to as *arguments*, and $D \subseteq A \times A$ is a binary relation over A, whose elements are referred to as *defeats* (or *attacks*). An argument is an abstract entity whose role is entirely determined by its relationships with other arguments. Given an AAF $\mathcal{A}$, we also refer to the set of its arguments and the set of its defeats as $Arg(\mathcal{A})$ and $Def(\mathcal{A})$, respectively.

Given arguments $a, b \in A$, we say that a *defeats* b iff there is $(a, b) \in D$. Similarly, a set $S \subseteq A$ *defeats* an argument $b \in A$ iff there is $a \in S$ such that a *defeats* b.

A set $S \subseteq A$ of arguments is said to be *conflict-free* if there are no $a, b \in S$ such that a *defeats* b. An argument a is said to be *acceptable* w.r.t. $S \subseteq A$ iff $\forall b \in A$ such that b *defeats* a, there is $c \in S$ such that c *defeats* b.

Several semantics for AAFs have been proposed to identify "reasonable" sets of arguments, called *extensions*. We consider the following well-known semantics [Dung, 1995; Dung *et al.*, 2007]: *admissible* ($\texttt{ad}$), *stable* ($\texttt{st}$), *complete* ($\texttt{co}$), *grounded* ($\texttt{gr}$), *preferred* ($\texttt{pr}$), and *ideal* ($\texttt{id}$).

A set $S \subseteq A$ of arguments is said to be

- an *admissible* extension iff S is conflict-free and all its arguments are acceptable w.r.t. S;
- a *stable* extension iff S is conflict-free and S defeats each argument in $A \setminus S$;
- a *complete* extension iff S is admissible and S contains all the arguments that are acceptable w.r.t. S;
- a *grounded* extension iff S is a minimal (w.r.t. $\subseteq$) complete set of arguments;
- a *preferred* extension iff S is a maximal (w.r.t. $\subseteq$) admissible set of arguments;
- an *ideal* extension iff S is admissible and S is contained in every preferred set of arguments.

Note that, with a little abuse of notation, we denote as ideal extension what in the literature is often denoted as *ideal set*.

Example 1 *Consider the AAF $\langle A, D\rangle$ obtained by extending the AAF $\mathcal{A} = \langle\{a, b\}, \{\delta_1 = (b, a)\}\rangle$ presented in the introduction as follows. The set A of arguments is $\{a, b, c\}$, where c is the new argument "Saturday will be sunny" (according to the Telegraph weather forecasting service). The set D of defeats is $\{\delta_1 = (b, a), \delta_2 = (b, c), \delta_3 = (c, b)\}$, where δ_2 and δ_3 encode the fact that arguments b and c attack each other. As $S = \{a, c\}$ is conflict-free and every argument in S is acceptable w.r.t. S, it is the case that S is admissible. It is easy to see that the sets $\{b\}$, $\{c\}$, and $\emptyset$ are admissible extensions as well. Since S is conflict-free and defeats b (the only argument in $A \backslash S$) it is stable. As S is maximally admissible, it a preferred extension, while $\{c\}$ is not, since a is acceptable w.r.t $\{c\}$. It is easy to check that S is complete, while it neither is grounded (since it is not minimally complete, as the emptyset is complete) nor ideal (since it is not contained in $\{b\}$ which is a preferred extension).*

Given an AAF $\mathcal{A}$, a set $S \subseteq Arg(\mathcal{A})$ of arguments, and a semantics *sem* $\in \{\texttt{ad}, \texttt{st}, \texttt{co}, \texttt{gr}, \texttt{pr}, \texttt{id}\}$, we define the function *ext*($\mathcal{A}$, *sem*, S) which returns *true* if S is an extension according to sem, *false* otherwise.

2.3 Probabilistic Abstract Argumentation

We now review the *probabilistic* abstract argumentation framework proposed in [Li *et al.*, 2011].

Definition 1 (PrAF) *A probabilistic argumentation framework (PrAF) is a tuple $\langle A, P_A, D, P_D\rangle$ where $\langle A, D\rangle$ is an AAF, and P_A and P_D are, respectively, functions assigning a non-zero*[1] *probability value to each argument in A and defeat in D, that is, $P_A : A \to (0, 1] \cap \mathbb{Q}$ and $P_D : D \to (0, 1] \cap \mathbb{Q}$.*

Basically, the value assigned by P_A to an argument a represents the probability that a actually occurs, whereas the value assigned by P_D to a defeat (a, b) represents the conditional probability that a defeats b given that both a and b occur.

The meaning of a PrAF is given in terms of possible worlds, each of them representing a scenario that may occur in the reality. Given a PrAF $\mathcal{F}$, a possible world is modeled by an AAF which is derived from $\mathcal{F}$ by considering only a subset of its arguments and defeats. More formally, given a PrAF $\mathcal{F} = \langle A, P_A, D, P_D\rangle$, a possible world w of $\mathcal{F}$ is an AAF $\langle A', D'\rangle$ such that $A' \subseteq A$ and $D' \subseteq D \cap (A' \times A')$. The set of the possible worlds of $\mathcal{F}$ will be denoted as $pw(\mathcal{F})$.

Example 2 *As a running example, consider the PrAF $\mathcal{F} = \langle A, P_A, D, P_D\rangle$ where A and D are those of Example 1, and assume that $P_D(\delta_2) = P_D(\delta_3) = 1$ (meaning that arguments b and c attack each other in all the possible scenarios), and that $P_A(c) = .2$ (this corresponds to the assumption that the Telegraph weather forecasting service has a low reliability). Furthermore, recall that $P_A(a) = .9$, $P_A(b) = .7$, $P_D(\delta_1) = .9$, as defined in the introduction. The set $pw(\mathcal{F})$ contains the following possible worlds:*

$w_1 = \langle\emptyset, \emptyset\rangle$ $w_2 = \langle\{a\}, \emptyset\rangle$ $w_3 = \langle\{b\}, \emptyset\rangle$ $w_4 = \langle\{c\}, \emptyset\rangle$
$w_5 = \langle\{a, b\}, \emptyset\rangle$ $w_6 = \langle\{a, c\}, \emptyset\rangle$ $w_7 = \langle\{b, c\}, \emptyset\rangle$
$w_8 = \langle A, \emptyset\rangle$ $w_9 = \langle\{a, b\}, \{\delta_1\}\rangle$ $w_{10} = \langle\{b, c\}, \{\delta_3\}\rangle$
$w_{11} = \langle\{b, c\}, \{\delta_2\}\rangle$ $w_{12} = \langle\{b, c\}, \{\delta_2, \delta_3\}\rangle$ $w_{13} = \langle A, \{\delta_1\}\rangle$
$w_{14} = \langle A, \{\delta_1, \delta_3\}\rangle$ $w_{15} = \langle A, \{\delta_1, \delta_2\}\rangle$ $w_{16} = \langle A, D\rangle$
$w_{17} = \langle A, \{\delta_2\}\rangle$ $w_{18} = \langle A, \{\delta_3\}\rangle$ $w_{19} = \langle A, \{\delta_2, \delta_3\}\rangle$

An interpretation for a PrAF $\mathcal{F} = \langle A, P_A, D, P_D\rangle$ is a probability distribution function I over the set $pw(\mathcal{F})$ of the possible worlds. Assuming that arguments represent pairwise independent events, and that each defeat represents an event conditioned by the occurrence of its argument events but independent from any other event, the interpretation for the PrAF $\mathcal{F} = \langle A, P_A, D, P_D\rangle$ is as follows. For each possible world $w \in pw(\mathcal{F})$, w is assigned by I the probability:

$$I(w) = \prod_{a\in Arg(w)} P_A(a) \times \prod_{a\in A\setminus Arg(w)} (1 - P_A(a)) \times \times \prod_{\delta\in Def(w)} P_D(\delta) \times \prod_{\delta\in \overline{D}(w)\setminus Def(w)} (1 - P_D(\delta))$$

[1] Assigning probability equal to 0 to arguments/defeats is useless.

where $\overline{D}(w)$ is the set of defeats that may appear in the possible world w, that is $\overline{D}(w) = D \cap (Arg(w) \times Arg(w))$. Hence, the probability of a possible world w is given by the product of four contributions: (*i*) the product of the probabilities of the arguments belonging to w; (*ii*) the product of the one's complements of the probabilities of the arguments that do not appear in w; (*iii*) the product of the conditional probabilities of the defeats in w (recall that a defeat $\delta = (a, b)$ may appear in w only if both a and b are in w); (*iv*) the product of the one's complements of the conditional probabilities of the defeats that, though they may appear in w, they do not.

Example 3 *Continuing our running example, the interpretation I for $\mathcal{F}$ is as follows. First of all, observe that, for each possible world $w \in pw(\mathcal{F})$, if both arguments b and c belong to $Arg(w)$ and $\delta_2 \notin Def(w)$ or $\delta_3 \notin Def(w)$, then $I(w) = 0$. The probabilities of the other possible worlds are the following:*
$I(w_1) = (1 - P_A(a)) \times (1 - P_A(b)) \times (1 - P_A(c)) = .024$;
$I(w_2) = .216$; $I(w_3) = .056$; $I(w_4) = .006$;
$I(w_5) = P_A(a) \times P_A(b) \times (1 - P_A(c)) \times (1 - P_D(\delta_1)) = .0504$;
$I(w_6) = .054$; $I(w_9) = .4536$; $I(w_{12}) = .014$;
$I(w_{16}) = .1134$; $I(w_{19}) = .0126$.

The probability that a set S of arguments is an extension according to a given semantics *sem* is defined as the sum of the probabilities of the possible worlds w for which S is an extension according to *sem*, i.e., *ext*(w, sem, S) = `true`.

Definition 2 ($\mathbf{Pr}^{sem}_{\mathcal{F}}(S)$) *Given a PrAF $\mathcal{F}$, a set S of arguments, and a semantics sem, the probability $Pr^{sem}_{\mathcal{F}}(S)$ that S is an extension according to sem is*

$$Pr^{sem}_{\mathcal{F}}(S) = \sum_{w \in pw(\mathcal{F}) \wedge ext(w,sem,S)} I(w).$$

The following example shows usages of this definition.

Example 4 *In our running example, the probabilities that the sets $S_1 = \{b\}$, and $S_2 = \{ac\}$ are admissible are as follows:*
$Pr^{\mathtt{ad}}_{\mathcal{F}}(S_1) = I(w_3) + I(w_5) + I(w_9) + I(w_{12}) + I(w_{16}) + I(w_{19}) = .7$
$Pr^{\mathtt{ad}}_{\mathcal{F}}(S_2) = I(w_6) + I(w_{16}) + I(w_{19}) = .18$

It is easy to see that, as S_2 is stable in the same possible worlds where it is admissible, $Pr^{\mathtt{st}}_{\mathcal{F}}(S_2) = Pr^{\mathtt{ad}}_{\mathcal{F}}(S_2)$. The probability that S_1 is stable is as follows: $Pr^{\mathtt{st}}_{\mathcal{F}}(S_1) = I(w_3) + I(w_9) + I(w_{12}) + I(w_{16}) = .637$.

Obviously, computing $Pr^{sem}_{\mathcal{F}}(S)$ by directly applying Definition 2 would require exponential time, since it relies on summing the probabilities of an exponential number of possible worlds. However, this does not rule out the possibility that efficient strategies for computing $Pr^{sem}_{\mathcal{F}}(S)$ exist. In fact, in the next section we determine tractable and intractable cases for the problem of computing $Pr^{sem}_{\mathcal{F}}(S)$.

3 Complexity of Probabilistic Argumentation

We characterize the complexity of the following problem.

Definition 3 (Problem $\textsc{Prob}^{sem}_{\mathcal{F}}(S)$) *Given a PrAF $\mathcal{F} = \langle A, P_A, D, P_D\rangle$, a set $S \subseteq A$ of arguments, and a semantics sem in $\{\mathtt{ad}, \mathtt{st}, \mathtt{co}, \mathtt{gr}, \mathtt{pr}, \mathtt{id}\}$, $\textsc{Prob}^{sem}_{\mathcal{F}}(S)$ is the problem of computing $Pr^{sem}_{\mathcal{F}}(S)$.*

In the complexity analysis, the size of the input of our problem is the sum of the sizes of A and D [2].

Our results are summarized in the rightmost column of Table 1. We first show that $\textsc{Prob}^{sem}_{\mathcal{F}}(S)$ is tractable for the *admissible* and *stable* semantics (Section 3.1). Next, we provide lower and upper bounds on the complexity $\textsc{Prob}^{sem}_{\mathcal{F}}(S)$ for the other semantics (Section 3.2).

Due to space limitations, we only give a hint on some of the proofs of our results.

3.1 Tractable Cases

We first show that $\textsc{Prob}^{\mathtt{ad}}_{\mathcal{F}}(S)$ can be solved in polynomial time. Then, we show a similar result for the stable semantics.

The tractability of $\textsc{Prob}^{\mathtt{ad}}_{\mathcal{F}}(S)$ follows from the following lemma, which states that $Pr^{\mathtt{ad}}_{\mathcal{F}}(S)$ can be determined by evaluating an expression which only involves the probabilities of the arguments and defeats of the PrAF.

Lemma 1 *Given a PrAF $\mathcal{F} = \langle A, P_A, D, P_D\rangle$ and a set $S \subseteq A$, $Pr^{\mathtt{ad}}_{\mathcal{F}}(S) = P_1(S) \cdot P_2(S) \cdot P_3(S)$, where[3]:*

$P_1(S) = \prod_{a \in S} P_A(a)$,

$P_2(S) = \prod_{\substack{\langle a,b\rangle \in D \\ \wedge a \in S \\ \wedge b \in S}} \Big(1 - P_D(\langle a,b\rangle)\Big)$, *and*

$P_3(S) = \prod_{d \in A \setminus S} \Big(P_{31}(S,d) + P_{32}(S,d) + P_{33}(S,d)\Big)$, *where:*

$P_{31}(S,d) = 1 - P_A(d)$,

$P_{32}(S,d) = P_A(d) \times \prod_{\substack{\langle d,b\rangle \in D \\ \wedge b \in S}} \Big(1 - P_D(\langle d,b\rangle)\Big)$,

$$P_{33}(S,d) = P_A(d) \times \Big(1 - \prod_{\substack{\langle d,b\rangle \in D \\ \wedge b \in S}} \Big(1 - P_D(\langle d,b\rangle)\Big)\Big) \times \times \Big(1 - \prod_{\substack{\langle a,d\rangle \in D \\ \wedge a \in S}} \Big(1 - P_D(\langle a,d\rangle)\Big)\Big).$$

This lemma can be proved by observing that the fact that a set S of arguments is admissible can be expressed as the probabilistic event $E(S) = e_1(S) \wedge e_2(S) \wedge e_3(S)$, where:

- $e_1(S)$ is the event that all of the arguments in S occur;
- $e_2(S)$ is the event that, given that $e_1(S)$ holds, S is conflict-free; and
- $e_3(S)$ is the event that, given that $e_1(S)$ holds, for all the arguments d in $A \setminus S$, one of the following events holds:
  - $e_{31}(S, d)$: d does not occur,
  - $e_{32}(S, d)$: d occurs and no defeat (d, b), with $b \in S$, occurs,
  - $e_{33}(S, d)$: d occurs, there is at least one argument $b \in S$ such that (d, b) occurs, and there is at least one argument $a \in S$ such that (a, d) occurs.

We point out that the occurrence of events $e_2(S)$ and $e_3(S)$ is conditioned by the occurrence of event $e_1(S)$, and that $e_2(S)$ and $e_3(S)$ are independent from one another. Moreover, for each pair of distinct arguments $d', d'' \in A/S$,

[2]The sizes of the rational numbers encoding the probabilities of arguments and defeats are assumed to be bounded by a constant.

[3]Note that, an empty product evaluates to 1.

subevents $e_{31}(S,d'), e_{32}(S,d')$ and $e_{33}(S,d')$ are independent from subevents $e_{31}(S,d'')$, $e_{32}(S,d'')$ and $e_{33}(S,d'')$. Finally, note that, for each $d \in A/S$, subevents $e_{31}(S,d)$, $e_{32}(S,d)$ and $e_{33}(S,d)$ are mutually exclusive.

Starting from the above considerations, it can be shown that $Pr(E(S)) = Pr(e_1(S)) \cdot Pr(e_2(S)) \cdot Pr(e_3(S))$, where $Pr(e_i)$ denotes the probability of event e_i. Moreover, it is easy to see that $Pr(e_1(S))$ (resp., $Pr(e_2(S))$) is equal to the value of expression $P_1(S)$ (resp., $P_2(S)$) given in Lemma 1. Furthermore, it can be shown that $Pr(e_3(S))$ is equal to $P_3(S) = \prod_{d \in A \setminus S} \big(P_{31}(S,d) + P_{32}(S,d) + P_{33}(S,d)\big)$, where $P_{31}(S,d)$, $P_{32}(S,d)$ and $P_{33}(S,d)$ are the probabilities of the mutually exclusive events $e_{31}(S,d)$, $e_{32}(S,d)$, and $e_{33}(S,d)$, respectively.

Example 5 *Continuing our running example, where $S_1 = \{b\}$ and $S_2 = \{a,c\}$, $Pr^{\mathtt{ad}}_{\mathcal{F}}(S_1)$ can now be computed by applying Lemma 1 :*

$Pr^{\mathtt{ad}}_{\mathcal{F}}(S_1) = P_A(b) \times [(1 - P_A(a)) + P_A(a) \times 1 + P_A(a) \times (1-1) \times (1-(1-P_D(\delta_1)))] \times [(1-P_A(c)) + P_A(c) \times (1 - P_D(\delta_3)) + P_A(c) \times (1-(1-P_D(\delta_3))) \times (1-(1-P_D(\delta_2)))] = .7 \times [.1 + .9 + 0] \times [.8 + 0 + .2] = .7$

Similarly, we can derive that $Pr^{\mathtt{ad}}_{\mathcal{F}}(S_2) = .18$.

The result of Lemma 1 allows us to give a polynomial time algorithm to solve $\textsc{Prob}^{\mathtt{ad}}_{\mathcal{F}}(S)$: evaluate the expression of Lemma 1 by iterating on the arguments and defeats in the PrAF $\mathcal{F}$. The complexity this algorithm is as follows.

Theorem 1 $\textsc{Prob}^{\mathtt{ad}}_{\mathcal{F}}(S)$ *can be solved in time $O(|S| \cdot |A|)$.*

We now address the problem of computing the probability that a given set S of arguments is stable for a given PrAF $\mathcal{F}$. We show that $\textsc{Prob}^{\mathtt{st}}_{\mathcal{F}}(S)$ can be solved in polynomial time by exploiting the result of the following lemma.

Lemma 2 *Given a PrAF $\mathcal{F} = \langle A, P_A, D, P_D\rangle$, and a set $S \subseteq A$, $Pr^{\mathtt{st}}_{\mathcal{F}}(S) = P_1(S) \cdot P_2(S) \cdot P_3(S)$, where $P_1(S)$ and $P_2(S)$ are defined as in Lemma 1 and $P_3(S) = \prod_{d \in A \setminus S} \Big(P_{31}(S,d) + P_{32}(S,d)\Big)$, where*

$P_{31}(S,d) = 1 - P_A(d)$, and

$P_{32}(S,d) = P_A(d) \times \Big(1 - \prod_{\langle a,d\rangle \in D \wedge a \in S} (1 - P_D(\langle a,d\rangle))\Big)$.

Similar to the case of admissible semantics, the fact that a set S of arguments is stable can be expressed as the probabilistic event $E'(S) = e_1(S) \wedge e_2(S) \wedge e'_3(S)$, where:

- $e_1(S)$ and $e_2(S)$ are the events introduced after Lemma 1, and
- $e'_3(S)$ is the event that, given that $e_1(S)$ holds, for all the arguments d in $A \setminus S$, one of the following events holds:
  - $e_{31}(S,d)$: d does not occur,
  - $e'_{32}(S,d)$: d occurs and it is defeated by S, that is, at least one defeat (a,d), with $a \in S$, occurs.

Reasoning analogously to the case of Lemma 1 it can be shown that $Pr(E'(S))$ is equal to $P_1(S) \cdot P_2(S) \cdot P_3(S)$.

Example 6 *Considering again our running example, from Lemma 2 we have that:*

$Pr^{\mathtt{st}}_{\mathcal{F}}(S_1) = P_A(b) \times [(1 - P_A(a)) + P_A(a) \times (1-(1-P_D(\delta_1)))] \times [(1-P_A(c)) + P_A(c) \times (1-(1-P_D(\delta_2)))] = .637$.

Reasoning analogously, $Pr^{\mathtt{st}}_{\mathcal{F}}(S_2)$ results equal to .18.

The complexity of an algorithm for solving $\textsc{Prob}^{\mathtt{st}}_{\mathcal{F}}(S)$, by exploiting Lemma 2, is stated in the following theorem.

Theorem 2 $\textsc{Prob}^{\mathtt{st}}_{\mathcal{F}}(S)$ *can be solved in time $O(|S| \cdot |A|)$.*

Theorems 1 and 2 show that, for the admissible and stable semantics, the problem of computing $Pr^{sem}_{\mathcal{F}}(S)$ is tractable. Thus, for $\textsc{Prob}^{\mathtt{ad}}_{\mathcal{F}}(S)$ and $\textsc{Prob}^{\mathtt{st}}_{\mathcal{F}}(S)$ there is no need to resort to approximate techniques for estimating $Pr^{\mathtt{ad}}_{\mathcal{F}}(S)$ and $Pr^{\mathtt{st}}_{\mathcal{F}}(S)$, as instead proposed in [Li *et al.*, 2011] where a Monte-Carlo simulation approach is adopted. In fact, our results show that $\textsc{Prob}^{\mathtt{ad}}_{\mathcal{F}}(S)$ and $\textsc{Prob}^{\mathtt{st}}_{\mathcal{F}}(S)$ belong to the same complexity class as $\textsc{Ver}^{\mathtt{ad}}(S)$ and $\textsc{Ver}^{\mathtt{st}}(S)$.

3.2 Hard Cases

We now characterize the computational complexity of the problem of computing the probability that a given set of arguments is an extension according to the complete, grounded, preferred, and ideal semantics. Specifically, we show that the lower bound on the complexity of $\textsc{Prob}^{sem}_{\mathcal{F}}(S)$ for each of these semantics is $FP^{\#P}$. This result motivates the use of approximate strategies, such as those adopted in [Li *et al.*, 2011], when dealing with these kinds of extensions. Moreover, we show that $FP^{\#P}$ is also an upper bound on the complexity of $\textsc{Prob}^{sem}_{\mathcal{F}}(S)$ for each of these semantics, thus providing a tight complexity characterization for $\textsc{Prob}^{sem}_{\mathcal{F}}(S)$.

Theorem 3 *For sem $\in \{\mathtt{co}, \mathtt{gr}, \mathtt{pr}, \mathtt{id}\}$, it holds that $\textsc{Prob}^{sem}_{\mathcal{F}}(S)$ is $FP^{\#P}$-complete.*

We give a hint on the $FP^{\#P}$-hardness proof of Theorem 3. The $FP^{\#P}$ hardness of $\textsc{Prob}^{sem}_{\mathcal{F}}(S)$ for *sem* $\in \{\mathtt{co}, \mathtt{gr}\}$ is shown by providing a Cook reduction to our problem from the $\#P$-hard problem *#PP2DNF* (*Partitioned Positive 2DNF*) [Provan and Ball, 1983], that is, the problem of counting the number of satisfying assignments of a DNF formula $\phi = C_1 \vee C_2 \vee \ldots \vee C_k$ whose propositional variables are positive and can be partitioned into two sets $X = \{x_1, \ldots, x_n\}$ and $Y = \{y_1, \ldots, y_m\}$, and each clause C_i has the form $x_j \wedge y_\ell$, with $x_j \in X$ and $y_\ell \in Y$.

Given a *#PP2DNF* instance ϕ, we construct the PrAF $\mathcal{F}_\phi = \langle A, P_A, D, P_D\rangle$ such that (*i*) A contains an argument for each propositional variable in ϕ, an argument c_ℓ for each clause C_ℓ of ϕ, and an argument s; (*ii*) D contains the defeats (x_i, c_ℓ) and (y_j, c_ℓ) for each clause $C_\ell = x_i \wedge y_j$ of ϕ, and the defeats (s, x_i) and (x_i, x_i) (resp., (s, y_j) and (y_j, y_j)) for each variable x_i (resp., y_j) of ϕ; (*iii*) P_A assigns probability 1 to all the arguments in A; P_D assigns probability 1 to all the defeats in D except the defeats (s, x_i) and (s, y_j), which are assigned .5. As an example of this construction, the PrAF $\mathcal{F}_{\phi'}$ resulting from $\phi' = (x_1 \wedge y_1) \vee (x_2 \wedge y_1) \vee (x_3 \wedge y_2) \vee (x_3 \wedge y_3)$ is represented by the direct graph shown in Figure 1(a). In this graph, (*i*) every node represents an argument, (*ii*) every edge represents a defeat, and (*iii*) for each argument/defeat, its probability (if different from 1) is specified near to it.

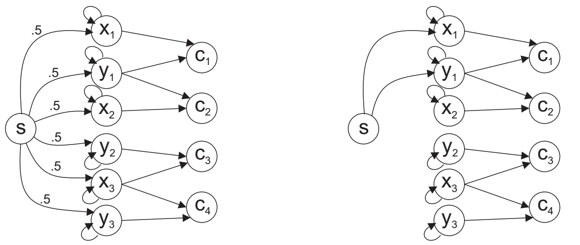

Figure 1: (a) PrAF $\mathcal{F}_{\phi'}$ (b) Possible world $w_{\tau'}$

It can be shown that there is a bijection $b : \mathcal{T} \to pw(\mathcal{F})$ between the set $\mathcal{T}$ of truth assignments of ϕ and the set $pw(\mathcal{F})$ of possible worlds. For instance, in the example above, the truth assignment $\tau' = x_1/1, x_2/0, x_3/0, y_1/1, y_2/0, y_3/0$ for ϕ' one-to-one corresponds to the world $w_{\tau'}$ shown in Figure 1(b), where the fact that variable x_1 (resp., y_1) is assigned 1 by τ is encoded by the occurrence of the defeat (s, x_1) (resp., (s, y_1)) in $w_{\tau'}$.

We proved that, given a truth assignment τ for the variables of ϕ, ϕ evaluates to *true* under τ iff $S = \{s\}$ is not a complete extension in the world $w = b(\tau)$. Continuing our example, it is easy to see that the truth assignment τ' makes ϕ' true and $\{s\}$ is not a complete extension in $w_{\tau'}$ since s defeats both x_1 and y_1 which makes c_1 acceptable w.r.t $\{s\}$.

Furthermore, it can be shown that the number of satisfying assignments $\#\phi$ of ϕ is equal to $2^{n+m} \cdot (1 - Pr_{\mathcal{F}}^{\mathsf{CO}}(\{s\}))$, which completes the $FP^{\#P}$-hardness proof for $\textsc{Prob}_{\mathcal{F}}^{\mathsf{CO}}(S)$.

To summarize, our complexity analysis shows that, while the problem $\textsc{Ver}^{sem}(S)$ of verifying whether a set S of arguments is an extension according to *sem* is intractable only for the preferred and the ideal semantics, its probabilistic counterpart $\textsc{Prob}_{\mathcal{F}}^{sem}(S)$ results to be intractable for the grounded and the complete semantics as well.

4 Related work

Recently approaches for handling uncertainty in AAFs by relying on probability theory have been proposed in [Dung and Thang, 2010; Rienstra, 2012; Li *et al.*, 2011; Thimm, 2012]. With the aim of modeling jury-based dispute resolutions, [Dung and Thang, 2010] proposed a PrAF where uncertainty is taken into account by specifying probability distribution functions (PDFs) over possible worlds and shown how an instance of the proposed PrAF can be obtained by specifying a probabilistic assumption-based argumentation framework (introduced by themselves). In the same spirit, [Rienstra, 2012] defined a PrAF as a PDF over the set of possible worlds, and introduced a probabilistic version of a fragment of ASPIC framework [Prakken, 2010] that can be used to instantiate the proposed PrAF. Differently from the two previous approaches, [Li *et al.*, 2011] proposed a PrAF where probabilities are directly associated with arguments and defeats, instead of being associated with possible worlds. [Li *et al.*, 2011] claimed that computing the probability $Pr(S)$ that a set S of arguments belongs to an extension requires exponential time for every semantics, and then proposes a Monte-Carlo simulation approach to approximate $Pr(S)$. In [Li *et al.*, 2011], as well as in [Dung and Thang, 2010], [Rienstra, 2012], $Pr(S)$ is defined as the sum of the probabilities of the possible worlds where S is an extension, according to a given semantics. [Thimm, 2012], instead, did not define a probabilist version of a classical semantics, but introduced a new probabilistic semantics. This semantics is based on specifying a class of PDFs, called *p-justifiable* PDFs, over sets of possible AAFs, and shown that this probabilistic semantics generalizes the complete semantics.

Though in the above-cited works probability theory is recognized as a fundamental tool to model uncertainty, a deeper understanding of the role of probability theory in abstract argumentation was developed only later in [Hunter, 2012; 2013], where the connection among argumentation theory, classical logic, and probability theory was investigated.

Besides the approaches that model uncertainty in AAFs by relying on probability theory, many proposals have been made where uncertainty is represented by exploiting weights or preferences on arguments and/or defeats [Bench-Capon, 2003; Amgoud and Vesic, 2011; Amgoud and Cayrol, 2002; Modgil, 2009; Dunne *et al.*, 2011; Coste-Marquis *et al.*, 2012], or by relying on the possibility theory [Amgoud and Prade, 2004; Alsinet *et al.*, 2008a; 2008b].

Although the approaches based on weights, preferences, possibilities, or probabilities to model uncertainty have been proved to be effective in different contexts, there is no common agreement on what kind of approach should be used in general. In this regard, [Hunter, 2012; 2013] observed that the probability-based approaches may take advantage from relying on a well-established and well-founded theory, whereas the approaches based on weights or preferences do not conform to well-established theories yet.

The computational complexity of computing extensions has been throughly investigated for AAFs [Dunne and Wooldridge, 2009; Dunne, 2009], for the case of adding weights to AAFs [Dunne *et al.*, 2011; Coste-Marquis *et al.*, 2012], and for the case of using preferences [Amgoud and Vesic, 2011]. To the best of our knowledge, this work is the first one characterizing the computational complexity of the problem of computing the probability that a set of arguments is an extension according to a given semantics.

5 Conclusions and future work

In this paper we focused on the probabilistic argumentation framework proposed in [Li *et al.*, 2011], and characterized the complexity of computing the probability that a set of arguments is an extension according to a given semantics. We showed that the complexity of this problem is either $PTIME$ or $FP^{\#P}$, depending on the considered semantics. Extending the complexity study presented in this paper to other AAF semantics, such as the *semi-stable* [Caminada, 2006], the *stage* [Caminada, 2010], the *CF2* [Baroni *et al.*, 2005] and the *prudent* [Coste-Marquis *et al.*, 2005] semantics, is an interesting direction for future work. Another promising direction is that of characterizing the complexity of the probabilistic version of the *credulous/sceptical acceptance* problem, that is, the problem of computing the probability that an argument belongs to any/every extension according to a given semantics.

References

[Alsinet *et al.*, 2008a] Teresa Alsinet, Carlos Iván Chesñevar, Lluis Godo, Sandra Sandri, and Guillermo Ricardo Simari. Formalizing argumentative reasoning in a possibilistic logic programming setting with fuzzy unification. *Int. J. Approx. Reasoning*, 48(3), 2008.

[Alsinet *et al.*, 2008b] Teresa Alsinet, Carlos Iván Chesñevar, Lluis Godo, and Guillermo Ricardo Simari. A logic programming framework for possibilistic argumentation: Formalization and logical properties. *Fuzzy Sets and Systems*, 159(10):1208–1228, 2008.

[Amgoud and Cayrol, 2002] Leila Amgoud and Claudette Cayrol. A reasoning model based on the production of acceptable arguments. *Ann. Math. Artif. Intell.*, 34(1-3):197–215, 2002.

[Amgoud and Prade, 2004] Leila Amgoud and Henri Prade. Reaching agreement through argumentation: A possibilistic approach. In *KR*, pages 175–182, 2004.

[Amgoud and Vesic, 2011] Leila Amgoud and Srdjan Vesic. A new approach for preference-based argumentation frameworks. *Ann. Math. Artif. Intell.*, 63(2):149–183, 2011.

[Baroni and Giacomin, 2009] Pietro Baroni and Massimiliano Giacomin. Semantics of abstract argument systems. In *Argumentation in Artificial Intelligence*, pages 25–44. 2009.

[Baroni *et al.*, 2005] Pietro Baroni, Massimiliano Giacomin, and Giovanni Guida. Scc-recursiveness: a general schema for argumentation semantics. *Artif. Intell.*, 168(1-2):162–210, 2005.

[Bench-Capon and Dunne, 2007] Trevor J. M. Bench-Capon and Paul E. Dunne. Argumentation in artificial intelligence. *Artif. Intell.*, 171(10-15):619–641, 2007.

[Bench-Capon, 2003] Trevor J. M. Bench-Capon. Persuasion in practical argument using value-based argumentation frameworks. *J. Log. Comput.*, 13(3):429–448, 2003.

[Besnard and Hunter, 2008] Philippe Besnard and Anthony Hunter, editors. *Elements Of Argumentation*. The MIT Press, 2008.

[Caminada, 2006] Martin Caminada. Semi-stable semantics. In *COMMA*, pages 121–130, 2006.

[Caminada, 2010] Martin Caminada. An algorithm for stage semantics. In *COMMA*, pages 147–158, 2010.

[Coste-Marquis *et al.*, 2005] Sylvie Coste-Marquis, Caroline Devred, and Pierre Marquis. Prudent semantics for argumentation frameworks. In *ICTAI*, pages 568–572, 2005.

[Coste-Marquis *et al.*, 2012] Sylvie Coste-Marquis, Sébastien Konieczny, Pierre Marquis, and Mohand Akli Ouali. Weighted attacks in argumentation frameworks. In *KR*, 2012.

[Dung and Thang, 2010] Phan Minh Dung and Phan Minh Thang. Towards (probabilistic) argumentation for jury-based dispute resolution. In *COMMA*, pages 171–182, 2010.

[Dung *et al.*, 2007] Phan Minh Dung, Paolo Mancarella, and Francesca Toni. Computing ideal sceptical argumentation. *Artif. Intell.*, 171(10-15):642–674, 2007.

[Dung, 1995] Phan Minh Dung. On the acceptability of arguments and its fundamental role in nonmonotonic reasoning, logic programming and n-person games. *Artif. Intell.*, 77(2):321–358, 1995.

[Dunne and Wooldridge, 2009] Paul E. Dunne and Michael Wooldridge. Complexity of abstract argumentation. In *Argumentation in Artificial Intelligence*, pages 85–104. 2009.

[Dunne *et al.*, 2011] Paul E. Dunne, Anthony Hunter, Peter McBurney, Simon Parsons, and Michael Wooldridge. Weighted argument systems: Basic definitions, algorithms, and complexity results. *Artif. Intell.*, 175(2):457–486, 2011.

[Dunne, 2009] Paul E. Dunne. The computational complexity of ideal semantics. *Artif. Intell.*, 173(18):1559–1591, 2009.

[Hunter, 2012] Anthony Hunter. Some foundations for probabilistic abstract argumentation. In *COMMA*, pages 117–128, 2012.

[Hunter, 2013] Anthony Hunter. A probabilistic approach to modelling uncertain logical arguments. *Int. J. Approx. Reasoning*, 54(1):47–81, 2013.

[Li *et al.*, 2011] Hengfei Li, Nir Oren, and Timothy J. Norman. Probabilistic argumentation frameworks. In *TAFA*, pages 1–16, 2011.

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

[Prakken, 2010] Henry Prakken. An abstract framework for argumentation with structured arguments. *Argument & Computation*, 1(2):93–124, 2010.

[Provan and Ball, 1983] J. Scott Provan and Michael O. Ball. The complexity of counting cuts and of computing the probability that a graph is connected. *SIAM J. Comput.*, 12(4):777–788, 1983.

[Rahwan and Simari, 2009] Iyad Rahwan and Guillermo R. Simari, editors. *Argumentation in Artificial Intelligence*. Springer, 2009.

[Rienstra, 2012] Tjitze Rienstra. Towards a probabilistic dung-style argumentation system. In *AT*, pages 138–152, 2012.

[Thimm, 2012] Matthias Thimm. A probabilistic semantics for abstract argumentation. In *ECAI*, pages 750–755, 2012.

[Valiant, 1979] Leslie G. Valiant. The complexity of computing the permanent. *Theor. Comput. Sci.*, 8:189–201, 1979.