Artificial Intelligence 323 (2023) 103967

Contents lists available at ScienceDirect

Artificial Intelligence

journal homepage: www.elsevier.com/locate/artint

Explainable acceptance in probabilistic and incomplete abstract argumentation frameworks ☆

Gianvincenzo Alfano [a,*], Marco Calautti [b], Sergio Greco [a], Francesco Parisi [a], Irina Trubitsyna [a]

[a] *Department of Informatics, Modeling, Electronics and System Engineering, University of Calabria, Italy*
[b] *Computer Science Department, University of Milan, Italy*

ARTICLE INFO

Article history:
Received 26 October 2022
Received in revised form 16 June 2023
Accepted 23 June 2023
Available online 27 June 2023

Keywords:
Formal argumentation
Explanations
Probabilistic argumentation framework
Incomplete argumentation framework

ABSTRACT

Dung's Argumentation Framework (AF) has been extended in several directions, including the possibility of representing uncertainty about the existence of arguments and attacks. In this regard, two main proposals have been introduced in the literature: Probabilistic Argumentation Framework (PrAF) and Incomplete Argumentation Framework (iAF). PrAF is an extension of AF with probability theory, thus representing quantified uncertainty. In contrast, iAF represents unquantified uncertainty, that is it can be seen as a special case where we only know that some elements (arguments or attacks) are uncertain. In this paper, we first address the problem of computing the probability that a given argument is accepted in PrAF. This is carried out by introducing the concept of probabilistic explanation for any given (probabilistic) extension. We show that the complexity of the problem is $FP^{\#P}$-hard and propose polynomial approximation algorithms with bounded additive error for PrAFs where odd-length cycles are forbidden. We investigate the approximate complexity of the related $FP^{\#P}$-hard problems of credulous and skeptical acceptance in PrAF, showing that they are generally harder than the problem of computing the probability that a given argument is accepted. Next we consider iAF and, after showing some equivalence properties among classes of iAFs, we study iAF as a special case of PrAF where uncertain elements have associated a probability equal to 1/2. Finally, given this result, we investigate the relationships between iAF acceptance problems and probabilistic acceptance in PrAF.

1. Introduction

As humans, we use argumentation to explain our reasons for or against claims in our discussions, persuade other people, and derive conclusions in a step-wise fashion. Most of the situations where argumentation takes place are inherently characterized by the presence of controversial information. Enabling automated systems to process such kind of information, much in the same way as organized human discussions are carried out, is an important challenge that has deserved increasing attention from the Artificial Intelligence community in the last decades. This has led to the development of an important and active research area called *formal argumentation* [2–4], that has been explored in several application contexts,

☆ This paper is a substantially revised and expanded version of [1].
* Corresponding author.
E-mail address: g.alfano@dimes.unical.it (G. Alfano).

https://doi.org/10.1016/j.artint.2023.103967

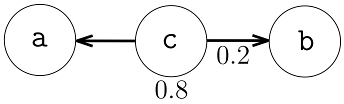

Fig. 1. Probabilistic argumentation framework modeling the robbery case.

e.g. legal reasoning [5], decision support systems [6], E-Democracy [7], healthcare [8], medical applications [9], financial analysis [10], explanation of results [11,1], as well as multi-agent systems and social networks [12].

In particular, an abstract Argumentation Framework (AF) is a simple, yet powerful formalism for modeling disputes between two or more agents [13]. An AF consists of a set of *arguments* and a binary *attack* relation over the set of arguments that specifies the interactions between arguments: intuitively, if argument a attacks argument b, then b is acceptable only if a is not. Hence, arguments are abstract entities whose role is entirely determined by the interactions specified by the attack relation.

Recently, there has been an increasing interest in extending argumentation frameworks to manage uncertain information. This has been carried out by either considering quantified uncertainty about the existence of arguments and attacks, thus combining formal argumentation with probability theory, or considering unquantified uncertainty by explicitly denoting the elements (arguments and attacks) which are uncertain. In fact, Probabilistic Argumentation [14] can be viewed as part of the several proposals that have been made in the last decades for extending reasoning tasks in AI frameworks with probabilities. These include for instance Probabilistic SAT (PSAT) [15], Probabilistic Logic [16], Probabilistic Logic Programming [17], and Probabilistic Databases [18].

One of the most popular approaches based on probability theory for modeling the uncertainty is the so called *constellations* approach [19–23], where alternative scenarios, called *possible worlds*, are associated with probabilities. In particular, in a *Probabilistic Argumentation Framework* (PrAF) [23–29] a probability distribution function (PDF) on the set of possible worlds is entailed by the probabilities that are associated with arguments and attacks.

Consider for instance the following scenario (inspired by an example in [30] that has been then revisited in [24]), where the defense attorney of John and Peter wants to model the situation of a robbery case involving his clients. The arguments of the case are the following, where Harry is a potential witness:

- a = John says he was not in town when the robbery took place, and therefore he claims to be innocent.
- b = Peter says he was at home watching TV when the robbery took place, and therefore he also claims to be innocent.
- c = Harry says that he is certain to have seen John outside the bank just before the robbery took place, and he also saw a second person outside the bank that could be Peter with probability 20%.

Considering what the three people have declared, the lawyer knows that nobody accused Harry to be involved in the robbery, whereas Harry could testify against both John and Peter. However, he is also confident that the probability that Harry will testify is 80% since he heard that Harry is a bit reluctant to testify. This can be modeled by means of an abstract argumentation framework where arguments and attacks represent probabilistic events, as shown in Fig. 1 where probabilities of arguments and attacks are specified only if they are less than 1. Herein, argument a (resp. b) represents the fact that John (resp. Peter) is innocent, whereas argument c represents the fact that Harry testifies. Moreover, the attack from c to a (resp. from c to b) represents the that fact that Harry blames John (resp. Peter). The probability of a and b is 1 as they are assumed to be certain, while a probability of 0.8 is associated to c. Moreover, associating probability 1 to the attack $(\mathtt{c}, \mathtt{a})$ means that argument c attacks a with certainty, while associating probability 0.2 to the attack $(\mathtt{c}, \mathtt{b})$ means that c attacks b with a degree of uncertainty represented by that probability, since Harry is sure he saw John but he is not sure he saw Peter. As it will be clear in what follows, the lawyer can conclude that John (resp. Peter) will be judged as innocent with probability 20% (resp. 84%).

Intuitively, PrAF is a combination of two powerful approaches to reasoning and decision making: probabilistic reasoning and abstract argumentation. Probabilities are assigned to arguments and attacks to indicate their degree of uncertainty. One of the benefits of probabilistic abstract argumentation is its ability to handle quantified uncertainty in the analysis. In fact, PrAF can help to model and analyze situations where there is uncertainty by capturing both the relationships between arguments and the uncertainty degrees of arguments and attacks. In this regard, it is worth mentioning that the need for probabilistic argumentation has been also supported by some empirical evaluations [31].

The next example introduces a PrAF that will be often used in the rest of the paper to explain new concepts and definitions.

Example 1. Consider a PrAF $\Delta = \langle\{\mathtt{fish}, \mathtt{meat}, \mathtt{white}, \mathtt{red}\}, \{(\mathtt{fish}, \mathtt{meat}), (\mathtt{meat}, \mathtt{fish}), (\mathtt{meat}, \mathtt{white}), (\mathtt{white}, \mathtt{red}), (\mathtt{red}, \mathtt{white})\}, \{\mathtt{fish}/0.6, \mathtt{white}/\ 0.8\}\rangle$, whose corresponding graph is shown in Fig. 2, where nodes and edges represent arguments and attacks, respectively, and probabilities different from 1 are specified nearby them. For the sake of brevity, we do not specify the probabilities of certain elements in Δ (all the other elements different from fish and white have probability 1). Intuitively, Δ describes what a person is going to have for lunch as follows. They will have either fish or meat, and will drink either white wine or red wine. However, if they will have meat, then they will not

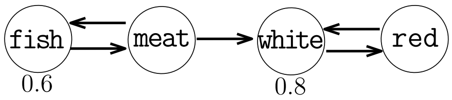

Fig. 2. Probabilistic argumentation framework Δ of Example 1.

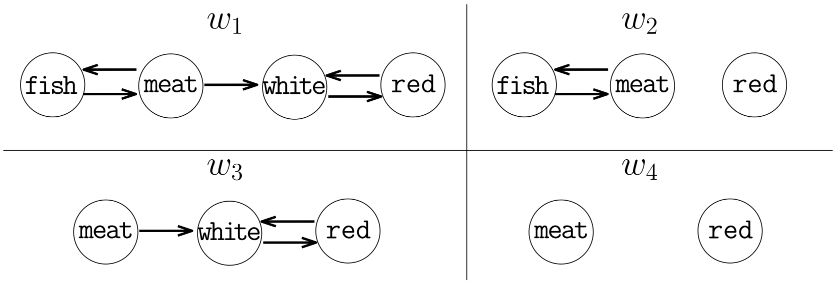

Fig. 3. Possible worlds of the probabilistic argumentation framework Δ of Example 1.

drink white wine. Furthermore, the probability that fish is available is 0.6, whereas the probability that white wine is available is 0.8. □

In this paper we do not address the problem of assigning probabilities to arguments or attacks, as instead done e.g. in [22,32], and assume they are given.

Several argumentation semantics—e.g. *grounded* (gr), *complete* (co), *preferred* (pr), *stable* (st), and *semi-stable* (sst)—have been defined for AFs, leading to the characterization of *σ-extensions*, which intuitively consist of the sets of arguments that can be collectively accepted under semantics σ. Consider for instance the deterministic version of the PrAF in Example 1, obtained by assuming that all arguments are certain (i.e. they have probability 1). Considering the preferred semantics, the pr-extensions are $E_1 = \{\mathtt{fish}, \mathtt{white}\}$, $E_2 = \{\mathtt{fish}, \mathtt{red}\}$, and $E_3 = \{\mathtt{meat}, \mathtt{red}\}$.

The semantics of a PrAF is given by considering all possible worlds (i.e. AFs) obtained by removing consistent subsets of the probabilistic elements. Here, for consistent subset we mean any subset of probabilistic elements (arguments and attacks) whose deletion from the initial framework results in an AF (for instance we cannot delete an argument without also deleting the attacks towards or from that argument). Every possible world has associated a probability value derived from the probabilities of the elements that have been kept or removed. Moreover, every possible world admits a set of σ-extensions. As we shall see, removing elements with probability 1 makes no sense as it would give rise to an AF which would make no contribution in the calculation of the probability of acceptance of arguments or of the existence of extensions, being its probability (of existing) equal to 0. As shown in the next example, the probability of a possible world w is computed by multiplying the probabilities of the elements occurring in w and the complement to 1 of the probabilities of the elements not occurring in w (see Equation (1) for the formal definition of probability of a possible world).

Example 2. Continuing with Example 1, the (non-zero probability) possible worlds of Δ (see Fig. 3) are as follows, where arguments are denoted by their initials for the sake of brevity:

- $w_1 = \langle\{\mathtt{f}, \mathtt{m}, \mathtt{w}, \mathtt{r}\}, \{(\mathtt{f}, \mathtt{m}), (\mathtt{m}, \mathtt{f}), (\mathtt{m}, \mathtt{w}), (\mathtt{w}, \mathtt{r}), (\mathtt{r}, \mathtt{w})\}\rangle$;
- $w_2 = \langle\{\mathtt{f}, \mathtt{m}, \mathtt{r}\}, \{(\mathtt{f}, \mathtt{m}), (\mathtt{m}, \mathtt{f})\}\rangle$;
- $w_3 = \langle\{\mathtt{m}, \mathtt{w}, \mathtt{r}\}, \{(\mathtt{m}, \mathtt{w}), (\mathtt{w}, \mathtt{r}), (\mathtt{r}, \mathtt{w})\}\rangle$;
- $w_4 = \langle\{\mathtt{m}, \mathtt{r}\}, \{\}\rangle$.

For instance, w_1 is the AF obtained from Δ by keeping all the arguments and attacks, while w_2 is obtained from Δ by removing white and, consistently with this, the attacks towards/from it. The probability of a possible world w_i is obtained by multiplying the probabilities $P(a)$ of each argument a occurring in w_i and the probabilities $(1 - P(b))$ of every argument b not occurring in w_i. For instance, the probability of world w_2 is $P(\mathtt{fish}) \cdot P(\mathtt{meat}) \cdot (1 - P(\mathtt{white})) \cdot P(\mathtt{red}) = 0.6 \cdot 1 \cdot (1 - 0.8) \cdot 1 = 0.12$. Thus, the probabilities of w_1, w_2, w_3, and w_4 are 0.48, 0.12, 0.32, and 0.08, respectively (hence they are called non-zero probability possible worlds). Since w_1 coincides with the deterministic version of Δ, its pr-extensions are E_1, E_2, and E_3 given earlier. The pr-extensions of w_2 are E_2 and E_3, while w_3 and w_4 admit only E_3 as their preferred extension. □

Interesting problems recently investigated in the context of probabilistic argumentation are *probabilistic credulous acceptance* (PrCA) and *probabilistic skeptical acceptance* (PrSA) [33,26]. In particular, given a PrAF Δ whose set of arguments is A, a goal argument $g \in A$ and a semantics σ, PrCA is the problem of computing the probability $PrCA^{\sigma}_{\Delta}(g)$ that the goal g is credulously accepted, that is, there is a possible world w of Δ such that g belongs to a σ-extension of w. Moreover, PrSA is the problem of computing the probability $PrSA^{\sigma}_{\Delta}(g)$ that the goal g is skeptically accepted, that is, g is credulously accepted and belongs to all σ-extensions of w (the probabilities $PrCA^{\sigma}_{\Delta}(g)$ and $PrSA^{\sigma}_{\Delta}(g)$ are formally introduced in Definition 2).

However, the answer to these problems does not reflect our intuition of probability that a goal argument is accepted under a given semantics. For instance, considering the PrAF Δ of Fig. 2, the probability that meat is credulously accepted under preferred semantics is 1, whereas the probability that meat is skeptically accepted under preferred semantics is 0.4. However, the fact that $PrCA^{\mathtt{pr}}_{\Delta}(\mathtt{meat}) = 1$ does not mean that the person in our example will surely have meat in any scenario (i.e. possible world). In fact, even if meat belongs to at least one preferred extension of every world of Δ, we expect that the probability of acceptance of meat should be lower than 1. Indeed, in any possible world, the presence of multiple extensions is an additional source of uncertainty that should be taken into account.

To better grasp the issue behind the probability of credulous acceptance, consider the following AF (where all elements are certain): $\Lambda = \langle\{\mathtt{fish}, \mathtt{meat}\}, \{(\mathtt{fish}, \mathtt{meat}), (\mathtt{meat}, \mathtt{fish})\}\rangle$ saying that fish and meat are mutually exclusive. Again, the probability that a person will have meat is 1, under probabilistic credulous acceptance, when considering the preferred semantics, whereas we believe that the expected answer should be 0.5. Moreover, if we consider AF w_1 of Example 2 (that can be obtained from Λ by adding arguments white and red and attacks (white, red), (red, white) and (meat, white)) we expect that the probability of having meat does not change.

With the aim of providing more intuitive answers for probabilistic acceptance, in this paper we investigate a new problem that we call *Probabilistic Acceptance* (denoted as PrA, or PrA[σ] when considering a given semantics σ), i.e. given a PrAF Δ and a goal argument g, compute the probability that g is accepted under semantics $\sigma \in \{\mathtt{gr}, \mathtt{co}, \mathtt{pr}, \mathtt{st}, \mathtt{sst}\}$. In our framework, acceptance still relies on σ-extensions but, differently from credulous acceptance, we get rid of the assumption that no uncertainty exists at the level of the extensions of a world (i.e. AF). In more detail, PrA[σ] implicitly assumes that a PDF over the set of σ-extensions of any AF (and thus of any possible world of PrAF Δ) is defined. Thus, a concrete instance of PrA is obtained after defining such a PDF. This can be carried out by exploiting the concept of *explanation* for an extension.

In general, in abstract argumentation an explanation for an extension E can be viewed as a (possibly minimal) subset $S \subseteq E$ such that, by assuming that the elements in S are acceptable, it turns out that all elements in $E \setminus S$ are "univocally" determined as acceptable (w.r.t. the underlying semantics). For instance, considering AF w_1 of Example 2, for the preferred extension $E = \{\mathtt{meat}, \mathtt{red}\}$, the set $S_1 = \{\mathtt{meat}\}$ is an explanation for E, whereas the set $S_2 = \{\mathtt{red}\}$ is not. In our perspective, explanations are sequences of "choices" to be made to justify how an extension is obtained and they provide a tool to assign probabilities to extensions. Integrating explanations in argumentation systems is important for enhancing the argumentation and persuasion capabilities of software agents [34–36]. For these reasons, several researchers have explored how to deal with explanations in formal argumentation [37–39].

In this paper, we explore an instantiation of PrA[σ] where the PDF over the set of σ-extensions of a world relies on the concept of *explanation*. We call this problem *Explanation-based Probabilistic Acceptance*, and denote it by PrEA (and PrEA[σ] for a specific semantics σ). Intuitively, an explanation for an σ-extension E is a *sequence of arguments* occurring in E that "justify" E. Every explanation is associated with a probability entailed by the possible choices that can be made when building it. These choices must be consistent with an ordering entailed by the strongly connected components of the given AF, and they are used to guide the construction of an extension. The sum of the probabilities of the explanations for an extension E gives the probability of E. Thus, we still assign to each possible world w of Δ a probability as in the standard way, but in addition propose to distinguish among extensions of a given world w by associating with them a probability based on explanations.

Example 3. Continuing with Example 1, take for instance the possible world w_1 having probability 0.48. As shown in Example 2, w_1 has three pr-extensions, namely E_1, E_2 and E_3. As we shall see, in this case, for each extension there is only one explanation. In particular, $X_1 = \langle\mathtt{fish}, \mathtt{white}\rangle$ is the explanation for E_1. The intuition of explanation X_1 is that, considering that the AF consists of two strongly connected components, we first choose fish (with probability 1/2 as we can only choose between fish and meat) in the first component and determine that meat cannot belong to the extension; then we choose white (with probability 1/2 as we can only choose between white and red) in the second component, obtaining that X_1 has probability $1/2 \cdot 1/2 = 1/4$. Analogously, $X_2 = \langle\mathtt{fish}, \mathtt{red}\rangle$ is the only explanation for E_2 with probability $1/2 \cdot 1/2 = 1/4$. Considering explanation $X_3 = \langle\mathtt{meat}\rangle$ for extension E_3, we have that we first choose meat with probability 1/2 as it belongs to the first component, and we can only choose between fish and meat. Next, since we determine that fish and white cannot belong to the extension, whereas red does, the probability of X_3 turns out to be 1/2. Since the probabilities of X_1, X_2 and X_3 are 1/4, 1/4 and 1/2, respectively, the probabilities associated with E_1, E_2 and E_3 in the world w_1 are 1/4, 1/4 and 1/2, respectively. Moreover, since E_1 is not an extension of any other possible world, the probability of E_1 in Δ is $1/4 \cdot 0.48 = 0.12$. In Example 15, we will give the probabilities of the pr-extensions of every possible world of the probabilistic AF Δ of Example 1, from which it turns out that the answer to PrEA[pr] for meat is 0.70, while that for fish is 0.30. $\square$

Another argumentation framework extending AF, that has received an increasing attention in the last years and is tightly related to PrAF, is that of incomplete AF (iAF) [40]. An iAF consists of arguments and attacks, where some of them are uncertain. With respect to PrAF, in iAF we only know that some elements are uncertain. Thus, an iAF can be viewed as a

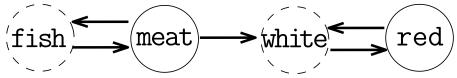

Fig. 4. iAF Δ of Example 4.

directed graph with certain and uncertain elements (from the graphical point of view, uncertain arguments and attacks will be depicted by using dashed lines).

The semantics of an iAF is given by considering all *completions*, i.e. AFs obtained by removing consistent subsets of the uncertain elements, and for each completion its extensions under a given semantics σ.

Example 4. The iAF derived from the PrAF of Example 1, by replacing the arguments with probability less than 1 with uncertain arguments, is shown in Fig. 4 (where dashed nodes represent uncertain arguments).

For this iAF, there are 4 completions that correspond to the possible worlds w_1, w_2, w_3 and w_4 of Example 2. □

Acceptance problems have been recently extended to iAF: an argument g is *possibly credulously* (resp. *skeptically*) accepted under semantics σ if there exists a completion where it is credulously (resp. skeptically) accepted under σ; an argument is *necessarily credulously* (resp. *skeptically*) accepted under semantics σ if for all completions it is credulously (resp. skeptically) accepted under σ. In the second part of the paper, we introduce the concept of (explanation-based) probabilistic acceptance for iAF, extending our proposal for PrAF to the case of iAF.

Contributions. We make the following main contributions.

- We first formally define the problem of Probabilistic Acceptance PrA[σ], for any semantics $\sigma \in \{\texttt{gr}, \texttt{co}, \texttt{pr}, \texttt{st}, \texttt{sst}\}$. Given a PrAF Δ and an argument g, the problem asks the probability that g is accepted in Δ, by means of some fixed PDF over the σ-extensions of the possible worlds of Δ.
- We introduce our notion of explanation, and exploit it to provide a PDF over the σ-extensions of an AF. This leads to an instantiation of PrA[σ], dubbed PrEA[σ].
- We investigate the complexity of PrA[σ], showing that it is $FP^{\#P}$-hard for $\sigma \in \{\texttt{gr}, \texttt{co}, \texttt{pr}, \texttt{st}, \texttt{sst}\}$ even for acyclic PrAFs and regardless of the way a PDF is defined on σ-extensions. This entails that PrEA[σ] is as hard as the problem of computing credulous and skeptical acceptance (i.e. PrCA[σ] and PrSA[σ]) [33]. Moreover, we show that PrEA[σ] remains $FP^{\#P}$-hard even for AF (PrAF where all arguments and attacks have probability 1) for $\sigma \in \{\texttt{co}, \texttt{pr}, \texttt{st}, \texttt{sst}\}$, while PrEA[gr] is polynomial (it is polynomial also if $\sigma \in \{\texttt{co}, \texttt{pr}, \texttt{st}, \texttt{sst}\}$ and the AF is acyclic). These results are summarized in Table 3.
- To deal with the intractability of PrA (and of PrEA), we propose an *additive error approximation algorithm* for PrEA[σ] that works for (*i*) PrAFs without odd-length cycles and semantics $\sigma \in \{\texttt{co}, \texttt{pr}, \texttt{st}, \texttt{sst}\}$, and (*ii*) general PrAFs (without the restriction on odd-length cycles) and grounded semantics.
- We show that our approximation results are the best that can be achieved (under standard theoretical assumptions) for $\sigma \in \{\texttt{gr}, \texttt{pr}, \texttt{st}, \texttt{sst}\}$, since (*i*) no *relative* error approximation algorithm exists for PrA[σ] (and thus for PrEA[σ]) for $\sigma \in \{\texttt{gr}, \texttt{co}, \texttt{pr}, \texttt{st}, \texttt{sst}\}$, even considering acyclic PrAFs, and (*ii*) if we admit odd-length cycles, then no *additive* error approximation algorithm exists for PrA[σ] (and thus for PrEA[σ]) with $\sigma \in \{\texttt{pr}, \texttt{st}, \texttt{sst}\}$. Table 1 summarizes the approximability results for PrEA[σ].
- We investigate the approximate complexity of PrCA and PrSA. As for the case of PrEA, no relative error approximation algorithm exists for PrCA[σ] and PrSA[σ] with $\sigma \in \{\texttt{gr}, \texttt{co}, \texttt{pr}, \texttt{st}, \texttt{sst}\}$, even considering acyclic PrAFs. The main difference with PrEA lies in the approximability via additive schemes. In particular, we can show that PrCA and PrSA are harder than PrEA in this regard, as no additive error approximation algorithm exists for PrCA[σ] and PrSA[σ] with $\sigma \in \{\texttt{pr}, \texttt{st}, \texttt{sst}\}$ even when PrAFs have no odd-length cycles.
- We also study the case where probabilities are given by intervals, instead of specific values. In this case the probabilistic acceptance problem gives as a result an interval. We show that all the complexity results still hold for PrAFs where probabilities are given by intervals.
- Finally, we study the probabilistic acceptance problem for iAFs. After presenting some equivalence transformations between iAFs, we introduce the probabilistic acceptance problem for iAFs. Notably, the complexity and approximation results obtained for PrAF carry over to iAF. Using the equivalence result between iAF, we investigate the relationships between iAF and PrAF and relate the (possible/necessary credulous/skeptical) acceptance problems in iAF to probabilistic acceptance in PrAF.

We believe that this paper is relevant for the AI community as it introduces the concept of probabilistic acceptance for AF and an approach for defining the acceptance value of an argument. Moreover, we provide a concrete solution to the probabilistic acceptance problem by introducing a way for defining the probability of extensions and then our concept of explanation for a given extension.

Plan of the paper. The rest of the paper is organized as follows. We first recall the abstract argumentation framework (Section 2) and the probabilistic argumentation framework (Section 3). In Section 4 we introduce explanations for extensions

Table 1
Approximability of PrEA[σ], depending on the semantics σ and on whether the input PrAF admits odd-length cycles. Non-existence (resp., existence) of an FP(A)RAS, i.e. a Fully Polynomial-time (Additive) Randomized Approximation Scheme, is denoted with $\times$ (resp., $\checkmark$) in the corresponding column.

	General PrAF		PrAF w/o odd cycles	
	FPRAS	FPARAS	FPRAS	FPARAS
gr	$\times$	$\checkmark$	$\times$	$\checkmark$
co	$\times$	open	$\times$	$\checkmark$
pr	$\times$	$\times$	$\times$	$\checkmark$
st	$\times$	$\times$	$\times$	$\checkmark$
sst	$\times$	$\times$	$\times$	$\checkmark$

and define the probability of explanations, extensions and arguments, under a given semantics. Then, in Section 5, we investigate the exact and approximate complexity of PrA and PrEA, and introduce algorithms for classes of PrAFs where an additive approximation algorithm for PrEA exists. We also investigate the approximate complexity of PrCA and PrSA and show that no additive error approximation algorithm exists under semantics $\sigma \in \{\mathtt{pr}, \mathtt{st}, \mathtt{sst}\}$. In Section 6 we extend the probabilistic framework by considering probabilistic intervals, and show that complexity results obtained for PrAF still hold for the extended PrAF. Then, in Section 7, we provide equivalence results between iAF and special classes of iAF, and investigate the relationships between iAF and PrAF by relating iAF acceptance problems to probabilistic acceptance in PrAF. Related work is discussed in Section 8 where conclusions are drawn and directions for future work are outlined.

This paper refines and substantially extends the work in [1]. In particular, the techniques defined in Section 4 have been revised and extended to also deal with the complete semantics. Due to the revision of the proposed technique, the results given in Section 5 are significantly different. Sections 6 and 7 are new and we provide the proofs of all results stated in the paper. To ease readability, the proofs of the results are given in the appendix.

2. Argumentation frameworks

An abstract *Argumentation Framework* (AF) is a pair $\langle A, \Sigma\rangle$, where A is a set of *arguments* and $\Sigma \subseteq A \times A$ is a set of *attacks*. An AF can be seen as a directed graph, whose nodes represent arguments and edges represent attacks. We shall use the notations a^+ and a^- for the sets $\{b \mid (a,b) \in \Sigma\}$ and $\{b \mid (b,a) \in \Sigma\}$, respectively. That is, a^+ denotes the set of arguments attacked by a, and a^- denotes the set of arguments attacking a. Further, for any $S \subseteq A$, we use S^+ and S^- to denote the sets $\bigcup_{a\in S} a^+$ and $\bigcup_{a\in S} a^-$, respectively. We use S^* to denote $S \cup S^+$, that is the set of arguments in S or attacked by (an argument in) S.

In the following, given an AF $\Lambda = \langle A, \Sigma\rangle$ and a set $S \subseteq A$ of arguments, we define $\Lambda{\downarrow_S} = \langle S, \Sigma \cap (S \times S)\rangle$ as the *restriction* of Λ to the set S.

Different argumentation semantics have been defined leading to the characterization of collectively acceptable sets of arguments, called *extensions* [13]. Given an AF $\Lambda = \langle A, \Sigma\rangle$ and a set $S \subseteq A$ of arguments, an argument $a \in A$ is said to be *i*) *defeated* w.r.t. S iff $\exists b \in S$ such that $(b,a) \in \Sigma$, and *ii*) *acceptable* w.r.t. S iff for every argument $b \in A$ with $(b,a) \in \Sigma$, there is $c \in S$ such that $(c,b) \in \Sigma$. The sets of arguments defeated and acceptable w.r.t. S are as follows (where Λ is understood):

- $Def(S) = S^+ = \{a \in A \mid \exists b \in S\,.\,(b,a) \in \Sigma\}$;
- $Acc(S) = \{a \in A \mid \forall b \in A\,.\,(b,a) \in \Sigma \Rightarrow b \in Def(S)\}$.

Given an AF $\langle A, \Sigma\rangle$, a set $S \subseteq A$ of arguments is said to be *conflict-free* iff $S \cap Def(S) = \emptyset$. Moreover, $S \subseteq A$ is said to be a *complete* extension iff it is conflict-free and $S = Acc(S)$. A complete extension S for a given AF $\langle A, \Sigma\rangle$, is said to be:

- *preferred* (pr) iff it is maximal (w.r.t. $\subseteq$);
- *stable* (st) iff it is a preferred extension such that $S \cup Def(S) = A$;
- *semi-stable* (sst) iff it is a preferred extension with a maximal set of decided elements, i.e. a preferred extension such that $S \cup Def(S)$ is maximal;
- *grounded* (gr) iff it is minimal (w.r.t. $\subseteq$).

In the following, if not specified otherwise, σ denotes any semantics in $\{\mathtt{gr}, \mathtt{co}, \mathtt{pr}, \mathtt{st}, \mathtt{sst}\}$. For any AF Λ and semantics σ, $\sigma(\Lambda)$ denotes the set of σ-extensions of Λ. Hereafter we say that an argument a is *true* (resp., *false*) w.r.t. a σ-extension E iff $a \in E$ (resp., $a \in Def(E)$). All the above-mentioned semantics except the stable admit at least one extension (i.e. $\sigma(\Lambda) \neq \emptyset$ for $\sigma \in \{\mathtt{gr}, \mathtt{co}, \mathtt{pr}, \mathtt{sst}\}$), and the grounded admits exactly one extension (i.e. $|\mathtt{gr}(\Lambda)| = 1$) [13,41]. Moreover, the grounded semantics is called *deterministic* (or *unique status*), whereas the other semantics are called *non-deterministic* (or *multiple status*). The stable semantics is said to be *total* as every argument of Λ belongs to either E or $Def(E)$

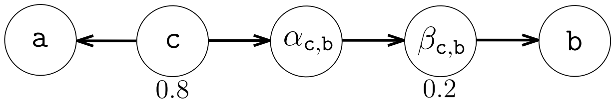

Fig. 5. PrAF equivalent to that of Fig. 1 (obtained using the translation in [43]).

(i.e. it is either *true* or *false*) for each extension $E \in \mathtt{st}(\Lambda)$. For any AF Λ, it holds that $\mathtt{st}(\Lambda) \subseteq \mathtt{sst}(\Lambda) \subseteq \mathtt{pr}(\Lambda) \subseteq \mathtt{co}(\Lambda)$, and $\mathtt{gr}(\Lambda) \subseteq \mathtt{co}(\Lambda)$ [42]. With a little abuse of notation, we also use $\mathtt{gr}(\Lambda)$ to denote the grounded extension of Λ.

Example 5. Consider the AF Λ derived from the probabilistic AF of Example 1 by assigning to all arguments probability equal to 1 (Λ coincides with the possible world w_1 of Example 2). The set of all complete extensions of Λ is $\{\emptyset, \{\mathtt{fish}\}, \{\mathtt{red}\}, \{\mathtt{fish, white}\}, \{\mathtt{fish, red}\}, \{\mathtt{meat, red}\}\}$. Thus, $\mathtt{pr}(\Lambda) = \mathtt{st}(\Lambda) = \mathtt{sst}(\Lambda) = \{\{\mathtt{fish, white}\}, \{\mathtt{meat, red}\}, \{\mathtt{fish, red}\}\}$, whereas $\mathtt{gr}(\Lambda) = \emptyset$. $\square$

For any AF $\Lambda = \langle A, \Sigma\rangle$, semantics σ, and argument $g \in A$, we say that g is *credulously* (resp. *skeptically*) accepted (under semantics σ) if g belongs to at least a σ-extension of Λ (resp. g is credulously accepted and belongs to every σ-extension of Λ).[1] We use $CA_\sigma(\Lambda, g)$ (resp. $SA_\sigma(\Lambda, g)$) to denote the fact that g is *credulously* (resp. *skeptically*) accepted or not (under semantics σ), that is $CA_\sigma(\Lambda, g)$ (resp. $SA_\sigma(\Lambda, g)$) is either true or false.

We use CA_σ (resp. SA_σ), or simply CA (resp. SA) whenever σ is understood, to denote the credulous (resp. skeptical) acceptance problem, that is, the problem of deciding whether an argument is credulously (resp. skeptically) accepted. Clearly, for the grounded semantics, which prescribes exactly one extension, the two problems are identical (i.e. $\mathsf{CA}_{\mathtt{gr}} \equiv \mathsf{SA}_{\mathtt{gr}}$).

Let $\Lambda = \langle A, \Sigma\rangle$ be an AF. A *strongly connected component (SCC)* of Λ is a maximal subset C of A such that, for every pair of arguments $a, b \in C$, there is a path from a to b along the attack relation in the graph representing Λ.[2]

Note that, differently from the standard definition, we use SCC to denote a set of nodes (i.e. arguments), not a subgraph.

An AF is *acyclic* (resp. *odd-cycle free*) if the associated graph is acyclic (resp. odd-cycle free). For acyclic AFs all the considered semantics coincide.

3. Probabilistic argumentation frameworks

In general, a probabilistic argumentation framework consists of probabilistic arguments and probabilistic attacks [23,24,26]. However, w.l.o.g. we can focus on *Probabilistic Argumentation Frameworks* (PrAFs) where only arguments are uncertain (and attacks are certain, i.e. their probability is 1), since, as shown in [43], an argumentation framework with probabilities on both arguments and attacks can be transformed into an equivalent PrAF (w.r.t. the computational tasks considered in this paper). For instance, the PrAF equivalent to that in Fig. 1 is shown in Fig. 5, where the chain of the certain attacks between $\mathtt{c}$ and $\mathtt{b}$ that passes through two new meta-arguments (one of which is probabilistic) replaces the probabilistic attack $(\mathtt{c}, \mathtt{b})$ of the PrAF of Fig. 1.

Definition 1. A Probabilistic Argumentation Framework (PrAF) is a triple $\langle A, \Sigma, P\rangle$, where $\langle A, \Sigma\rangle$ is an *AF* and P is a function assigning a probability value to every argument in A, that is, $P : A \to (0, 1]$.

Observe that assigning probability equal to 0 to arguments is useless. Basically, the value assigned by P to any argument a represents the probability that a actually occurs. Moreover, every attack (a, b) occurs with conditional probability 1, that is, a attacks b whenever both a and b occur.

The meaning of a PrAF is given in terms of *possible worlds*. Formally, given a PrAF $\Delta = \langle A, \Sigma, P\rangle$, a possible world of Δ is an AF $w = \langle A', \Sigma'\rangle$ such that $A' \subseteq A$ and $\Sigma' = \Sigma \cap (A' \times A')$. We use $pw(\Delta)$ to denote the set of all possible worlds of Δ.

An argument $a \in A$ can be viewed as a probabilistic event which is independent from the other events associated with other arguments $b \in A$ (with $b \neq a$).

The *interpretation* of a PrAF $\Delta = \langle A, \Sigma, P\rangle$ is a probability distribution function (PDF) $\mathcal{I}$ over the set $pw(\Delta)$ of the possible worlds. Each $w = \langle A', \Sigma'\rangle \in pw(\Delta)$ is assigned by $\mathcal{I}$ the probability

$$\mathcal{I}(w) = \prod_{a \in A'} P(a) \cdot \prod_{a \in A \setminus A'} (1 - P(a)). \tag{1}$$

It is worth noting that we can restrict our attention only to possible worlds w such that $\mathcal{I}(w) > 0$, as only these worlds contribute to compute probabilistic extensions and probabilistic acceptance of arguments. Therefore, in Equation (1) we can

[1] An alternative definition for skeptical acceptance provided in the literature does not require that g is also credulously accepted. The two definitions give different results only for stable semantics when no stable extensions exist. The definition used here is more consistent with the probabilistic acceptance discussed in the paper (introduced in Definition 3).

[2] If $a = b$, a path trivially exists.

only consider arguments $a \in A$ such that $0 < P(a) < 1$, as the deletion of an argument a with $P(a) = 1$ (resp. the addition to an argument a with $P(a) = 0$) results in a possible world w with $\mathcal{I}(w) = 0$. The next example shows how probabilities of possible worlds are computed.

Example 6. The (non-zero probability) possible worlds of the PrAF Δ of Example 1 are w_1, w_2, w_3 and w_4 given in Example 2. Then, interpretation $\mathcal{I}$ is as follows:
- $\mathcal{I}(w_1) = P(\mathtt{fish}) \cdot P(\mathtt{white}) = 0.6 \cdot 0.8 = 0.48$,
- $\mathcal{I}(w_2) = P(\mathtt{fish}) \cdot (1 - P(\mathtt{white})) = 0.6 \cdot 0.2 = 0.12$,
- $\mathcal{I}(w_3) = (1 - P(\mathtt{fish})) \cdot P(\mathtt{white}) = 0.4 \cdot 0.8 = 0.32$,
- $\mathcal{I}(w_4) = (1 - P(\mathtt{fish})) \cdot (1 - P(\mathtt{white})) = 0.4 \cdot 0.2 = 0.08$, and
- $\mathcal{I}(w) = 0$ for any other world $w \in pw(\Delta)$ obtained by deleting elements with probability 1. □

Given a PrAF Δ and a semantics $\sigma \in \{\mathtt{gr}, \mathtt{co}, \mathtt{pr}, \mathtt{st}, \mathtt{sst}\}$, we shall denote by $\sigma(\Delta) = \{E \mid \exists w \in pw(\Delta) \wedge \mathcal{I}(w) > 0 \wedge E \in \sigma(w)\}$ the set of σ-extensions for Δ.

As mentioned earlier, relevant problems for AF are those concerning *credulous* and *skeptical acceptance*. The analogous problems in the context of a probabilistic AF are the following.

Definition 2 *(Probabilistic credulous/skeptical acceptance).* Given a PrAF $\Delta = \langle A, \Sigma, P\rangle$, an argument $g \in A$, the probability $PrCA^{\sigma}_{\Delta}(g)$ that g is credulously acceptable w.r.t. semantics σ is

$$PrCA^{\sigma}_{\Delta}(g) = \sum_{\substack{w \in pw(\Delta) \text{ s.t.} \\ \exists E \in \sigma(w) \text{ with } g \in E}} \mathcal{I}(w). \tag{2}$$

The probability $PrSA^{\sigma}_{\Delta}(g)$ that g is skeptically acceptable w.r.t. semantics σ is

$$PrSA^{\sigma}_{\Delta}(g) = \sum_{\substack{w \in pw(\Delta) \text{ s.t.} \\ \sigma(w) \neq \emptyset \ \wedge\ \forall E \in \sigma(w),\ g \in E}} \mathcal{I}(w). \tag{3}$$

We use PrCA[σ] and PrSA[σ] (or simply PrCA and PrSA whenever σ is understood) to denote the problems of computing $PrCA^{\sigma}_{\Delta}(g)$ and $PrSA^{\sigma}_{\Delta}(g)$, respectively. Both PrCA[σ] and PrSA[σ] are $FP^{\#P}$-hard for all semantics $\sigma \in \{\mathtt{gr}, \mathtt{co}, \mathtt{pr}, \mathtt{st}, \mathtt{sst}\}$ [33].[3] Considering the PrAF concerning the robbery case discussed in the Introduction, we have that $PrCA^{\sigma}_{\Delta}(\mathtt{a}) = PrSA^{\sigma}_{\Delta}(\mathtt{a}) = 0.2$, $PrCA^{\sigma}_{\Delta}(\mathtt{b}) = PrSA^{\sigma}_{\Delta}(\mathtt{b}) = 0.84$ and $PrCA^{\sigma}_{\Delta}(\mathtt{c}) = PrSA^{\sigma}_{\Delta}(\mathtt{c}) = 0.8$, under any semantics $\sigma \in \{\mathtt{gr}, \mathtt{co}, \mathtt{pr}, \mathtt{st}, \mathtt{sst}\}$. In this case, the probabilistic credulous and skeptical acceptance coincide under any semantics because every possible world is acyclic. A more general case is considered below in Example 7.

As discussed in the introduction, probabilistic credulous and skeptical acceptance does not express the probability that a given argument is accepted as both definitions do not properly take into account the probabilistic values associated with the elements of the argumentation framework. Therefore, in this paper we study a new problem, called *Probabilistic Acceptance*, which can be intuitively stated as follows. Given a (probabilistic) framework Δ, a semantics σ, and a goal argument g, compute the probability that g is accepted. However, differently from previously proposed probabilistic measures, considering a possible world w having probability $\mathcal{I}(w)$, under the given semantics σ, every extension $E \in \sigma(w)$ has associated a probability $Pr(E, w, \sigma)$ so that $\sum_{E \in \sigma(w)} Pr(E, w, \sigma) = 1$ (the sum of the probabilities of the σ-extensions of w is equal to 1) and $Pr(E, w, \sigma) = 0$ for all $E \notin \sigma(w)$. In more detail, as stated next, we require that a PDF over the set of extensions is given.

Definition 3 *(Probabilistic acceptance).* Given a PrAF $\Delta = \langle A, \Sigma, P\rangle$ and an argument $g \in A$, the probability $PrA^{\sigma}_{\Delta}(g)$ that g is acceptable w.r.t. semantics σ is

$$PrA^{\sigma}_{\Delta}(g) = \sum_{\substack{w \in pw(\Delta) \wedge \\ E \in \sigma(w) \wedge g \in E}} \mathcal{I}(w) \cdot Pr(E, w, \sigma) \tag{4}$$

where $Pr(\cdot, w, \sigma)$ is a PDF over the set $\sigma(w)$.

It is worth noting that it makes sense to use the concept of probabilistic acceptance even for standard AF, once we have defined the PDF $Pr(\cdot, \Lambda, \sigma)$. Moreover, notice that, if for a given possible world w no extension $E \in \sigma(w)$ exists such that $g \in E$, then the contribution of the addend corresponding to that world in the summation of Equation (4) is null. In

[3] It is worth noting that the complexity result for PrSA[st] given in [33] still holds for our definition of PrSA[st], as the PrAF used in their reduction is acyclic (thus a stable extension always exists, as required by our definition, and the two variants of the problem coincide).

Table 2
Skeptical, credulous, and probabilistic acceptance for the arguments of PrAF Δ of Example 1 w.r.t. the preferred semantics (assuming a uniform PDF over the set of preferred extensions of each possible world of Δ).

Argument g	$PrSA^{\mathtt{pr}}_{\Delta}(g)$	$PrCA^{\mathtt{pr}}_{\Delta}(g)$	$PrA^{\mathtt{pr}}_{\Delta}(g)$
fish	0.00	0.60	0.38
meat	0.40	1.00	0.62
white	0.00	0.48	0.16
red	0.52	1.00	0.84

particular, if $\sigma = \mathtt{st}$ and Δ is an AF such that $\mathtt{st}(\Delta) = \emptyset$, then $PrA^{\sigma}_{\Delta}(g) = 0$ for any argument g, regardless of the chosen PDF.

Example 7. Consider the PrAF Δ of Example 1. The values of skeptical probabilistic acceptance, credulous probabilistic acceptance, and probabilistic acceptance, assuming the uniform PDF over the set of extensions, for each argument w.r.t. the preferred semantics are reported in the second, third, and fourth columns of Table 2, respectively. For instance, considering that the probabilities of the (non-zero-probability) possible worlds w_1, w_2, w_3, w_4 of Δ are 0.48, 0.12, 0.32, and 0.08 (cf. Example 6), and recalling that the $\mathtt{pr}$-extensions of w_1 are $E_1 = \{\mathtt{fish}, \mathtt{white}\}$, $E_2 = \{\mathtt{fish}, \mathtt{red}\}$, and $E_3 = \{\mathtt{meat}, \mathtt{red}\}$, the $\mathtt{pr}$-extensions of w_2 are E_2 and E_3, and w_3 and w_4 have only E_3 as their preferred extension, we have that $PrCA^{\mathtt{pr}}_{\Delta}(\mathtt{fish}) = 0.48 \cdot 2/3 + 0.12 \cdot 1/2 = 0.38$ (fish belongs to two of the three $\mathtt{pr}$-extensions of w_1, and to one of the two $\mathtt{pr}$-extensions of w_2). $\square$

Considering Example 7, we can observe that, assuming a uniform PDF over the set of extensions, we have that for any argument g, $PrA^{\mathtt{pr}}_{\Delta}(g)$ is in the interval $[PrSA^{\mathtt{pr}}_{\Delta}(g), PrCA^{\mathtt{pr}}_{\Delta}(g)]$. In some sense, for any semantics σ, $PrSA^{\sigma}_{\Delta}(g)$ and $PrCA^{\sigma}_{\Delta}(g)$ define a range for each reasonable $PrA^{\sigma}_{\Delta}(g)$, whose value depends on the specific PDF. However, the uniform distribution used in the previous example could give unintuitive results. Considering the world w_1 of our running example (cf. Example 2), one expects that the probability of having meat is $1/2$ and that it does not change if in our menu there are no drinks. However, this does not hold if a uniform PDF over the set of extensions is considered.

Our definition of probabilistic acceptance differs from the notion of (probabilistic) credulous acceptance for deterministic argumentation frameworks (i.e. AF) proposed in [44], where a PDF over the set of σ-extensions is assumed to be given. Besides considering PrAFs, we propose an approach based on explanations which entails a PDF on the set of σ-extensions of a PrAF.

4. Explanations

In this section, we show how the probability $Pr(E, \Lambda, \sigma)$ of an extension E for an AF Λ under semantics σ can be defined. Based on this, we obtain a PDF over the set of σ-extensions for every possible world w of a PrAF that will be then used to provide a concrete instantiation for the probabilistic acceptance problem. The idea is to assign, for every possible world w, a probability to every extension in $\sigma(w)$, according to a sequence of choices (called explanation) made to compute it. These choices are made by following the topological order of arguments in the graph representing the AF. Once we have assigned a probability to possible worlds and extensions of possible worlds, the probabilistic acceptance of an argument a is defined as the sum of the products of the probabilities of the possible worlds w and extensions of w containing a. One of the innovative aspects of this work is how probabilities are assigned to extensions by exploiting the concept of explanation. Compared to the uniform distribution, the proposed PDF assigns probabilities to extensions and arguments (i.e. probabilities of being accepted) that appear to be more intuitive. For instance, considering the AF obtained from the deterministic version of the PrAF in Fig. 1 (i.e. world w_1 shown in Fig. 3 (top-left), and reported for the sake of readability in Fig. 6 (left)), the probability of accepting one of the two mutually exclusive arguments fish or meat (under stable/preferred/semi-stable semantics) is the same and equal to $1/2$. In contrast, under the uniform distribution, the probability of having fish (resp. meat) is $2/3$ (resp. $1/3$). Moreover, as stated in Theorem 6, the proposed PDF based on explanations allows us to obtain a tractable sampling strategy for a large class of PrAFs.

To define $Pr(E, \Lambda, \sigma)$ we introduce the concept of *explanation* consisting of a sequence of necessary suggestions useful to construct a given extension E, that is a sequence of choices made to obtain the extension. In particular, the choices we consider are guided by an ordering entailed by the strongly connected components (SCCs) of the given AF.

We start by introducing some notations. Given an AF $\Lambda = \langle A, \Sigma \rangle$, a subset $\Omega \subseteq A$ of arguments, and an argument $a \in A \setminus \Omega$, let $G = \mathtt{gr}(\Lambda)$ be the grounded extension of Λ, we denote by:

- Λ^*, the restriction of Λ to the set $A \setminus G^*$, that is, Λ^* is the AF $\Lambda\downarrow_{A \setminus G^*}$ obtained from Λ by removing the arguments in G or attacked by G (i.e. G^+) as well as the attacks involving these arguments;
- Λ_a, the restriction of Λ obtained through the deletion of attacks whose target is a;
- $\Lambda^*_a = (\Lambda_a)^*$, the restriction of Λ_a obtained by deleting, letting $G_a = \mathtt{gr}(\Lambda_a)$, all arguments (and related attacks) in G^*_a;

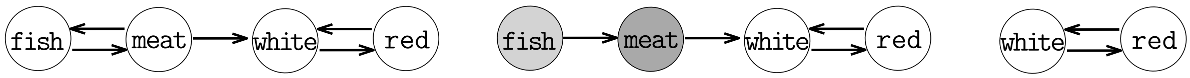

Fig. 6. AF Λ of Example 5 (left) and AFs $\Lambda_{\mathtt{fish}}$ (center) and $\Lambda^*_{\mathtt{fish}}$ (right) of Example 8, where light (resp. dark) grey-colored arguments are those that appear as true (resp. false) in $G_{\mathtt{fish}} = \mathtt{gr}(\Lambda_{\mathtt{fish}})$.

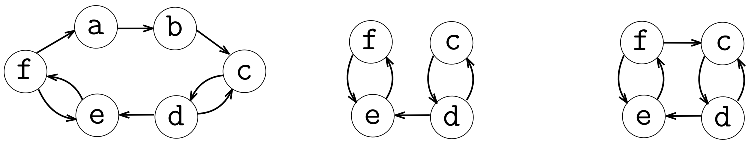

Fig. 7. AFs Λ (left) $\Lambda^*_{\mathtt{a}}$ (center) and $\widehat{\Lambda_{\mathtt{a}}}$ (right) of Example 9.

- $\widehat{\Lambda_a}$ the AF obtained from Λ^*_a by adding, letting $G_a = \mathtt{gr}(\Lambda_a)$ and a^- be the set of attackers of a w.r.t. Λ, the *(virtual) attacks* in $(a^- \setminus G^+_a) \times ((G^+_a)^+ \setminus G^*_a)$, where $G^+_a = (G_a)^+$;
- $\Omega_{\Lambda,a} = (\Omega \cup a^-) \setminus G^+_a$ (also denoted as Ω_a whenever Λ is understood).

Intuitively, Ω denotes a set of arguments whose status has not yet been determined but, based on previous assumptions, will have to be derived as defeated; a is an argument we are assuming to be accepted; and $G_a = \mathtt{gr}(\Lambda)$ (resp. G^+_a) is the set of arguments that are derived to be accepted (resp. defeated) as a consequence of the assumption about a. Moreover, Ω_a and Λ^*_a are the updated set of arguments to be derived as defeated and the updated AF (obtained by removing the arguments in G^*_a whose status has been determined), after assuming a is accepted. Indeed, assuming that a is accepted, all arguments in a^- (i.e. those attacking a) should be defeated and added to Ω (yielding Ω_a), though the arguments in G^*_a, whose status has been determined, are deleted from Ω_a. Moreover, Λ^*_a is obtained from Λ_a by deleting the arguments in $G_a = \mathtt{gr}(\Lambda_a)$ and G^+_a whose status (accepted or defeated) has been determined. Finally, virtual attacks are added to Λ^*_a, obtaining the updated AF $\widehat{\Lambda_a}$, to ensure that if a belonged to a SCC C, after eliminating the arguments in G^*_a and the related attacks, the remaining arguments in $C \setminus G^*_a$ continue to form a SCC in $\widehat{\Lambda_a}$. Regarding virtual attacks, it is worth noting that, since the source arguments must be derived as defeated, they have no real effect on the status of target arguments, but ensure that if there was a SCC C in Λ with $a \in C$, the remaining arguments in C continue to form a SCC of $\widehat{\Lambda_a}$.

The following examples illustrate the concepts introduced above.

Example 8. Consider the AF $\Lambda = \langle A, \Sigma\rangle$ of Example 5 (that is, possible world w_1 of Example 2) shown in Fig. 6 (left). The AF $\Lambda_{\mathtt{fish}}$, obtained from Λ by removing attack $(\mathtt{meat}, \mathtt{fish})$, is shown in Fig. 6 (center). In the figure, the nodes representing the arguments in the grounded extension $G_{\mathtt{fish}} = \mathtt{gr}(\Lambda_{\mathtt{fish}}) = \{\mathtt{fish}\}$ are colored in light-grey, while those in $G^*_{\mathtt{fish}} = G_{\mathtt{fish}} \cup \{\mathtt{meat}\}$ are colored in light or dark grey. Then, Fig. 6 (right) shows the AF $\widehat{\Lambda_{\mathtt{fish}}} = \Lambda^*_{\mathtt{fish}} = \Lambda{\downarrow}_{A \setminus G^*_{\mathtt{fish}}} = \langle\{\mathtt{white}, \mathtt{red}\}, \{(\mathtt{white}, \mathtt{red}), (\mathtt{red}, \mathtt{white})\}\rangle$.

As another example, by considering argument meat, we would have $G_{\mathtt{meat}} = \mathtt{gr}(\Lambda_{\mathtt{meat}}) = \{\mathtt{meat}, \mathtt{red}\}$, $G^*_{\mathtt{meat}} = \{\mathtt{meat}, \mathtt{red}, \mathtt{fish}, \mathtt{white}\}$ and, thus, $\widehat{\Lambda_{\mathtt{meat}}} = \Lambda^*_{\mathtt{meat}} = \langle\emptyset, \emptyset\rangle$. □

Example 9. Consider the AF $\Lambda = \langle\{\mathtt{a}, \mathtt{b}, \mathtt{c}, \mathtt{d}, \mathtt{e}, \mathtt{f}\}, \{(\mathtt{a}, \mathtt{b}), (\mathtt{b}, \mathtt{c}), (\mathtt{c}, \mathtt{d}), (\mathtt{d}, \mathtt{c}), (\mathtt{d}, \mathtt{e}), (\mathtt{e}, \mathtt{f}), (\mathtt{f}, \mathtt{e}), (\mathtt{f}, \mathtt{a})\}\rangle$ shown in Fig. 7 (left). Since the grounded extension of Λ is empty (i.e. $\mathtt{gr}(\Lambda) = \emptyset$), we have that $\Lambda^* = \Lambda$. Taking argument a, we have that $\mathtt{a}^- = \{\mathtt{f}\}$, $G_{\mathtt{a}} = \mathtt{gr}(\Lambda_{\mathtt{a}}) = \{\mathtt{a}\}$, $G^+_{\mathtt{a}} = \{\mathtt{b}\}$, $G^*_{\mathtt{a}} = \{\mathtt{a}, \mathtt{b}\}$ and $(G^+_{\mathtt{a}})^+ = \{\mathtt{c}\}$. Consequently, $\Lambda^*_{\mathtt{a}}$ (shown in Fig. 7 (center)) is obtained by deleting arguments a and b (and related edges), whereas $\widehat{\Lambda_{\mathtt{a}}}$ (shown in Fig. 7 (right)) is derived from $\Lambda^*_{\mathtt{a}}$ by adding attack $(\mathtt{f}, \mathtt{c})$. □

The idea underlying the derivation of $\Lambda^*_{\mathtt{a}}$ is that if we assume that argument a is accepted, then we can determine the status of other arguments (those in G^*_a) which can be removed from the AF Λ. Moreover, to preserve the topology of the AF we add attacks from arguments in $a^- \setminus G^*_a$ to arguments in $(G^+_a)^+ \setminus G^*_a$ (recall that arguments in G^*_a are not in Λ^*_a anymore). For instance, considering the AF of Example 9 whose graph has only one SCC, after the addition of edge $(\mathtt{f}, \mathtt{c})$, the resulting graph still consists of a single SCC. It is worth noting that, in the derivation of $\widehat{\Lambda_a}$, arguments in a^- must be defeated and arguments in $(G^+_a)^+$ are attacked by arguments which have been derived as defeated.

The following proposition states that, for any extension $E \in \sigma(\Lambda)$ and argument $a \in E$, let G_a be the grounded extension of Λ_a, $E \setminus G_a$ is a σ-extension for the AF Λ^*_a. We will rely on this result later in the definition of explanation.

Proposition 1. *Let $\Lambda = \langle A, \Sigma\rangle$ be an AF, $\sigma \in \{\mathtt{co}, \mathtt{gr}, \mathtt{pr}, \mathtt{st}, \mathtt{sst}\}$ a semantics, $E \in \sigma(\Lambda)$ an extension, $a \in E$ an argument and $G_a = \mathtt{gr}(\Lambda_a)$, Then, $E' = E \setminus G_a$ is a σ-extension of Λ^*_a.*

For instance, considering the AF Λ of Example 8 and the extension $E_1 = \{\mathtt{fish}, \mathtt{white}\} \in \mathtt{pr}(\Lambda)$, we have that the set $E_1 \setminus G_{\mathtt{fish}} = \{\mathtt{white}\}$ is a preferred extension for $\widehat{\Lambda_{\mathtt{fish}}}$.

In order to define explanations, we assume that SCCs are linearly ordered according to the topological ordering of the graph representing the AF. We use the notation $C \prec C'$ to denote that SCC C precedes SCC C' (w.r.t. the fixed linear order). The linear order on the SCCs induces a partial order on the arguments, that is $a \prec b$ only if $a \in C$, $b \in C'$ and $C \prec C'$. As it will be clear in the following, fixing a linear ordering on the SCCs has no impact on the probabilities assigned to extensions by PDF $Pr(E, \Lambda, \sigma)$. Given a set of SCCs $\mathcal{C} = \{\mathcal{C}_1, ..., \mathcal{C}_n\}$, we say that a SCC $\mathcal{C}_i \in \mathcal{C}$ is the *first* SCC (in $\mathcal{C}$) if every argument in $\mathcal{C}_i$ precedes every argument in $\mathcal{C}_j$ for all $j \neq i$.

Example 10. Consider the AF $\Lambda = \langle\{\mathtt{a}, \mathtt{b}, \mathtt{c}, \mathtt{d}\}, \{(\mathtt{a}, \mathtt{b}), (\mathtt{b}, \mathtt{a}), (\mathtt{a}, \mathtt{d}), (\mathtt{c}, \mathtt{d})\}\rangle$. There are tree SCCs $C_1 = \{\mathtt{a}, \mathtt{b}\}$, $C_2 = \{\mathtt{c}\}$ and $C_3 = \{\mathtt{d}\}$. Assuming the linear ordering $C_1 \prec C_2 \prec C_3$, the partial order on arguments $\mathtt{a} \prec \mathtt{c}$, $\mathtt{b} \prec \mathtt{c}$, $\mathtt{c} \prec \mathtt{d}$ is induced. Alternatively, if we had assumed the linear ordering $C_2 \prec C_1 \prec C_3$ then we would have had the partial order $\mathtt{c} \prec \mathtt{a}$, $\mathtt{c} \prec \mathtt{b}$, $\mathtt{a} \prec \mathtt{d}$, $\mathtt{b} \prec \mathtt{d}$. □

The next definition introduces the concept of explanation for an extension E w.r.t. an AF Λ, an underlying semantics σ and a set Ω of arguments whose status is assumed to be false, i.e. it is assumed that $\Omega \subseteq Def(E)$. The set Ω is said to be a *set of assumptions*. Initially Ω is empty and it must be empty at the end of the process determining an explanation, meaning that we should end up with no assumptions on the status of arguments (as their status must be determined).

An explanation $X = \langle a_1, ..., a_n\rangle$ is a sequence of arguments occurring in E which are assumed to be accepted (in particular, an argument at each step is considered). In contrast, at each step, Ω consists of a set of arguments that, on the basis of the choices previously made, must be derived as defeated in the next steps, as they attack arguments in X (whose status has been assumed to be true); once the status of some of these arguments is determined, they are removed from Ω. Intuitively, Ω is used for guiding next choices so that assumptions previously made are not retracted.

The explanation is determined recursively for (i) an AF Λ^*, which is derived from Λ by discarding all arguments in the grounded extension (which are always accepted) and those attacked by these arguments (which are always defeated), (ii) an extension E and (iii) an (initially empty) set of arguments Ω, whose status must be determined as false. It is also worth recalling that, for any AF Λ and extension $E \in \mathtt{co}(\Lambda) \setminus \mathtt{gr}(\Lambda)$ there must be some argument $a \in E$ occurring in an even cycle of Λ [45].

Let ε be a fresh symbol, not used for arguments names, we define an explanation for an extension as follows.

Definition 4 *(Explanation).* Let $\Lambda = \langle A, \Sigma\rangle$ be an AF, $\sigma \in \{\mathtt{co}, \mathtt{gr}, \mathtt{pr}, \mathtt{st}, \mathtt{sst}\}$ a semantics, $E \in \sigma(\Lambda)$, and $\Omega \subset A$ a set of assumptions about the (false) status of some arguments. A sequence $X = \langle a_1, \ldots, a_n\rangle$, where $a_i \in E \cup \{\varepsilon\}$ (with $i \in [1...n]$), is an explanation for E (w.r.t. Λ and Ω) if, letting $\Lambda = \Lambda^*$ and C be the first SCC of Λ, one of the following three conditions hold:

1. (Final Step) $X = \langle\rangle$, $\Omega = \emptyset$, and $\Lambda = \langle\emptyset, \emptyset\rangle$.
2. (Choice Step) (i) a_1 occurs in an even cycle of C, (ii) $\Omega \neq \emptyset \Rightarrow G^+_{a_1} \cap \Omega \neq \emptyset$, and (iii) $\langle a_2, ..., a_n\rangle$ is an explanation for $E' = E \setminus G_{a_1}$ w.r.t. $\widehat{\Lambda_{a_1}}$ and Ω_{a_1}.
3. (Skip Step) $a_1 = \varepsilon$, $\Omega = \emptyset$ and $\langle a_2, \ldots, a_n\rangle$ is an explanation for E w.r.t. $\Lambda{\downarrow_{A\setminus C}}$ and Ω.

For any explanation X for E, letting $set(X) = \{x \mid x \text{ occurs in } X\}$, $set(X) \cap A$ be an *explanation-set* for E.

Therefore, an explanation $X = \langle a_1, ..., a_n\rangle$ for an extension $E \in \sigma(\Lambda)$, w.r.t. a set of (false) assumptions Ω, is (recursively) determined as follows:

1. If $\Omega \neq \emptyset$, an argument a_1 in E occurring in an even cycle is chosen (Item 2.(i)) such that, after removing all attacks targeting a, at least one argument in Ω must be derived as defeated (Item 2.(ii)). Moreover, if $\Omega = \emptyset$, either an argument in E occurring in an even cycle is chosen (Item 2) or the whole SCC C is skipped, assigning (implicitly) to all its elements the undecided truth value (Item 3).
2. After that: (a) compute Ω_{a_1} from Ω by adding arguments in a_1^- and then deleting arguments whose status is false in the grounded extension of Λ_{a_1} (i.e. $G^+_{a_1}$), (b) compute $\widehat{\Lambda_{a_1}}$ from Λ by deleting all arguments in $G^*_{a_1}$, whose status has been determined as accepted or defeated w.r.t. $\mathtt{gr}(\Lambda_{a_1})$, and then adding virtual attacks from arguments in $(a_1^- \setminus G^+_{a_1})$ to arguments in $((G^+_{a_1})^+ \setminus G^*_{a_1})$.

As said before, added virtual attacks do not play any role in determining the status of arguments in the resulting AF.

Observe that: (i) when the computation of a SCC terminates, i.e. the status of all arguments in it has been defined, Ω becomes empty, (ii) under stable semantics Item 3 of Definition 4 is never applied, and (iii) under preferred semantics Item 3 is applied only if Item 2 cannot be applied.

Intuitively, an explanation-set $set(X)$ for an extension E (derived from an explanation X) consists of a set of arguments which, if assumed to be true, allow us to derive all arguments in E. An explanation defines also an order on the arguments. Clearly, an explanation-set may be derived from more than one explanation (see also Example 12).

Example 11. Continuing with the AF Λ of Examples 5 and 8 (reported in Fig. 6 (left)), recall that it has six complete extensions: $E_0 = \emptyset$, $E_1 = \{\mathtt{fish}, \mathtt{white}\}$, $E_2 = \{\mathtt{fish}, \mathtt{red}\}$, $E_3 = \{\mathtt{meat}, \mathtt{red}\}$, $E_4 = \{\mathtt{red}\}$ and $E_5 = \{\mathtt{fish}\}$. E_0 is the grounded extension, whereas E_1, E_2 and E_3 are preferred, stable and semi-stable extensions (that we also considered in Example 3).

For the preferred extension E_1 there is only one explanation $X_1 = \langle \mathtt{fish}, \mathtt{white}\rangle$ obtained as described in what follows, where superscripts are used to distinguish the different steps. At the beginning we have an empty set of assumptions (i.e. $\Omega^0 = \emptyset$) and, as the grounded extension of Λ is empty, we have that $\Lambda = \Lambda^*$ and, therefore, our initial AF is $\Lambda^0 = \Lambda$, whereas our initial extension is denoted by $E_1^0 = E_1$. Then, the following steps are performed.

1. Argument $\mathtt{fish}$ is chosen in the first SCC of Λ^0. By removing the attacks toward $\mathtt{fish}$ we get the AF $\Lambda^0_{\mathtt{fish}}$. Then, we compute the grounded extension of $\Lambda^0_{\mathtt{fish}}$ which is $G_{\mathtt{fish}} = \{\mathtt{fish}\}$ and remove arguments in $G_{\mathtt{fish}}$ from E_1^0, obtaining $E_1^1 = \{\mathtt{white}\}$, and arguments in $G^*_{\mathtt{fish}}$ from $\Lambda^0_{\mathtt{fish}}$, obtaining $\Lambda^{0\,*}_{\mathtt{fish}} = \langle\{\mathtt{white}, \mathtt{red}\}, \{(\mathtt{white}, \mathtt{red}), (\mathtt{red}, \mathtt{white})\}\rangle$. As $\mathtt{fish}^- \setminus G^*_{\mathtt{fish}} = \emptyset$, we have that $\Omega^0_{\mathtt{fish}} = \emptyset$ and $\widehat{\Lambda^0_{\mathtt{fish}}} = \Lambda^{0*}_{\mathtt{fish}}$.
2. Let $\Lambda^1 = \widehat{\Lambda^0_{\mathtt{fish}}}$ and $\Omega^1 = \Omega^0_{\mathtt{fish}}$, the first SCC of Λ^1 is $C = \{\mathtt{white}, \mathtt{red}\}$. Therefore, the only argument that can be chosen is $\mathtt{white}$, obtaining the extension $E_1^2 = \emptyset$. After removing attacks to $\mathtt{white}$, the grounded extension of the resulting AF $\Lambda^1_{\mathtt{white}}$ is $G_{\mathtt{white}} = \{\mathtt{white}\}$. By further removing arguments in $G^*_{\mathtt{white}}$, we get the AF $\Lambda^{1*}_{\mathtt{white}} = \langle\emptyset, \emptyset\rangle$, $\Omega^1_{\mathtt{white}} = \emptyset$ and $\widehat{\Lambda^1_{\mathtt{white}}} = \Lambda^{1*}_{\mathtt{white}}$.
3. Let $\Lambda^2 = \widehat{\Lambda^1_{\mathtt{white}}}$ and $\Omega^2 = \Omega^1_{\mathtt{white}}$, as $E_2^2 = \emptyset$, $\Lambda^2 = \langle\emptyset, \emptyset\rangle$ and $\Omega^2 = \emptyset$, the process terminates.

Consider now explanation $X_3 = \langle \mathtt{meat}\rangle$ for the preferred extension $E_3 = \{\mathtt{meat}, \mathtt{red}\}$. As for the previous case, initially we have $\Omega^0 = \emptyset$, $\Lambda^0 = \Lambda^* = \Lambda$ and $E_3^0 = E_3$. Then, we perform the following steps.

1. Argument $\mathtt{meat}$ is chosen in the first SCC of Λ^0 obtaining $\Lambda^0_{\mathtt{meat}}$ by deleting attacks to $\mathtt{meat}$. After having computed the grounded extension $G_{\mathtt{meat}} = \{\mathtt{meat}, \mathtt{red}\}$ of $\Lambda^0_{\mathtt{meat}}$, and having deleted arguments in $G_{\mathtt{meat}}$ from E_3 and arguments in $G^*_{\mathtt{meat}}$ from Λ^0, we get $E_3^1 = \emptyset$, $\Omega^0_{\mathtt{meat}} = \emptyset$ and the AF $\widehat{\Lambda^0_{\mathtt{meat}}} = \Lambda^{0*}_{\mathtt{meat}} = \langle\emptyset, \emptyset\rangle$.
2. Let $\Lambda^1 = \widehat{\Lambda^0_{\mathtt{meat}}}$ and $\Omega^1 = \Omega^0_{\mathtt{meat}}$, as $E_3^1 = \emptyset$, $\Lambda^1 = \langle\emptyset, \emptyset\rangle$ and $\Omega^1 = \emptyset$, the process terminates.

Regarding the complete extension $E_4 = \{\mathtt{red}\}$, there is only one explanation $\langle \varepsilon, \mathtt{red}\rangle$ obtained as follows, after setting $\Omega^0 = \emptyset$, $\Lambda^0 = \Lambda$ and $E_4^0 = E_4$.

1. At the first step, ε is chosen and we have that $E_4^1 = \{\mathtt{red}\}$, $\Omega^1 = \Omega^0 = \emptyset$ and $\Lambda^1 = \Lambda^0\downarrow_{A\setminus C}$ is the AF obtained by deleting the first C of Λ^0.
2. At the second step, the only possible choice is $\mathtt{red}$ and we get $E_4^2 = \emptyset$, $\Omega^2 = \Omega^1_{\mathtt{red}} = \emptyset$ and $\Lambda^2 = \widehat{\Lambda^1_{\mathtt{red}}} = \langle\emptyset, \emptyset\rangle$.
3. Finally, the process terminates with the explanation $\langle \varepsilon, \mathtt{red}\rangle$.

Considering the grounded extension E_0, its explanation is $\langle \varepsilon, \varepsilon\rangle$, whereas the corresponding explanation-set is, obviously, $\emptyset$. $\square$

In the previous example, the set Ω of assumptions remained empty for all explanations considered. The next example shows the role of the set of assumptions for deriving explanations.

Example 12. Consider the AF Λ of Example 9 shown in Fig. 7 (left) and having $\mathtt{pr}(\Lambda) = \{E_1 = \{\mathtt{a}, \mathtt{c}, \mathtt{e}\}, E_2 = \{\mathtt{b}, \mathtt{d}, \mathtt{f}\}\}$. As shown next, according to Definition 4, $\langle \mathtt{a}, \mathtt{e}, \mathtt{c}\rangle$ is an explanation for E_1, whereas $\langle \mathtt{a}, \mathtt{c}, \mathtt{e}\rangle$ is not an explanation for E_1. Initially we have $\Omega^0 = \emptyset$, $\Lambda^0 = \Lambda$, as the grounded extension of Λ is empty, and $E_1^0 = E_1$. Then we have the following steps:

1. At the first step, we may choose $\mathtt{a}$. Thus, we derive $G_{\mathtt{a}} = \mathtt{gr}(\Lambda^0_{\mathtt{a}}) = \{\mathtt{a}\}$, $G^*_{\mathtt{a}} = \{\mathtt{a}, \mathtt{b}\}$, $E_1^1 = E_1^0 \setminus G_{\mathtt{a}} = \{\mathtt{c}, \mathtt{e}\}$, $\Omega_{\mathtt{a}} = \mathtt{a}^- = \{\mathtt{f}\}$, $\Lambda^{0*}_{\mathtt{a}} = \langle A^1 = \{\mathtt{c}, \mathtt{d}, \mathtt{e}, \mathtt{f}\}, \Sigma^1 = \{(\mathtt{c}, \mathtt{d}), (\mathtt{d}, \mathtt{c}), (\mathtt{e}, \mathtt{f}), (\mathtt{f}, \mathtt{e}), (\mathtt{d}, \mathtt{e})\}$ (see Fig. 7 (center)), and $\widehat{\Lambda^0_{\mathtt{a}}}$ is derived from $\Lambda^{0*}_{\mathtt{a}}$ by adding the virtual edge $(\mathtt{f}, \mathtt{c})$ (see Fig. 7 (right)) so that the resulting component continues to be strongly connected.
2. At the second step, let $\Omega^1 = \Omega^0_{\mathtt{a}}$ and $\Lambda_1 = \widehat{\Lambda^0_{\mathtt{a}}}$, as $\Omega^1 \neq \emptyset$, we have to choose an argument $\mathtt{x} \in E_1^1$ such that, $(\mathtt{gr}(\Lambda^1_{\mathtt{x}}))^+ \cap \Omega^1 \neq \emptyset$. Therefore, the only feasible choice is $\mathtt{e}$ and, as $\mathtt{gr}(\Lambda^1_{\mathtt{e}}) = \{\mathtt{c}\}$, we get $E_1^2 = E_1^1 \setminus \{\mathtt{c}\}$, $\Omega^1_{\mathtt{e}} = \{\mathtt{d}\}$ and $\Lambda^{1*}_{\mathtt{e}} = \widehat{\Lambda^1_{\mathtt{e}}} = \langle\{\mathtt{c}, \mathtt{d}\}, \{(\mathtt{c}, \mathtt{d}), (\mathtt{d}, \mathtt{c})\}\rangle$.
3. At the third step, let $\Omega^2 = \Omega^1_{\mathtt{e}}$ and $\Lambda_2 = \widehat{\Lambda^1_{\mathtt{e}}}$, since $\Omega^2 \neq \emptyset$, the only feasible choice in E_1^2 is $\mathtt{c}$ and we get $E_1^3 = \emptyset$, $\Omega^2_{\mathtt{c}} = \emptyset$ and $\Lambda^{2*} = \widehat{\Lambda^2_{\mathtt{c}}} = \langle\emptyset, \emptyset\rangle$.
4. At the fourth step, let $\Omega^3 = \Omega^2_{\mathtt{c}}$ and $\Lambda_3 = \widehat{\Lambda^2_{\mathtt{c}}}$, as $\Omega_3 = E_1^3 = \emptyset$ and $\Lambda^3 = \langle\emptyset, \emptyset\rangle$, the process terminates giving explanation $\langle \mathtt{a}, \mathtt{e}, \mathtt{c}\rangle$.

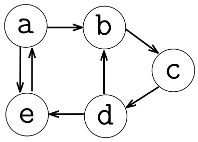

Fig. 8. AF Λ of Example 13.

For the extension E_1, there are three explanations $\langle \mathtt{a}, \mathtt{e}, \mathtt{c} \rangle$, $\langle \mathtt{c}, \mathtt{a}, \mathtt{e} \rangle$, $\langle \mathtt{e}, \mathtt{c} \rangle$ and only two explanation-set $\{\mathtt{a}, \mathtt{c}, \mathtt{e}\}$ and $\{\mathtt{e}, \mathtt{c}\}$. $\square$

The reason for discarding the sequence $\langle \mathtt{a}, \mathtt{c}, \mathtt{e} \rangle$ will be clearer in the next section where we show how to compute both explanations and extensions together and, therefore, explanations will be not "guided" by extensions. Recall also that explanations are mainly used for assigning probabilities to extensions.

The following example shows that by restricting the choice of arguments to those occurring in even cycles we reduce the set of explanations. This restriction does not impact on the computation of the probabilities assigned to extensions.

Example 13. Consider the AF $\Lambda = \langle \{\mathtt{a}, \mathtt{b}, \mathtt{c}, \mathtt{d}, \mathtt{e}\}, \{(\mathtt{a}, \mathtt{b}), (\mathtt{b}, \mathtt{c}), (\mathtt{c}, \mathtt{d}), (\mathtt{d}, \mathtt{e}), (\mathtt{e}, \mathtt{a}), (\mathtt{a}, \mathtt{e}), (\mathtt{d}, \mathtt{b})\} \rangle$ shown in Fig. 8. Λ has two complete extensions $E_0 = \emptyset$, $E_1 = \{\mathtt{a}, \mathtt{c}\}$. E_0 is the grounded extension, whereas E_1 is a preferred (stable, and semi-stable) extension. According to Definition 8, E_1 has only one explanation $X_1 = \langle \mathtt{a} \rangle$. However, if we had removed the restriction in Item 2 that a_1 must occur in an even cycle, both sequences $X_1 = \langle \mathtt{a} \rangle$ and $X_2 = \langle \mathtt{c}, \mathtt{a} \rangle$ would have been explanations for E_1. $\square$

In the following, the set of explanations for a σ-extension E of an AF Λ is denoted by $Exp^{\sigma}_{\Lambda}(E)$. Moreover, $Exp^{\sigma}(\Lambda) = \bigcup_{E \in \sigma(\Lambda)} Exp^{\sigma}_{\Lambda}(E)$ is the set of explanations of Λ under semantics σ.

Proposition 2. *Let Λ be an AF and $\sigma \in \{\mathtt{gr}, \mathtt{co}, \mathtt{pr}, \mathtt{st}, \mathtt{sst}\}$ a semantics. Then:*

i) for every $E \in \sigma(\Lambda)$, $Exp^{\sigma}_{\Lambda}(E) \neq \emptyset$;
ii) for every $E_i, E_j \in \sigma(\Lambda)$ with $E_i \neq E_j$, $Exp^{\sigma}_{\Lambda}(E_i) \cap Exp^{\sigma}_{\Lambda}(E_j) = \emptyset$.

The previous proposition states that the relationship between explanations and extensions is a total and surjective function from the set of explanations $Exp^{\sigma}(\Lambda)$ to the set of extensions $\sigma(\Lambda)$. Thus, under a given semantics σ, every explanation identifies a unique extension. In the following, for any explanation X for an extension E under semantics σ, $ext^{\sigma}(X)$ denotes the extension associated to X.

The next theorem states that any explanation defines a set of assumptions to be made for computing the extension E it explains, in the sense that having some arguments in the explanation means that some other arguments must be assumed to be false (that is, in E^{+}).

Theorem 1. *Let $\Lambda = \langle A, \Sigma \rangle$ be an AF, σ a semantics in $\{\mathtt{gr}, \mathtt{co}, \mathtt{pr}, \mathtt{st}, \mathtt{sst}\}$ and E a σ-extension. Then, for any $X \in Exp^{\sigma}_{\Lambda}(E)$ and $\widetilde{X} = set(X) \cap A$ we have that $E = \mathtt{gr}(\Lambda_{\widetilde{X}})$ and $\widetilde{X}^{-} \subseteq E^{+}$, where $\Lambda_{\widetilde{X}}$ is the AF derived from Λ by deleting attacks to arguments in $\widetilde{X}$.*

Since a given extension may have multiple explanations of different length, it is reasonable to assume that some explanations are preferred to others. We now introduce probabilities for explanations. As said before, the grounded semantics has a unique explanation which has probability 1. To define probabilities of explanations, we exploit the concept of probabilistic trie.

Definition 5. Given an AF $\Lambda = \langle A, \Sigma \rangle$ and a semantics σ, the probabilistic trie for Λ under semantics σ is the triple $\mathcal{T}^{\sigma}_{\Lambda} = \langle N, H, \pi \rangle$ of nodes N and edges H where $\langle N, H \rangle$ is the trie of all sequences in $Exp^{\sigma}(\Lambda)$, $\pi : N \to (0, 1]$ is the function inductively defined as follows:

$$\pi(\langle\rangle) = 1 \qquad \text{and} \qquad \pi(x) = \frac{\pi(parent(x))}{|children(parent(x))|}$$

where $parent(x)$ denotes the parent of x, whereas $|children(x)|$ denotes the number of children of x.

Since the set of leaves of the probabilistic trie $\mathcal{T}^{\sigma}_{\Lambda} = \langle N, H, \pi \rangle$ coincides with $Exp^{\sigma}(\Lambda)$ (i.e. $leaves(\mathcal{T}^{\sigma}_{\Lambda}) = Exp^{\sigma}(\Lambda)$) hereafter, with a little abuse of notation, we assume that π is a function from $Exp^{\sigma}(\Lambda)$ to $(0, 1]$. By definition, we have that $\sum_{X \in Exp^{\sigma}(\Lambda)} \pi(X) = 1$.

As defined next, the probability value associated with a σ-extension E of an AF Λ is given by the sum of the probabilities of the explanations for E under semantics σ.

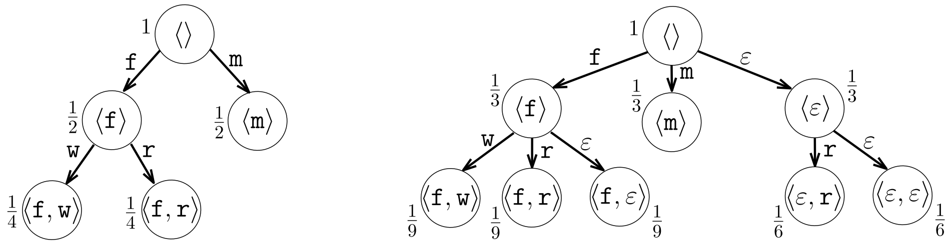

Fig. 9. Probabilistic trie for the AF Λ of Example 5 under preferred/stable/semi-stable semantics (left), and complete semantics (right).

Definition 6. Given an AF $\Lambda = \langle A, \Sigma\rangle$, a semantics σ, and an extension $E \in \sigma(\Lambda)$, the *Explanation-based Probability* associated with extension E is:

$$PrE(E, \Lambda, \sigma) = \sum_{X \in Exp^{\sigma}_{\Lambda}(E)} \pi(X). \tag{5}$$

The intuition behind the previous definition is that, as extensions do not share explanations, the probability of the explanations is carried over to the extensions. As a consequence we have that $\sum_{E\in\sigma(\Lambda)} PrE(E, \Lambda, \sigma) = 1$.

With a little effort, it can be checked that function $PrE(\cdot, \Lambda, \sigma)$ is a PDF over the set of σ-extensions of Λ. The following example illustrates how probabilities are associated to extensions of an AF.

Example 14. Let Λ be the AF of Example 5. The explanations for the preferred (stable and semi-stable) extensions are represented by the leaf nodes of the trie shown in Fig. 9 (left), where arguments are denoted by their initials. For instance, the probability of (the unique) explanation $\langle \mathtt{fish}, \mathtt{white}\rangle$ for extension $E_1 = \{\mathtt{fish}, \mathtt{white}\}$ (cf. Example 11) is $\mathtt{1/4}$, whereas the probability of (the unique) explanation $\langle \mathtt{meat}\rangle$ for extension $E_3 = \{\mathtt{meat}, \mathtt{red}\}$ is $\mathtt{1/2}$. Therefore under preferred (stable and semi-stable) semantics, the probability of E_1 is $\mathtt{1/4}$, whereas that of E_3 is $\mathtt{1/2}$.

Considering the complete semantics, there are six extensions whose explanations are represented by the leaf nodes of the trie shown in Fig. 9 (right). All explanations containing $\mathtt{fish}$ have probability $\mathtt{1/9}$, while $\langle \mathtt{meat}\rangle$ has probability $\mathtt{1/3}$, and $\langle \varepsilon, \varepsilon\rangle$ and $\langle \epsilon, \mathtt{red}\rangle$ have probability $\mathtt{1/6}$. Thus, under complete semantics, since in our example every extension has exactly one explanation, the extensions $\{\mathtt{fish}, \mathtt{white}\}$, $\{\mathtt{fish}, \mathtt{red}\}$, and $\{\mathtt{fish}\}$ have probability $\mathtt{1/9}$, while $\{\mathtt{meat}, \mathtt{red}\}$ has probability $\mathtt{1/3}$, and the empty set and $\{\varepsilon, \mathtt{red}\}$ have probability $\mathtt{1/6}$. $\square$

The following proposition states that, the fact that we have chosen a fixed but arbitrary linear ordering on the SCCs (according to the topological ordering of the graph representing the AF) has no effect on the computation of the explanation-based probability of extensions.

Proposition 3. *Let $\Lambda = \langle A, \Sigma\rangle$ be an AF, σ a semantics in $\{\mathtt{gr}, \mathtt{co}, \mathtt{pr}, \mathtt{st}, \mathtt{sst}\}$ and E a σ-extension. Let $\mathcal{C}$ and $\mathcal{D}$ be two linear orderings of the SCCs of Λ (according to the topological ordering of the graph representing the AF Λ). Let $PrE_{\mathcal{O}}(E, \Lambda, \sigma)$ be the probability associated with extension E under a linear ordering $\mathcal{O}$. Then, it holds that $PrE_{\mathcal{C}}(E, \Lambda, \sigma) = PrE_{\mathcal{D}}(E, \Lambda, \sigma)$.*

Probabilistic AFs As for probabilistic AF Δ, we have to consider all AFs w in the set $pw(\Delta)$ of possible worlds. Thus, each extension in $\sigma(\Delta)$ has associated an (explanation-based) probabilistic value defined as follows:

$$PrE(E, \Delta, \sigma) = \sum_{w \in pw(\Delta)} \mathcal{I}(w) \cdot PrE(E, w, \sigma) \tag{6}$$

Moreover, we define $PrEA^{\sigma}_{\Delta}(g)$ as the probability obtained by using $PrE(E, w, \sigma)$ as an instantiation of the PDF required in Definition 3. That is, we have obtained an instantiation of our Probabilistic Acceptance problem, that we call *Explanation-based Probabilistic Acceptance* problem, whose output is $PrEA^{\sigma}_{\Delta}(g)$. In particular, the Explanation-based Probabilistic Acceptance of a goal argument is as follows.

Definition 7 *(Explanation-based probabilistic acceptance).* Given a PrAF $\Delta = \langle A, \Sigma, P\rangle$, a semantics σ and an argument $g \in A$, the probability $PrEA^{\sigma}_{\Delta}(g)$ that g is explanation-based acceptable w.r.t. semantics σ is

$$PrEA^{\sigma}_{\Delta}(g) = \sum_{E \in \sigma(\Delta) \wedge g \in E} PrE(E, \Delta, \sigma). \tag{7}$$

The next example shows how the explanations-based probabilistic acceptance is computed for extensions and for goal arguments.

Example 15. Consider the PrAF Δ of Example 1. As shown in Example 2, for Δ there are four (non-zero probability) possible worlds whose probabilities are given in Example 6.

Let $E_1 = \{\mathtt{fish}, \mathtt{white}\}$, $E_2 = \{\mathtt{fish}, \mathtt{red}\}$ and $E_3 = \{\mathtt{meat}, \mathtt{red}\}$. We have that $\mathtt{pr}(w_1) = \{E_1, E_2, E_3\}$, $\mathtt{pr}(w_2) = \{E_2, E_3\}$, $\mathtt{pr}(w_3) = \{E_3\}$, and $\mathtt{pr}(w_4) = \{E_3\}$. The following table reports for each world w the probability $\mathcal{I}(w)$ (second column) and, for each pair $\langle$world w, preferred extension $E\rangle$, the probability $PrE(E, w, \mathtt{pr})$ w.r.t. the AF w (last three columns). Finally, the last row of the table reports, for each $E \in \{E_1, E_2, E_3\}$, the probability $PrE(E, \Delta, \sigma)$ that E is a $\mathtt{pr}$-extension of the Δ. For instance, the probability that E_2 is a $\mathtt{pr}$-extension of Δ is $0.48 \cdot 1/4 + 0.12 \cdot 1/2 + 0.32 \cdot 0 + 0.08 \cdot 0 = 0.18$ (see the last row of following table).

		$PrE(E, w, \mathtt{pr})$		
w	$\mathcal{I}(w)$	$E_1 = \{\mathtt{fish}, \mathtt{white}\}$	$E_2 = \{\mathtt{fish}, \mathtt{red}\}$	$E_3 = \{\mathtt{meat}, \mathtt{red}\}$
w_1	0.48	1/4	1/4	1/2
w_2	0.12	0	1/2	1/2
w_3	0.32	0	0	1
w_4	0.08	0	0	1
$PrE(E, \Delta, \sigma)$		0.12	0.18	0.70

According to Definition 7, the probability of acceptance of a goal argument in Δ is reported in the following table. For instance,
$PrEA^{\mathtt{pr}}_{\Delta}(\mathtt{fish}) = 0.12 + 0.18 = 0.30$, whereas $PrEA^{\mathtt{pr}}_{\Delta}(\mathtt{meat}) = 0.70$.

	fish	meat	white	red
$PrEA^{\mathtt{pr}}_{\Delta}(g)$	0.30	0.70	0.12	0.88

□

5. Exact and approximate complexity

In this section, we discuss the exact and approximate complexity of probabilistic AF. We consider semantics $\sigma \in \{\mathtt{gr}, \mathtt{co}, \mathtt{pr}, \mathtt{st}, \mathtt{sst}\}$ and concentrate on the following two main problems.

PROBLEM :	PrA[σ]
INPUT :	A PrAF Δ and an argument g.
OUTPUT:	The number $PrA^{\sigma}_{\Delta}(g)$.

PROBLEM :	PrEA[σ]
INPUT :	A PrAF Δ and an argument g.
OUTPUT :	The number $PrEA^{\sigma}_{\Delta}(g)$.

We recall that PrA[σ] is defined after choosing an arbitrary but fixed PDF over the set of extensions of an AF, while PrEA[σ] uses the specific PDF $PrE(\cdot, \Lambda, \sigma)$ of Definition 6, and thus $PrEA^{\sigma}_{\Delta}(g)$ is computed according to Definition 7.

The organization of this section is as follows. Section 5.1 provides the exact complexity for general PrAF, as well as for acyclic PrAF, AF and acyclic AF (the results obtained are summarized in Table 3). Section 5.2 provides inapproximability results for different classes of PrAFs, that are, general PrAF, PrAF without odd cycles, and acyclic PrAF. The results obtained in Section 5.2, along with other results discussed below, are summarized in Table 4. Next, for the cases where approximability has not been ruled out, approximability is studied in Section 5.3. Finally, we conclude with a comparison with other notions of acceptance proposed in the literature in Section 5.4 where inapproximability of the probabilistic credulous and skeptical acceptance problems (PrCA/PrSA) are investigated.

5.1. Exact complexity

We show that for all semantics, the above problems are intractable. In particular, we have the following result.

Theorem 2. *For $\sigma \in \{\mathtt{gr}, \mathtt{co}, \mathtt{pr}, \mathtt{st}, \mathtt{sst}\}$, PrA[$\sigma$] is $FP^{\#P}$-hard, even for acyclic PrAFs and for any chosen PDF.*

Since PrEA is an instantiation of PrA where a specific PDF is used, we obtain the following corollary.

Table 3
Complexity of PrEA[σ] with semantics $\sigma \in \{\texttt{gr}, \texttt{co}, \texttt{st}, \texttt{pr}, \texttt{sst}\}$ for different classes of PrAFs, i.e. general PrAF, acyclic PrAF, AF, and acyclic AF.

	General PrAF	Acyclic PrAF	AF	Acyclic AF
gr	$FP^{\#P}$-h [Corollary 1]	$FP^{\#P}$-h [Corollary 1]	FP [Corollary 2]	FP [Corollary 2]
co, pr, st, sst	$FP^{\#P}$-h [Corollary 1]	$FP^{\#P}$-h [Corollary 1]	$FP^{\#P}$-h [Theorem 3]	FP [Corollary 2]

Corollary 1. *For $\sigma \in \{\texttt{gr}, \texttt{co}, \texttt{pr}, \texttt{st}, \texttt{sst}\}$, PrEA[$\sigma$] is $FP^{\#P}$-hard, even for acyclic PrAFs.*

Intuitively, the proof of Theorem 2 shows that one of the sources of complexity for PrA (and thus for PrEA) is the exponential number of possible worlds to be considered in the computation of $PrA^{\sigma}_{\Delta}(g)$ (and thus of $PrEA^{\sigma}_{\Delta}(g)$). However, as stated next, this is not the only source of complexity.

The following theorem states that, even if restricting PrAF to be an AF, the complexity of PrEA is still $FP^{\#P}$-hard for all semantics except for the grounded.

Theorem 3. *For $\sigma \in \{\texttt{co}, \texttt{pr}, \texttt{st}, \texttt{sst}\}$, PrEA[$\sigma$] is $FP^{\#P}$-hard for AFs (that is, for PrAFs where all probabilities are set to 1).*

Thus, even focusing on a single possible world (i.e. on AF), the proof of Theorem 3 suggests that, under multiple status semantics, an additional source of complexity is the exponential number of extensions to be considered in the computation of $PrA^{\sigma}_{\Delta}(g)$ (and thus of $PrEA^{\sigma}_{\Delta}(g)$).

The following proposition states that for AF the complexity becomes polynomial for the grounded semantics or for acyclic frameworks as they admit a single complete extension (coinciding with the grounded extension).

Proposition 4. *For any chosen PDF, i) PrA[$\texttt{gr}$] is in FP for AF, and ii) PrA[σ] is in FP for $\sigma \in \{\texttt{pr}, \texttt{co}, \texttt{st}, \texttt{sst}\}$ and acyclic AFs.*

Again, since PrEA is an instantiation of PrA where a specific PDF is used, we obtain the following corollary.

Corollary 2. *i) PrEA[$\texttt{gr}$] is in FP for AF, and ii) PrEA[σ] is in FP for $\sigma \in \{\texttt{pr}, \texttt{co}, \texttt{st}, \texttt{sst}\}$ and acyclic AFs.*

The high computational complexity of PrA[σ] (and thus of PrEA[σ]), for all semantics σ, even for very simple settings, such as acyclic PrAFs, suggests that one would need to focus on finding efficient algorithms that solve the problem approximately. Next, we present a quite complete picture of the approximability landscape of our problems, under different semantics and approximation schemes.

5.2. Inapproximability results

Before defining the kind of approximation schemes we are going to target, we recall some basic notions. A *(discrete) probability space* is a pair $\mathsf{PS} = (\Theta, \pi)$, where Θ is a finite set, called *sample space*, and elements therein are called *outcomes*. Furthermore, $\pi : \Theta \to [0, 1]$ is a function such that $\sum_{\theta \in \Theta} \pi(\theta) = 1$, called the *probability distribution* of PS. A subset $\xi \subseteq \Theta$ is called an *event*. The probability of an event ξ (w.r.t. Θ), denoted $\Pr_{\Theta}(\xi)$ is defined as $\sum_{\theta \in \xi} \pi(\theta)$. When Θ is clear from the context, we may simply write Pr. A *random variable* over PS is a function $X : \Theta \to \mathbb{Q}$, which intuitively maps outcomes to some value of interest. For every $x \in \mathbb{Q}$, $X \leq x$ denotes the event $\{\theta \in \Theta \mid X(\theta) \leq x\}$, i.e. all the outcomes whose value of interest, according to X, is at most x. The event $X \geq x$, and other more complex events involving inequalities and other constraints are defined in a similar way.

An algorithm is *randomized* if it outputs a random variable over some probability space PS. Equivalently, one can see a randomized algorithm as an algorithm that has access to a random stream of bits [46].

Definition 8. Consider a function $f : \{0, 1\}^* \to \mathbb{Q}$. A fully polynomial-time randomized approximation scheme (FPRAS) for f is a randomized algorithm A that given as input $x \in \{0, 1\}^*$, and numbers $\epsilon > 0$, $\delta \in (0, 1)$, outputs a random variable $A(x, \epsilon, \delta)$ over some probability space PS such that:

$$\Pr\Big(|A(x, \epsilon, \delta) - f(x)| \leq \epsilon \cdot f(x)\Big) \geq 1 - \delta,$$

and A runs in polynomial time in $|x|$, $1/\epsilon$, and $\ln(1/\delta)$.

Intuitively, in the above definition, the function f represents a certain problem (e.g. PrA), where the input x is the encoding of the PrAF Δ and the argument g, and the value $f(x)$ is the probability of accepting g (e.g. the number $PrA^{\sigma}_{\Delta}(g)$). An FPRAS for $f(x)$ is thus a randomized algorithm $A(x, \epsilon, \delta)$ that computes an approximation of $f(x)$ such that $|A(x, \epsilon, \delta) -$

$f(x)| \leq \epsilon \cdot f(x)$ with a probability greater than or equal to $1 - \delta$. Parameters ϵ and δ in $A(x, \epsilon, \delta)$ are used for example to determine the size of the sample used to compute the approximated value (cf. Algorithm 1).

Similarly to FPRAS, we can define the notion of *additive* FPRAS. A *fully polynomial-time additive randomized approximation scheme* (FPARAS) for a function f is defined as in Definition 8, where the inequality $|A(x, \epsilon, \delta) - f(x)| \leq \epsilon \cdot f(x)$ is replaced with $|A(x, \epsilon, \delta) - f(x)| \leq \epsilon$. The difference between FPARAS and FPRAS is in the kind of error guarantee we are asking. Indeed, an FPARAS provides *additive error* guarantees, i.e. the approximated value differs from the exact one by at most ϵ, while an FPRAS provides a *relative error* guarantee, i.e. the approximated value differs from the exact one by at most an $(\epsilon \cdot f(x))$-factor. In general, FPRAS is preferable to FPARAS as the approximation takes into account the value of the solution by a *relative error*. However, whenever an FPRAS does not exist, an FPARAS is the next natural option to consider.

To start discussing inapproximability results for $\mathsf{PrA}[\sigma]$ (and $\mathsf{PrEA}[\sigma]$) under both FPRAS and FPARAS schemes, we need to recall the class of decision problems BPP. A decision problem Π is in BPP iff there exists a polynomial-time randomized decision procedure A such that, for every instance $x \in \{0, 1\}^*$ of Π, if x is a yes (resp. no) instance of Π, then $Pr(A(x) = \text{yes})$ (resp. $Pr(A(x) = \text{no})$) is greater than or equal to $2/3$. It is known that NP $\subseteq$ BPP implies that the polynomial-time hierarchy collapses [47].

The first result states that $\mathsf{PrA}[\sigma]$ and $\mathsf{PrEA}[\sigma]$ admits no FPRAS for all semantics σ (see second, fourth and sixth column of Table 4).

Theorem 4. *Consider a semantics $\sigma \in \{\mathtt{gr}, \mathtt{co}, \mathtt{pr}, \mathtt{st}, \mathtt{sst}\}$. Unless NP $\subseteq$ BPP, there is no FPRAS for* $\mathsf{PrA}[\sigma]$*, even for acyclic PrAFs and for any chosen PDF.*

Next, we show that for all semantics except for the grounded, even approximation algorithms with bounded additive error (i.e. FPARAS) cannot be devised for PrAFs. For the proof, we rely on a technical lemma that shows a certain gap property of the problem of credulously accepting an argument. We first introduce some further notions.

We say that a pair (Λ, g) of an AF Λ and argument g is *σ-uniform*, for a semantics σ, if the existence of an extension $E \in \sigma(\Lambda)$ such that $g \in E$, implies that every extension $E \in \sigma(\Lambda)$ is such that $g \in E$. In other words, when a pair (Λ, g) is σ-uniform, then credulous and skeptical acceptance of g over Λ coincide.

Let us now consider the following restriction of the classical credulous acceptance problem, where σ is a semantics.

PROBLEM :	$\mathsf{UnCA}[\sigma]$
INPUT :	A σ-uniform pair (Λ, g).
QUESTION :	Is there $E \in \sigma(\Lambda)$ such that $g \in E$?

We show that even when restricting our attention to σ-uniform pairs of AFs and arguments, credulous acceptance is NP-hard, for all semantics in $\{\mathtt{pr}, \mathtt{st}, \mathtt{sst}\}$. The result is proved by providing a reduction from 3SAT by exploiting and adapting the construction (of an AF Λ) known in the literature for the credulous acceptance [48]. Particularly, we prove that Λ has exclusively non-empty σ-extensions if the given formula is satisfiable, and exclusively empty σ-extensions otherwise.

Lemma 1. *For $\sigma \in \{\mathtt{pr}, \mathtt{st}, \mathtt{sst}\}$,* $\mathsf{UnCA}[\sigma]$ *is NP-hard.*

We use the above lemma to prove the inapproximability results stated below.

Theorem 5. *Let $\sigma \in \{\mathtt{pr}, \mathtt{st}, \mathtt{sst}\}$. Unless NP $\subseteq$ BPP, there is no FPARAS for* $\mathsf{PrA}[\sigma]$*, for any chosen PDF.*

Corollary 3. *Let $\sigma \in \{\mathtt{pr}, \mathtt{st}, \mathtt{sst}\}$. There is no FPARAS and no FPRAS for* $\mathsf{PrA}[\sigma]$ *for AF (PrAF with probabilities equal to 1) and for any chosen PDF.*

Theorems 2, 4, and 5 rule out the existence of polynomial-time algorithms for solving $\mathsf{PrA}[\sigma]$. In terms of exact and approximate computation via FPRASes, this is not possible even for acyclic PrAFs, whereas in terms of approximate computation via FPARAS, this is not possible for general PrAFs as well as for AFs, for all semantics $\sigma \in \{\mathtt{pr}, \mathtt{st}, \mathtt{sst}\}$. Notably, our results highlight an intrinsic difficulty in providing efficient procedures (either exact or approximate) for *any* approach assigning a probability to an argument by means of a probability distribution over the extensions.

From the above discussion, it is clear that our efforts should be towards approximation schemes with bounded additive error guarantees, i.e. FPARAS. In particular, in the light of Theorem 5, one could still provide an FPARAS when either $\sigma \in \{\mathtt{gr}, \mathtt{co}\}$ or some restriction on the input PrAF is assumed. In fact, we are going to show that either when $\sigma = \mathtt{gr}$ or when the input PrAF has no odd-length cycles, the use of explanations for devising a PDF over extensions allows us to construct an FPARAS.

Algorithm 1 Apx.

Input: A PrAF $\Delta = \langle A, \Sigma, P\rangle$, a semantics σ, a goal argument $g \in A$, error parameter $\epsilon > 0$, and uncertainty parameter $0 < \delta < 1$.
Output: a random number p such that $PrEA^{\sigma}_{\Delta}(g) \in [p-\epsilon, p+\epsilon]$ with probability $1-\delta$.

1: $n := \lceil \frac{1}{2\epsilon^2} \times \ln(\frac{2}{\delta}) \rceil$;
2: $c := 0$;
3: **for** $i \in \{1, \ldots, n\}$ **do**
4: Choose $w \in pw(\Delta)$ with probability $\mathcal{I}(w)$;
5: Choose $X \in Exp^{\sigma}(w)$ with probability $\pi(X)$;
6: **if** $g \in ext(X)$ **then**
7: $c := c + 1$;
8: **return** $\frac{c}{n}$;

Algorithm 2

Input: An AF $\Lambda = \langle A, \Sigma\rangle$ and a semantics σ.
Output: An explanation for a σ-extension.

1: Let $X = \langle\rangle$; $\Omega = \emptyset$; $\Gamma = \emptyset$; $\Lambda = \Lambda^*$;
2: **while** $\Lambda \neq \langle\emptyset, \emptyset\rangle$ **do**
3: Let A' be the first SCC of Λ ;
4: Let $\mathcal{C} = \{a \in A' \setminus \Omega \mid \Omega \neq \emptyset \Rightarrow (\Omega \cap G_a^+) \neq \emptyset \quad \text{with } G_a = \mathtt{gr}(\Lambda_a)\} \setminus \Gamma$;
5: **if** $\sigma = \mathtt{gr}$ **then**
6: $\mathcal{C} = \{\varepsilon\}$;
7: **if** $\sigma = \mathtt{co} \wedge \Omega = \emptyset$ **then**
8: $\mathcal{C} = \mathcal{C} \cup \{\varepsilon\}$;
9: Select $a \in \mathcal{C}$ with probability $\frac{1}{|\mathcal{C}|}$;
10: Append a to X;
11: **if** $a = \varepsilon$ **then**
12: $\Gamma = \Gamma \cup \{x \in A \setminus A' \mid \exists\ (y, x) \in \Sigma \text{ with } y \in A'\}$;
13: $\Lambda = \Lambda_{\downarrow_{A \setminus A'}}$;
14: **else**
15: $\Omega = \Omega_a$;
16: $\Lambda = \widehat{\Lambda_a}$;
17: **return** X

5.3. Devising an FPARAS

We report now an FPARAS for the problem PrEA[σ] for the cases where either the semantics is the grounded or the input PrAF has no odd-length cycles.

The general structure of our algorithm is presented in Algorithm 1. Consider a PrAF Δ, a semantics σ and an argument g. The high-level idea is to perform a number of iterations n, and at each iteration sample a world w of Δ and an explanation X in $Exp^{\sigma}(w)$, and count the fraction of iterations for which the given argument g is in the σ-extension $ext(X)$ explained by X.

We point out that, besides line 5, all steps of our algorithm can be easily implemented in polynomial time regardless of the shape of the input PrAF and the semantics. Particularly, to prove that Algorithm 1 leads to an FPARAS in the cases described above, it suffices to prove that line 5 can be implemented in polynomial time when either $\sigma = \mathtt{gr}$ or Δ has no odd-length cycles. This is done via Algorithm 2 which will be discussed later. Thus, Algorithm 1 enjoys the probabilistic and error guarantees of an FPARAS (this can be proved via standard probabilistic inequalities [49]).

Algorithm 2, receives as input an AF $\Lambda = \langle A, \Sigma\rangle$ and a semantics σ, and returns an explanation in $Exp^{\sigma}(\Lambda)$. It uses variables (i) Ω, containing the set of arguments which must be determined as false in the next steps as they attack arguments which have been assumed to be true, i.e. every time we add an argument a to X we add to Ω the attackers of a and then delete the arguments whose status is determined by such a choice, and (ii) Γ, containing the set of arguments whose status cannot be determined as true in the next steps. Γ is relevant only in computation of complete extensions, as for stable (preferred and semi-stable) semantics, arguments are either accepted or defeated.[4]

It is worth noting that during the computation the AF changes as we delete from it arguments whose status has been determined, and add virtual attacks to maintain the topology of the SCCs (as discussed in Section 4)

Algorithm 2 proceeds as follows. At the beginning variables X, Ω and Γ are initialized (i.e. $X = \langle\rangle$, $\Omega = \Gamma = \emptyset$), whereas Λ is updated to Λ^* by deleting arguments occurring in the grounded extension or defeated by them (line 1). Clearly, if the source AF is acyclic we have that Λ (which is now equal to Λ^*) is empty and the algorithm terminates, returning the empty explanation (with probability 1). Then, it iterates until the current AF Λ becomes empty (line 2) and the following steps are executed iteratively. It determines the set $\mathcal{C}$ of arguments in the first SCC, from which it is possible to choose the next element (lines 3 and 4). Moreover, if $\sigma = \mathtt{gr}$, $\mathcal{C}$ is set to $\{\varepsilon\}$ (line 6) so that the only possible choice will be ε, whereas if $\sigma = \mathtt{co}$ and Ω is empty, ε is added to $\mathcal{C}$, as it is also an admissible choice (line 8). At this point an element a is nondeterministically selected from $\mathcal{C}$ with probability $1/|\mathcal{C}|$ (line 9) and it is added to X (line 10). The next steps

[4] Recall that for odd-cycle free AF, preferred, stable and semi-stable semantics coincide.

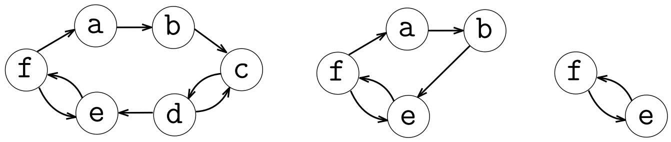

Fig. 10. AFs Λ (left), Λ' (center), and Λ'' (right) of Example 16.

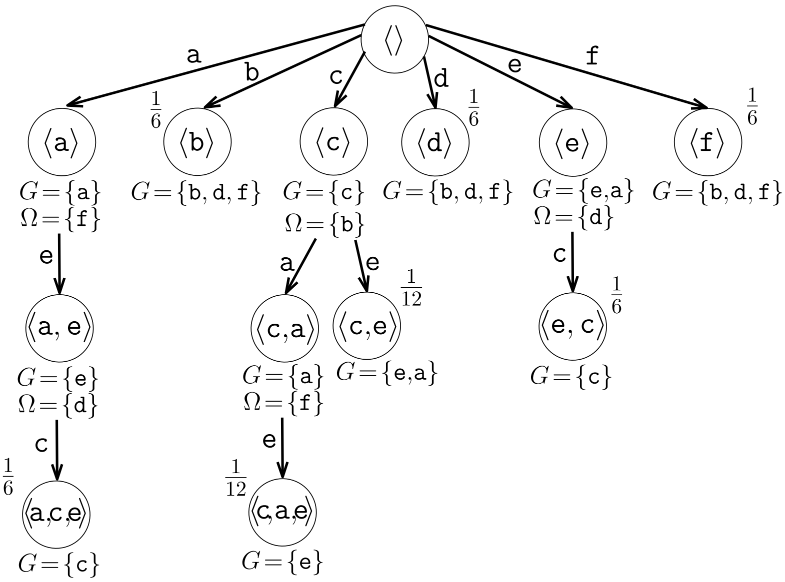

Fig. 11. Probabilistic trie for the AF Λ of Example 16 under preferred/stable/semi-stable semantics.

depend on the choice made (line 7). If $a = \varepsilon$ all arguments in the first component are deleted from Λ (line 13) and all arguments attacked by these arguments are added to Γ to remember that their status cannot be true and, thus, they cannot be chosen in the next steps (line 12). By deleting the whole components and adding ε to X we are stating that the status of all elements in the component is undefined. If the chosen element a is an argument, the following steps are executed: *i*) the attackers of a are added to Ω as we are assuming that their status must be false, and all elements whose status is determined (i.e. those in G_a^+) are deleted from Ω (line 15); *ii*) the AF Λ is updated to $\widehat{\Lambda_a}$ (line 16), by deleting arguments in G_a^* (and related attacks) whose status has been determined, and adding virtual attacks from nodes in $(a^- \setminus G_a^+)$ to nodes in $((G_a^+)^+ \setminus G_a^*)$, so that the resulting component continues to be strongly connected by reconstructing paths in the graph that were broken through the deletion of nodes.

The next two examples illustrate how explanations are computed under preferred and complete semantics, respectively. In the examples, superscripts with natural numbers are used to distinguish the values of variables at each iteration.

Example 16. Consider the AF $\Lambda = \langle A, \Sigma\rangle$ shown in Fig. 10 (left) with $A = \{\mathtt{a}, \mathtt{b}, \mathtt{c}, \mathtt{d}, \mathtt{e}, \mathtt{f}\}$ and $\Sigma = \{(\mathtt{a}, \mathtt{b}), (\mathtt{b}, \mathtt{c}), (\mathtt{c}, \mathtt{d}), (\mathtt{d}, \mathtt{e}), (\mathtt{e}, \mathtt{f}), (\mathtt{f}, \mathtt{a}), (\mathtt{d}, \mathtt{c}), (\mathtt{f}, \mathtt{e})\}$ whose preferred (and stable and semi-stable) extensions are $E_1 = \{\mathtt{a}, \mathtt{c}, \mathtt{e}\}$ and $E_2 = \{\mathtt{b}, \mathtt{d}, \mathtt{f}\}$. The set of possible explanations which can be computed by Algorithm 2 are shown in Fig. 11. Initially, the first SCC is $\mathcal{C} = A$ and, as $\Omega^0 = \emptyset$ and $\Lambda^0 = \Lambda$, there are 6 possible choices with probability 1/6.

The choice of an argument $x \in \{\mathtt{b}, \mathtt{d}, \mathtt{f}\}$ allows to have that $G_x = \{\mathtt{b}, \mathtt{d}, \mathtt{f}\}$ and, therefore, $G_x^+ = \{\mathtt{a}, \mathtt{c}, \mathtt{e}\}$. Consequently, both the updated extension and the updated AF are empty and we have three explanations $\langle\mathtt{b}\rangle$, $\langle\mathtt{d}\rangle$ and $\langle\mathtt{f}\rangle$ for E_2, all with probability $\mathtt{1/6}$.

By choosing $\mathtt{c}$ (with probability 1/6) we get $G_\mathtt{c} = \{\mathtt{c}\}$, $G_\mathtt{c}^+ = \{\mathtt{d}\}$ and $\Omega^1 = \Omega_\mathtt{c}^0 = \{\mathtt{b}\}$. The derived AF $\Lambda^1 = \widehat{\Lambda_\mathtt{c}^0}$ is obtained by deleting arguments in $\{\mathtt{c}, \mathtt{d}\}$ (i.e. arguments in $G_\mathtt{c}$ and $G_\mathtt{c}^+$) and adding the virtual attack $(\mathtt{b}, \mathtt{e})$ (see Fig. 10 (center)). After this choice we have that $\mathcal{C} = \{\mathtt{a}, \mathtt{e}\}$ as, since Ω^1 is not empty, we have two possible choices $x \in \{\mathtt{a}, \mathtt{e}\}$ such that $(\mathtt{gr}(\Lambda_x)^+ \cap \Omega) \neq \emptyset$, with probability 1/2. By choosing $\mathtt{e}$ we get $G_\mathtt{e} = \{\mathtt{e}, \mathtt{a}\}$ which is a preferred extension of Λ^1, whereas if we choose $\mathtt{a}$, we get $G_\mathtt{a} = \{\mathtt{a}\}$, $G_\mathtt{a}^+ = \{\mathtt{b}\}$ and $\Omega^2 = \Omega_\mathtt{a}^1 = \{\mathtt{f}\}$. The derived AF $\Lambda^2 = \widehat{\Lambda_\mathtt{a}^1}$ (shown in Fig. 10 (right)) is obtained by deleting arguments in $\{\mathtt{a}, \mathtt{b}\}$ and adding the edge $(\mathtt{f}, \mathtt{e})$ (which was already present). At this point $\mathcal{C} = \{\mathtt{e}\}$ and by choosing $\mathtt{e}$ we have that $(\mathtt{gr}(\Lambda_c^2) \cap \Omega^2) \neq \emptyset$. Therefore, we derive the explanations $\langle\mathtt{c}, \mathtt{e}\rangle$ and $\langle\mathtt{c}, \mathtt{a}, \mathtt{e}\rangle$, both with probability 1/12.

By choosing first $\mathtt{a}$ we get the explanations $\langle\mathtt{a}, \mathtt{e}, \mathtt{c}\rangle$, whereas by choosing $\mathtt{e}$ we get the explanation $\langle\mathtt{e}, \mathtt{c}\rangle$ both with probability 1/6.

Table 4
Approximability of PrEA[σ]/PrCA[σ]/PrSA[σ] for $\sigma \in \{\texttt{gr}, \texttt{co}, \texttt{pr}, \texttt{st}, \texttt{sst}\}$ and different classes of PrAFs, i.e. general PrAF, PrAF admitting even-length cycles only, and acyclic PrAF. Non-existence (resp. existence) of an FP(A)RAS, i.e. a Fully Polynomial-time (Additive) Randomized Approximation Scheme, is denoted with $\times$ (resp. ✓) in the corresponding column. Results for PrA[σ] coincide with that of PrEA[σ].

	General PrAF		PrAF w/o odd cycles		acyclic PrAF	
	FPRAS	FPARAS	FPRAS	FPARAS	FPRAS	FPARAS
gr	×/×/×	✓/✓/✓	×/×/×	✓/✓/✓	×/×/×	✓/✓/✓
co	×/×/×	open/×/✓	×/×/×	✓/×/✓	×/×/×	✓/✓/✓
pr, st, sst	×/×/×	×/×/×	×/×/×	✓/×/×	×/×/×	✓/✓/✓

Example 17. Considering the AF $\Lambda = \langle A, \Sigma \rangle$ of Example 5 (coinciding with possible world w_1 of Example 2) shown in Fig. 6 (left). The set of possible explanations which can be computed by Algorithm 2 under complete semantics can be represented through the trie shown in Fig. 9 (right), though the sets Ω, Γ and their evolution are not there. Initially, as Ω^0 and Γ^0 are empty, and $\Lambda^0 = \Lambda$, $\mathcal{C} = \{\texttt{fish}, \texttt{meat}, \varepsilon\}$ and thus there are 3 possible choices with probability $1/3$. The choice of argument meat allows to have that $G_{\texttt{meat}} = \{\texttt{meat}, \texttt{red}\}$ which corresponds to the extension E_3.

By choosing fish (with probability $1/3$) we get $G_{\texttt{fish}} = \{\texttt{fish}\}$, $G^+_{\texttt{fish}} = \{\texttt{meat}\}$ and $\Omega = \Gamma = \emptyset$. The derived AF $\Lambda^1 = \widehat{\Lambda^0_{\texttt{fish}}}$ is obtained by deleting arguments in $\{\texttt{fish}, \texttt{meat}\}$ (i.e. arguments in $G^*_{\texttt{fish}}$). After this choice we have that $\mathcal{C} = \{\texttt{white}, \texttt{red}, \varepsilon\}$. Thus we have again three possible choices with probability $1/3$. By choosing white (resp. red and ε) we get $G_{\texttt{white}} = \{\texttt{white}\}$ (resp. $G_{\texttt{red}} = \{\texttt{red}\}$, and $G_\varepsilon = \emptyset$) and, thus, we end with the three explanations $\langle \texttt{fish}, \texttt{red} \rangle$, $\langle \texttt{fish}, \texttt{white} \rangle$ and $\langle \texttt{fish}, \varepsilon \rangle$.

By choosing ε (with probability 1/3) we get $\Gamma^1 = \{\texttt{white}\}$ and the derived AF $\Lambda^1 = \Lambda^0_{\downarrow A \setminus \{\texttt{fish}, \texttt{meat}\}}$ is obtained by deleting arguments in $\{\texttt{fish}, \texttt{meat}\}$ (i.e. arguments in the first SCC of Λ). After this choice we have that $\mathcal{C} = \{\texttt{red}, \texttt{white}, \varepsilon\} \setminus \Gamma^1 = \{\texttt{red}, \varepsilon\}$. Thus we have two possible choices with probability 1/2. By choosing red (resp. ε) we get $G_{\texttt{red}} = \{\texttt{red}\}$ (resp. $\emptyset$) and, thus, we end with the two explanations $\langle \varepsilon, \texttt{red} \rangle$ and $\langle \varepsilon, \varepsilon \rangle$. $\square$

Theorem 6. *Algorithm 2 with input an AF Λ and a semantics $\sigma \in \{\texttt{gr}, \texttt{co}, \texttt{pr}, \texttt{st}, \texttt{sst}\}$ is such that:*

- *It outputs an $X \in Exp^\sigma(\Lambda)$ with probability $\pi(X)$, and*
- *it runs in polynomial time,*

whenever i) $\sigma = \texttt{gr}$, or ii) Λ has no odd-length cycles.

By exploiting the above result and standard probabilistic inequalities [49], we can prove our main approximability result.

Theorem 7. *Problem* PrEA[σ], *with $\sigma \in \{\texttt{gr}, \texttt{co}, \texttt{pr}, \texttt{st}, \texttt{sst}\}$, has an FPARAS if either i) $\sigma = \texttt{gr}$, or ii) the input PrAF has no odd cycles.*

5.4. Inapproximability for credulous and skeptical acceptance

We conclude Section 5 by investigating the approximate complexity of the problems of probabilistic credulous acceptance (PrCA) and probabilistic skeptical acceptance (PrSA). As for PrEA, we show that no relative error approximation algorithm (i.e. FPRAS) exists for PrCA and PrSA under any semantics, even considering acyclic PrAFs. However, we can show that PrCA and PrSA are harder than PrEA in the sense that no additive error approximation algorithm (i.e. FPARAS) exists for PrCA and PrSA under preferred, stable, and semi-stable semantics even when PrAFs have no odd-length cycles (in contrast, FPARAS schemes exist for PrEA under such conditions, cf. Theorem 7). Approximability results for PrEA, PrCA, and PrSA are summarized in Table 4.

We start by formally defining the problems considered in this section. For a semantics σ and a goal argument g, we consider the following two problems, where the numbers $PrCA^\sigma_\Delta(g)$ and $PrSA^\sigma_\Delta(g)$ are defined as shown in Definition 2:

PROBLEM :	PrCA[σ]
INPUT :	A PrAF Δ and an argument g.
OUTPUT :	The number $PrCA^\sigma_\Delta(g)$.

PROBLEM :	PrSA[σ]
INPUT :	A PrAF Δ and an argument g.
OUTPUT :	The number $PrSA^\sigma_\Delta(g)$.

It is known that PrCA[σ] and PrSA[σ] are FP$^{\#P}$-hard [33], and this holds even for acyclic PrAFs. As for PrEA[σ], we can show that PrCA[σ] and PrSA[σ] admits no FPRAS for all semantics σ (see the second, the fourth and the sixth column of Table 4).

Theorem 8. *For $\sigma \in \{\mathtt{gr}, \mathtt{co}, \mathtt{pr}, \mathtt{st}, \mathtt{sst}\}$, unless NP $\subseteq$ BPP, there is no FPRAS for* PrCA[σ] *and* PrSA[σ]*, even for acyclic PrAFs.*

The main difference with PrEA[σ] lies in the approximability via FPARAS. In particular, we can show that PrCA[σ] and PrSA[σ] are harder than PrEA[σ] in this regard, as no FPARAS exists, for all semantics $\sigma \in \{\mathtt{pr}, \mathtt{st}, \mathtt{sst}\}$, even when PrAFs have no odd-length cycles (see the third and fifth columns in the last row of Table 4).

Theorem 9. *Unless $NP \subseteq BPP$, i) there is no FPARAS for* PrCA[σ] *and for* PrSA[σ] *with $\sigma \in \{\mathtt{pr}, \mathtt{st}, \mathtt{sst}\}$, even for PrAFs without odd-length cycles; and ii) there is no FPARAS for* PrCA[$\mathtt{co}$]*, even for AFs (PrAFs with probabilities all equal to 1).*

However, a positive result we can obtain for the problem PrCA[σ] (resp. PrSA[σ]) is when $\sigma = \mathtt{gr}$ (resp. $\sigma \in \{\mathtt{gr}, \mathtt{co}\}$). In this case, PrEA[$\sigma$] = PrCA[$\sigma$], and PrSA[$\sigma$] = PrCA[$\mathtt{gr}$], and thus we obtain the following corollary from Theorem 7 (see the third and the fourth rows of the second column of Table 4).

Corollary 4. PrCA[$\mathtt{gr}$], PrSA[$\mathtt{gr}$] *and* PrSA[$\mathtt{co}$] *admit an FPARAS.*

6. Extended PrAFs

This section is devoted to studying an extension of PrAF, called *Extended PrAF* (EPrAF), where precise probabilities are not known, and each element has associated a probability interval. We consider probabilistic AF where only arguments may be uncertain. Thus, each argument has associated a probability interval $[p_1, p_2]$, with $p_1, p_2 \in [0, 1]$. The importance of imprecise probabilities has been observed by numerous researchers [50] and has generated an interest in several contexts, including Probabilistic SAT (PSAT) [15], Probabilistic Logic [16], Probabilistic Logic Programming [51], Probabilistic Databases [52], Spatio-Temporal Knowledge Bases [53,54]. Thus, inspired by the way of representing probabilities in these approaches, we extend the probabilistic argumentation framework so that arguments have associated probabilistic intervals instead of specific values.

Definition 9. An Extended Probabilistic Argumentation Framework (EPrAF) is a triple $\langle A, \Sigma, P_I\rangle$ where $\langle A, \Sigma\rangle$ is an *AF*, and P_I is a function assigning to every argument in A a probability interval $[p_1, p_2]$ with $p_1, p_2 \in [0, 1]$ and $p_1 \le p_2$.

Thus, $P_I(a)$ is a vector of two elements, which denotes an interval whenever $P_I(a)[1] \le P_I(a)[2]$, where $P_I(a)[i]$ denotes the i-th element of $P_I(a)$.

The (interval) *interpretation* for an EPrAF $\Delta = \langle A, \Sigma, P_I\rangle$ is a distribution function $\mathcal{I}_I$ over the set $pw(\Delta)$ of possible worlds assigning to each $w = \langle A', \Sigma'\rangle \in pw(\Delta)$ a probability interval defined as follows:

$$\mathcal{I}_I(w) = \left[\prod_{a\in A'} P_I(a)[1] \cdot \prod_{a\in A\setminus A'} (1-P_I(a)[1]),\ \prod_{a\in A'} P_I(a)[2] \cdot \prod_{a\in A\setminus A'} (1-P_I(a)[2])\right]^*$$

where $[p_1, p_2]^* = [p_1, p_2]$ if $p_1 \le p_2$, otherwise $[p_1, p_2]^* = [p_2, p_1]$.

Recalling that the only possible worlds that contribute to compute probabilistic acceptance of arguments are those with $\mathcal{I}_I(w) > 0$, in the above Equation we can only consider arguments $a \in A$ such that $0 < P_I(a) < 1$, as the deletion of an argument a with $P_I(a) = [1, 1]$ (resp. the addition to an argument a with $P_I(a) = [0, 0]$) gives rise to a possible world w with $\mathcal{I}_I(w) = [0, 0]$.

Example 18. Consider the EPrAF derived from the PrAF of Example (1) by replacing the probability distribution function with the function P_I defined as follows: $P_I(\mathtt{f}) = [0.55, 0.65]$, $P_I(\mathtt{w}) = [0.75, 0.85]$ (where arguments are denoted by their initials for the sake of brevity). Considering the possible worlds w_1, w_2, w_3, w_4 of Example 2, the probabilistic intervals of each interpretation (after rounding values to the second decimal) are as follows:

- $\mathcal{I}_I(w_1) = [P_I(\mathtt{f})[1] \cdot P_I(\mathtt{w})[1], P_I(\mathtt{f})[2] \cdot P_I(\mathtt{w})[2]]^*$
  $= [0.55 \cdot 0.75, 0.65 \cdot 0.85]^* = [0.41, 0.55]^* = [0.41, 0.55]$;
- $\mathcal{I}_I(w_2) = [P_I(\mathtt{f})[1] \cdot (1\text{-}P_I(\mathtt{w})[1]), P_I(\mathtt{f})[2] \cdot (1\text{-}P_I(\mathtt{w})[2])]^*$
  $= [0.55 \cdot 0.25, 0.65 \cdot 0.15]^* = [0.14, 0.10]^* = [0.10, 0.14]$;
- $\mathcal{I}_I(w_3) = [(1\text{-}P_I(\mathtt{f})[1]) \cdot P_I(\mathtt{w})[1], (1\text{-}P_I(\mathtt{f})[2]) \cdot P_I(\mathtt{w})[2]]^*$
  $= [0.45 \cdot 0.75, 0.35 \cdot 0.85]^* = [0.34, 0.31]^* = [0.30, 0.34]$;
- $\mathcal{I}_I(w_4) = [(1\text{-}P_I(\mathtt{f})[1]) \cdot (1 - P_I(\mathtt{w})[1]), (1\text{-}P_I(\mathtt{f})[2]) \cdot (1\text{-}P_I(\mathtt{w})[2])]^*$
  $= [0.45 \cdot 0.25, 0.35 \cdot 0.15]^* = [0.11, 0.05]^* = [0.05, 0.11]$. □

Analogously to what has been defined for PrAF, each extension in $\sigma(\Delta)$ has associated an (explanation-based) probabilistic interval defined by considering all possible worlds as follows:

$$EPr(E,\Delta,\sigma) = \left[\sum_{w \in pw(\Delta)} \mathcal{I}_I(w)[1] \cdot Pr(E,w,\sigma), \sum_{w \in pw(\Delta)} \mathcal{I}_I(w)[2] \cdot Pr(E,w,\sigma)\right]^* \tag{8}$$

Clearly, whenever Pr is replaced with PrE we obtain $EPr(E,\Delta,\sigma)$, that is the explanation-based probability interval associated to σ-extension E of Δ.

The next definition introduces the extended probabilistic acceptance for EPrAF.

Definition 10 *(Probabilistic acceptance for EPrAFs).* Given an EPrAF $\Delta = \langle A, \Sigma, P_I\rangle$ and an argument $g \in A$, the probability interval $EPrA^{\sigma}_{\Delta}(g)$ that g is acceptable w.r.t. semantics σ is:

$$EPrA^{\sigma}_{\Delta}(g) = \left[\sum_{E \in \sigma(\Delta) \wedge g \in E} EPr(E,\Delta,\sigma)[1], \sum_{E \in \sigma(\Delta) \wedge g \in E} EPr(E,\Delta,\sigma)[2]\right]^* \tag{9}$$

Moreover, we define $EPrEA^{\sigma}_{\Delta}(g)$ as the interval probability obtained as above but using the PDF $PrE(\cdot,w,\sigma)$. The corresponding problem is called *Extended Explanation-based Probabilistic Acceptance*.

Theorem 10. *Given an EPrAF $\Delta = \langle A, \Sigma, P_I\rangle$, for any PrAF $\Delta' = \langle A, \Sigma, P\rangle$ such that $P(a) \in P_I(a)$ for all $a \in A$, it holds that $PrA^{\sigma}_{\Delta'}(g) \in EPrA^{\sigma}_{\Delta}(g)$ for all $g \in A$, independently of the PDF $Pr(\cdot,w,\sigma)$ used.*

The theorem entails that the complexity results of the previous sections, stated for PrAFs, also hold for EPrAFs.

Example 19. Consider the EPrAF $\Delta = \langle A, \Sigma, P_I\rangle$ of Example 18 and the PrAF $\Delta' = \langle A, \Sigma, P\rangle$ of Example 1. As we have that $P(a) \in P_I(a)$ for all $a \in A$ (with $P(a)$ being the average value of $P_I(a)[1]$ and $P_I(a)[2]$ $\forall a \in A$), we have that $PrA^{\sigma}_{\Delta}(g) \in EPrA^{\sigma}_{\Delta}(g)$ for all $g \in A$ and any chosen PDF.

Indeed, considering the Extended Explanation-based Probabilistic Acceptance (i.e. *EPrEA*), the following table reports for each world w the probability $\mathcal{I}_I(w)$ (second column) and, for each pair $\langle$world w, preferred extension $E\rangle$, the probability $PrE(E,w,\mathtt{pr})$ w.r.t. AF w (last three columns). Finally, the last row of the table reports, for each set E, the probability interval that E is a pr-extension of the EPrAF Δ.

		$PrE(E,w,\mathtt{pr})$		
w	$\mathcal{I}_I(w)$	E_1={fish, white}	E_2={fish, red}	E_3={meat, red}
w_1	[.41, .55]	1/4	1/4	1/2
w_2	[.10, .14]	0	1/2	1/2
w_3	[.30, .34]	0	0	1
w_4	[.05, .11]	0	0	1
$EPrE(E,\Delta,\mathtt{pr})$		[.41/4, .55/4]* =[0.103, 0.138]	[.41/4+.10/2, .55/4+.14/2]* =[0.153, 0.208]	[.41/2+.10/2+.3+.05, .55/2+.14/2+.34+.11]* =[0.605, 0.795]

The probability of acceptance of a goal argument in Δ (resp. Δ') is reported in first (resp. second) row of the following table. For instance, using Definition 10, we obtain $EPrEA^{\mathtt{pr}}_{\Delta}(\mathtt{fish}) = [0.103 + 0.153, 0.138 + 0.208] = [0.258, 0.346]$.

	fish	meat	white	red
$EPrEA^{\mathtt{pr}}_{\Delta}(g)$	[0.258, 0.346]	[0.605, 0.795]	[0.103, 0.138]	[0.758, 1]
$PrEA^{\mathtt{pr}}_{\Delta}(g)$	0.30	0.70	0.12	0.88

□

7. Incomplete argumentation framework

Incomplete AF has been introduced in [40] and further investigated in several recent works [55–57].

Definition 11 *(Incomplete AF).* An *incomplete (abstract) Argumentation Framework* (iAF) is a tuple $\Delta = \langle A, B, R, T\rangle$, where A and B are disjoint sets of arguments, and R and T are disjoint sets of attacks between arguments in $A \cup B$. Arguments in A and attacks in R are said to be *certain*, while arguments in B and attacks in T are said to be *uncertain*.

Certain arguments in A are definitely known to exist, while uncertain arguments in B are not known for sure: they may occur or may not. Analogously, certain attacks in R are definitely known to exist if both the incident arguments exist, while for uncertain attacks in T it is not known for sure if they hold, even if both the incident arguments exist.

An iAF compactly represents alternative AF scenarios, called *completions*.

Definition 12 *(Completion).* A *completion* for an iAF $\Delta = \langle A, B, R, T\rangle$ is an AF $\Lambda = \langle A', R'\rangle$ where $A \subseteq A' \subseteq A \cup B$ and $R \cap (A' \times A') \subseteq R' \subseteq (R \cup T) \cap (A' \times A')$.

The set of completions of Δ is denoted by $comp(\Delta)$. An iAF $\langle A, B, R, T\rangle$ is *acyclic* (resp. *odd-cycle free*) iff the AF $\langle A \cup B, R \cup T\rangle$ is acyclic (resp. odd-cycle free).

Acceptance problems Credulous and skeptical acceptance for iAF have been proposed in [58], where the goal, i.e. the element for which acceptance is checked, is an argument.

Definition 13 *(Possible/necessary credulous/skeptical acceptance).* Let $\Delta = \langle A, B, R, T\rangle$ be an iAF and $\sigma \in \{\mathtt{gr}, \mathtt{co}, \mathtt{st}, \mathtt{pr}, \mathtt{sst}\}$. Then, an argument $g \in A \cup B$ is said to be:

1. *possibly credulously accepted* under σ, denoted as $PCA_\sigma(\Delta, g)$, iff there exists a completion Λ of Δ such that $CA_\sigma(\Lambda, g)$ is true;
2. *possibly skeptically accepted* under σ, denoted as $PSA_\sigma(\Delta, g)$, iff there exists a completion Λ of Δ such that $SA_\sigma(\Lambda, g)$ is true;
3. *necessarily credulously accepted* under σ, denoted as $NCA_\sigma(\Delta, g)$, iff for every completion Λ of Δ $CA_\sigma(\Lambda, g)$ is true;
4. *necessarily skeptically accepted* under σ, denoted as $NSA_\sigma(\Delta, g)$, iff for every completion Λ of Δ $SA_\sigma(\Lambda, g)$ is true.

We use PCA[σ] (resp. PSA[σ], NCA[σ], NSA[σ]), or simply PCA (resp. PSA, NCA, NSA) whenever σ is understood, to denote the problem of deciding acceptance according to Item 1 (resp. 2, 3, and 4) of Definition 13. For the grounded semantics we have PCA[$\mathtt{gr}$] $\equiv$ PSA[$\mathtt{gr}$] and NCA[$\mathtt{gr}$] $\equiv$ NSA[$\mathtt{gr}$].

Example 20. Consider the iAF $\Delta = \langle\{\mathrm{a}, \mathrm{b}, \mathrm{d}\}, \{\mathrm{c}\}, \{(\mathrm{a}, \mathrm{b}), (\mathrm{b}, \mathrm{a})\}, \emptyset\rangle$. Δ has 2 completions: $\Lambda_1 = \langle\{\mathrm{a}, \mathrm{b}, \mathrm{d}\}, \{(\mathrm{a}, \mathrm{b}), (\mathrm{b}, \mathrm{a})\}\rangle$ and $\Lambda_2 = \langle\{\mathrm{a}, \mathrm{b}, \mathrm{c}, \mathrm{d}\}, \{(\mathrm{a}, \mathrm{b}), (\mathrm{b}, \mathrm{a})\}\rangle$. Under semantics $\sigma \in \{\mathtt{st}, \mathtt{pr}, \mathtt{sst}\}$, Λ_1 has two extensions $E'_1 = \{\mathrm{a}, \mathrm{d}\}$ and $E''_1 = \{\mathrm{b}, \mathrm{d}\}$, while Λ_2 has two extensions $E'_2 = \{\mathrm{a}, \mathrm{c}, \mathrm{d}\}$ and $E''_2 = \{\mathrm{b}, \mathrm{c}, \mathrm{d}\}$. Thus, a, b, c, d satisfy PCA, c, d satisfy PSA, a, b, d satisfy NCA and only d satisfies NSA. □

7.1. Equivalent forms of iAFs

In this section we show that iAFs can be rewritten into equivalent iAFs where uncertainty is restricted to either attacks or (unattacked) arguments.

Definition 14 *(arg-iAF and att-iAF).* An iAF $\Delta = \langle A, B, R, T\rangle$ is said to be *argument-incomplete* (arg-iAF for short) if $T = \emptyset$, whereas it is said to be *attack-incomplete* (att-iAF for short) if $B = \emptyset$.

Given an iAF Δ, we denote by $arg(\Delta)$ the arg-iAF derived from Δ by replacing every uncertain attack (a, b) with the certain attacks $(a, \alpha_{ab}), (\alpha_{ab}, \beta_{ab}), (\beta_{ab}, b)$, where α_{ab} (resp. β_{ab}) is a fresh certain (resp. uncertain) argument. Analogously, we denote by $att(\Delta)$ the att-iAF derived from Δ as follows: for each uncertain argument b, make b certain and add an uncertain attack (α, b), where α is a fresh certain argument—it is sufficient to add only one fresh argument α.

Example 21. Consider the iAF $\Delta = \langle\{\mathrm{b}, \mathrm{c}\}, \{\mathrm{a}\}, \{(\mathrm{a}, \mathrm{b}), (\mathrm{b}, \mathrm{a})\}, \{(\mathrm{b}, \mathrm{c})\}\rangle$ shown in Fig. 12, where dashed nodes (resp. edges) represent uncertain arguments (resp. attacks). The arg-iAF derived from Δ is $arg(\Delta) = \Delta' = \langle\{\mathrm{b}, \mathrm{c}, \alpha_{\mathrm{bc}}\}, \{\mathrm{a}, \beta_{\mathrm{bc}}\}, \{(a, b), (b, a), (\mathrm{b}, \alpha_{\mathrm{bc}}), (\alpha_{\mathrm{bc}}, \beta_{\mathrm{bc}}), (\beta_{\mathrm{bc}}, \mathrm{c})\}, \emptyset\rangle$, whereas the att-iAF derived from Δ is $att(\Delta) = \Delta'' = \langle\{\mathrm{b}, \mathrm{c}, \mathrm{a}, \alpha\}, \emptyset, \{(\mathrm{a}, \mathrm{b}), (\mathrm{b}, \mathrm{a})\}, \{(\mathrm{b}, \mathrm{c}), (\alpha, \mathrm{a})\}\rangle$. The iAFs $arg(\Delta)$ and $att(\Delta)$ are shown if Fig. 13, where the mapping of uncertain attacks in $arg(\Delta)$ and of uncertain arguments in $att(\Delta)$ is shown in blue. □

The transformations described above to eliminate uncertain attacks/arguments are inspired by those proposed in [43] to eliminate attacks/arguments with probability less than 1 in probabilistic AF. We now introduce a special class of arg-iAFs.

Definition 15 *(farg-iAF).* An arg-iAF $\Delta = \langle A, B, R, \emptyset\rangle$ is said to be *fact-uncertain* (farg-iAF) iff $\forall (a, b) \in R$, $b \notin B$.

Thus, in *farg*-iAFs uncertain arguments are not attacked by other arguments. Given an arg-iAF Δ, $farg(\Delta)$ denotes the farg-iAF derived from Δ as follows: for each uncertain argument b which is attacked in Δ, make b certain and add the

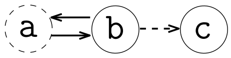

Fig. 12. Incomplete AF Δ of Example 21.

Fig. 13. Arg-iAF $arg(\Delta)$ and att-iAF $att(\Delta)$ for the iAF Δ of Example 21. (For interpretation of the colors in the figure(s), the reader is referred to the web version of this article.)

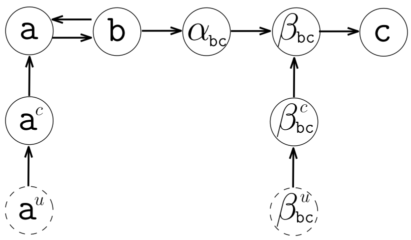

Fig. 14. Farg-iAF $farg(arg(\Delta))$ for the iAF Δ of Example 21.

attacks (b^u, b^c), (b^c, b), where b^c (resp. b^u) is a fresh certain (resp. uncertain) argument. With a little abuse of notation, for any iAF Δ we use $farg(\Delta)$ to denote $farg(arg(\Delta))$.

Example 22. Consider the iAF Δ of Example 21. The derived farg-iAF is $farg(\Delta) = \langle A = \{\mathtt{a}, \mathtt{b}, \mathtt{c}, \mathtt{a}^{\mathtt{c}}, \alpha_{\mathtt{bc}}, \beta_{\mathtt{bc}}, \beta^{\mathtt{c}}_{\mathtt{bc}}\},\ B = \{\mathtt{a}^{\mathtt{u}},\ \beta^{\mathtt{u}}_{\mathtt{bc}}\},\ R = \{(\mathtt{a}, \mathtt{b}), (\mathtt{b}, \mathtt{a}), (\mathtt{a}^{\mathtt{c}}, \mathtt{a}),\ (\mathtt{a}^{\mathtt{u}}, \mathtt{a}^{\mathtt{c}}),\ (\mathtt{b}, \alpha_{\mathtt{bc}}), (\alpha_{\mathtt{bc}}, \beta_{\mathtt{bc}}),\ (\beta_{\mathtt{bc}}, \mathtt{c}),\ (\beta^{\mathtt{c}}_{\mathtt{bc}}, \beta_{\mathtt{bc}}),\ (\beta^{\mathtt{u}}_{\mathtt{bc}}, \beta^{\mathtt{c}}_{\mathtt{bc}})\},\ T = \emptyset\rangle$ is shown in Fig. 14. $\square$

In order to define the relationship between a given iAF Δ and a derived iAF $\varphi(\Delta)$, where $\varphi \in \{arg, att, farg\}$, we first relate the completions of $\varphi(\Delta)$ with the completions of Δ.

Given an iAF $\Delta = \langle A, B, R, T\rangle$, let $\varphi \in \{arg, att, farg\}$, for any $\Lambda' = \langle A', R'\rangle \in comp(\varphi(\Delta))$, $af_\Delta(\Lambda')$ denotes the AF $\Lambda'' = \langle A'', R''\rangle \in comp(\Delta)$ with:

- $A'' = A \cup ((B \cap A') \setminus \{a \mid (\alpha, a) \in R' \vee a^u \notin A'\})$, and
- $R'' = (R \cap (A'' \times A''))\ \cup\ ((T \cap (A'' \times A'')) \setminus \{(a, b) \mid (\beta_{ab} \notin A') \vee (\beta^c_{ab} \in A' \wedge \beta^u_{ab} \notin A')\})$.

Herein, the set $\{a \mid (\alpha, a) \in R' \vee a^u \notin A'\}$ is introduced in the formula to avoid considering arguments either (i) attacked by α in $comp(att(\Delta))$ or (ii) always false in $comp(farg(\Delta))$ as $a^u \notin A'$. Analogously, the set $\{(a, b) \mid (\beta_{ab} \notin A') \vee (\beta^c_{ab} \in A' \wedge \beta^u_{ab} \notin A')\}$ is introduced in the formula to avoid considering uncertain attacks that are chosen to not occur in either (i) $comp(arg(\Delta))$, as $\beta_{ab} \notin A'$, or (ii) $comp(farg(\Delta))$ as $(\beta^c_{ab} \in A' \wedge \beta^u_{ab} \notin A')$.

Example 23. Consider the iAF Δ of Example 21 and $farg(\Delta) = \langle A, B, R, T\rangle$ of Example 22 (see Fig. 14). For $\Lambda = \langle (A \cup B) \setminus \{\beta^u_{bc}\},\ R \setminus \{(\beta^{\mathtt{u}}_{\mathtt{bc}}, \beta^{\mathtt{c}}_{\mathtt{bc}})\}\rangle \in comp(farg(\Delta))$, containing the uncertain argument $\mathtt{a}^{\mathtt{u}}$ but not $\beta^{\mathtt{u}}_{\mathtt{bc}}$, $af_\Delta(\Lambda) = \langle\{\mathtt{a}, \mathtt{b}, \mathtt{c}\}, \{(\mathtt{a}, \mathtt{b}), (\mathtt{b}, \mathtt{a})\}\rangle \in comp(\Delta)$, that is it contains the uncertain argument $\mathtt{a}$, but does not contain the uncertain attack $(\mathtt{b}, \mathtt{c})$. $\square$

Lemma 2. *For any iAF Δ and $\varphi \in \{arg, att, farg\}$, $af_\Delta : comp(\varphi(\Delta)) \rightarrow comp(\Delta)$ is a surjective function.*

The next theorem states the 'equivalence' between iAFs and the iAFs derived by applying the previous mappings.

Theorem 11. *Let $\Delta = \langle A, B, R, T\rangle$ be an iAF, $\sigma \in \{\mathtt{gr}, \mathtt{co}, \mathtt{st}, \mathtt{pr}, \mathtt{sst}\}$, and $\varphi \in \{arg, att, farg\}$ and let $\Lambda \in comp(\Delta)$. Then,*

- *$comp(\Delta) = \{af_\Delta(\Lambda) \mid \Lambda \in comp(\varphi(\Delta))\}$, and*
- *$\sigma(\Lambda) = \{E \cap (A \cup B) \mid \Lambda' \in comp(\varphi(\Delta))\ \wedge\ \Lambda = af_\Delta(\Lambda')\ \wedge\ E \in \sigma(\Lambda')\}$.*

Thus, any iAF Δ is equivalent to an arg-iAF (resp. farg-iAF, att-iAF) Δ' in the sense that there is mapping between the completions of $\varphi(\Delta)$ and the completions of Δ, and for any pair of AFs for which the mapping holds, the two AFs have the same (modulo arguments added in the rewriting) set of σ-extensions, for $\sigma \in \{\mathtt{gr}, \mathtt{co}, \mathtt{st}, \mathtt{pr}, \mathtt{sst}\}$. This result entails that arg-iAFs (resp. farg-iAF, att-iAF) have the same expressivity of general iAFs, though arg-iAFs (resp. farg-iAF, att-iAF) have a simpler structure.

The next example shows that the function af_Δ is not injective.

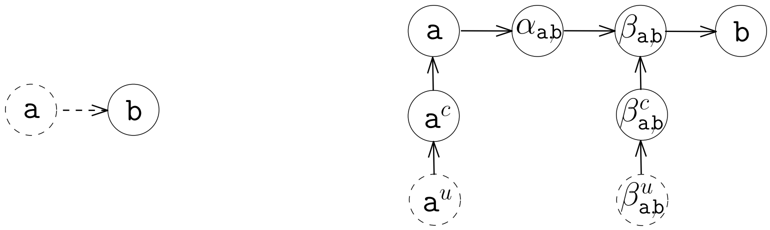

Fig. 15. Incomplete AF of Example 24 (left) and corresponding farg-AF (right).

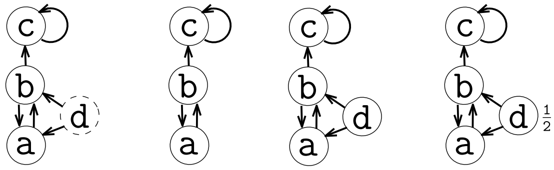

Fig. 16. (From left to right:) iAF Δ of Example 25, its completions Λ_1 and Λ_2, and its derived PrAF Δ^p.

Example 24. Consider the iAF $\Delta = \langle\{\mathtt{b}\}, \{\mathtt{a}\}, \emptyset, \{(\mathtt{a}, \mathtt{b})\}\rangle$ shown in Fig. 15 (left). Δ has three completions: $\Lambda_1 = \langle\{\mathtt{a}, \mathtt{b}\}, \{(\mathtt{a}, \mathtt{b})\}\rangle$, $\Lambda_2 = \langle\{\mathtt{a}, \mathtt{b}\}, \emptyset\rangle$, and $\Lambda_3 = \langle\{\mathtt{b}\}, \emptyset\rangle$.

The derived farg-iAF is $\Delta' = \langle A = \{\mathtt{a}, \mathtt{b}, \alpha_{\mathtt{ab}}, \beta_{\mathtt{ab}}, \mathtt{a}^{\mathtt{c}}, \beta^{\mathtt{c}}_{\mathtt{ab}}\},\ B = \{\mathtt{a}^{\mathtt{u}}, \beta^{\mathtt{u}}_{\mathtt{ab}}\},\ R = \{(\mathtt{a}, \alpha_{\mathtt{ab}}), (\alpha_{\mathtt{ab}}, \beta_{\mathtt{ab}}), (\beta_{\mathtt{ab}}, \mathtt{b}), (\mathtt{a}^{\mathtt{c}}, \mathtt{a}), (\mathtt{a}^{\mathtt{u}}, \mathtt{a}^{\mathtt{c}}), (\beta^{\mathtt{c}}_{\mathtt{ab}}, \beta_{\mathtt{ab}}), (\beta^{\mathtt{u}}_{\mathtt{ab}}, \beta^{\mathtt{c}}_{\mathtt{ab}})\}, T = \emptyset\rangle$ shown in Fig. 15 (right).

Δ' has four completions:

- $\Lambda'_1 = \langle A \cup B, R\rangle$,
- $\Lambda'_2 = \langle A \cup \{a^u\}, R \setminus \{(\beta^u_{ab}, \beta^c_{ab})\}\rangle$,
- $\Lambda'_3 = \langle A \cup \{\beta^u_{a,b}\}, R \setminus \{(a^u, a^c)\}\rangle$,
- $\Lambda'_4 = \langle A, R \setminus \{(a^u, a^c)(\beta^u_{ab}, \beta^c_{ab})\}\rangle$.

Moreover we have that $af_\Delta(\Lambda'_1) = \Lambda_1$, $af_\Delta(\Lambda'_2) = \Lambda_2$ and $af_\Delta(\Lambda'_3) = af_\Delta(\Lambda'_4) = \Lambda_3$. $\square$

7.2. Probabilistic acceptance in iAF

In this section, using the equivalence results given earlier, we investigate the relationships between iAF and PrAF by relating iAF acceptance problems to probabilistic acceptance in PrAF.

We start by defining a PrAF Δ^p encoding an arg-iAF Δ.

Definition 16 *(Derived PrAF)*. Given an arg-iAF $\Delta = \langle A, B, R, \emptyset\rangle$, the PrAF derived from Δ is $\Delta^p = \langle A \cup B, R, P\rangle$, where $P : A \cup B \to \{1/2, 1\}$ with $P(a) = 1$ for $a \in A$ and $P(b) = 1/2$ for $b \in B$.

It is easy to check that, given an arg-iAF $\Delta = \langle A, B, R, \emptyset\rangle$, for every $\Lambda \in pw(\Delta^p)$, $\mathcal{I}(\Lambda)$ is equal to either 0 or $\frac{1}{2^{|B|}}$. As stated next, non-zero probability possible worlds of derived PrAF Δ^p one-to-one correspond to completions of Δ.

Proposition 5. *For any arg-iAF Δ, $comp(\Delta) = \{\Lambda \mid \Lambda \in pw(\Delta^p) \wedge \mathcal{I}(\Lambda) > 0\}$.*

Example 25. Consider the arg-iAF $\Delta = \langle A = \{\mathtt{a}, \mathtt{b}, \mathtt{c}\},\ B = \{\mathtt{d}\},\ R = \{(\mathtt{a}, \mathtt{b}), (\mathtt{b}, \mathtt{a}), (\mathtt{b}, \mathtt{c}), (\mathtt{c}, \mathtt{c}), (\mathtt{d}, \mathtt{a}), (\mathtt{d}, \mathtt{b})\}, \emptyset\rangle$ shown on left hand-side of Fig. 16. The derived PrAF $\Delta^p = \langle A \cup B, R, P\rangle$, with $P(x) = 1$ for $x \in \{\mathtt{a}, \mathtt{b}, \mathtt{c}\}$ and $P(\mathtt{d}) = 1/2$, is shown on the right-hand side of Fig. 16. There are only two possible worlds with probability greater than 0: $\Lambda_1 = \langle A, R \setminus \{(\mathtt{d}, \mathtt{a}), (\mathtt{d}, \mathtt{b})\}\rangle$ and $\Lambda_2 = \langle A \cup B, R\rangle$ with $\mathcal{I}(\Lambda_1) = \mathcal{I}(\Lambda_2) = 1/2$. $\square$

Therefore, for any arg-iAF Δ we take the associated probabilistic AF Δ^p and assume that a PDF over $pw(\Delta^p)$ is given. The probabilistic acceptance of a goal argument g is defined as follows.

Definition 17. Given an arg-iAF $\Delta = \langle A, B, R, \emptyset\rangle$ and an argument $g \in A \cup B$, the probability $PrA^\sigma_\Delta(g)$ that g is acceptable w.r.t. semantics σ is:

$$PrA^\sigma_\Delta(g) = \sum_{\substack{\Lambda \in pw(\Delta^p) \wedge \\ E \in \sigma(\Lambda) \wedge g \in E}} \mathcal{I}(\Lambda) \cdot Pr(E, \Lambda, \sigma),$$

where $Pr(\cdot, \Lambda, \sigma)$ is a PDF over the set $\sigma(\Lambda)$.

Table 5
Explanation-based acceptance probability for arguments of iAF of Example 26.

σ	a	b	c	d
gr	0	0	0	1/2
co	1/6	1/6	0	1/2
st	0	1/2	0	0
pr	1/4	1/4	0	1/2
sst	0	1/2	0	1/2

The explanation-based acceptance probability $PrEA^{\sigma}_{\Delta}(g)$ for a goal argument g is then defined by replacing $Pr(E, \Lambda, \sigma)$ with the explanation-based probability $PrE(E, \Lambda, \sigma)$ associated with σ-extensions (cf. Definition 6).

Example 26. Consider again the iAF of Example 25. There are only two possible worlds Λ_1 and Λ_2 (see Fig. 16) having both probability equal to $1/2$. Λ_1 has 3 complete extensions $E'_1 = \emptyset$, $E''_1 = \{\mathtt{a}\}$ and $E'''_1 = \{\mathtt{b}\}$; E'_1 is grounded, E''_1 and E'''_1 are preferred, and E'''_1 is stable and semi-stable. As for Λ_2, it has only one complete extension $E'_2 = \{\mathtt{d}\}$. With a little effort, it can be checked that, for each possible world Λ_i (with $i \in \{1, 2\}$) and semantics σ, the explanation-based probability associated with extensions $PrE(\cdot, \Lambda, \sigma)$ follows a uniform distribution. For instance, under preferred semantics, $PrE(E''_1, \Lambda_1, \mathtt{pr}) = PrE(E''_2, \Lambda_1, \mathtt{pr}) = 1/2$ and $PrE(E'_2, \Lambda_2, \mathtt{pr}) = 1$. Therefore, $PrEA^{\mathtt{pr}}_{\Delta}(\mathtt{a}) = 1/2 \times 1/2 = 1/4$ since $\mathtt{a}$ is only in one of the two preferred extensions of Λ_1. The explanation-based acceptance probability for all arguments in the iAF is reported in Table 5. □

Given an arg-iAF Δ and an argument g, the problem of computing the value $PrA^{\sigma}_{\Delta}(g)$ (under a given semantics σ) is denoted by PrA[σ], or simply PrA whenever σ is understood. We recall that PrA[σ] is defined after choosing an arbitrary but fixed PDF over the set of extensions of the possible worlds of the derived PrAF Δ^p (cf. Definition 17). The problem of computing the value $PrEA^{\sigma}_{\Delta}(g)$ is denoted by PrEA[σ]; in this case the PDF $PrE(\cdot, \Lambda, \sigma)$ of Definition 6 is used.

As stated next, the results we have for PrAF carry over to iAF. In particular, PrA[σ] is $FP^{\#P}$-hard, regardless of the chosen PDF and semantics σ.

Corollary 5. *For $\sigma \in \{\mathtt{gr}, \mathtt{co}, \mathtt{pr}, \mathtt{st}, \mathtt{sst}\}$, PrA[$\sigma$] is $FP^{\#P}$-hard, even for acyclic arg-iAF and for any chosen PDF.*

We also have the following inapproximability results for PrA[σ] (and thus for PrEA[σ]) under FPRAS and FPARAS schemes.

Corollary 6. *Consider a semantics $\sigma \in \{\mathtt{gr}, \mathtt{co}, \mathtt{pr}, \mathtt{st}, \mathtt{sst}\}$. Unless NP ⊆ BPP, there is no FPRAS for PrA[σ], even for acyclic arg-iAFs and for any chosen PDF.*

Corollary 7. *Let $\sigma \in \{\mathtt{pr}, \mathtt{st}, \mathtt{sst}\}$. Unless NP ⊆ BPP, there is no FPARAS for PrA[σ], for any chosen PDF.*

Notably, we have the following positive result concerning the computation of explanation-based acceptance probability in iAFs.

Corollary 8. *Problem PrEA[σ], with $\sigma \in \{\mathtt{gr}, \mathtt{co}, \mathtt{pr}, \mathtt{st}, \mathtt{sst}\}$, has an FPARAS if either i) $\sigma = \mathtt{gr}$, or ii) the input arg-iAF has no odd cycles.*

We conclude this section with the following propositions that highlight some relationships between iAFs and PrAFs, assuming that the uniform PDF over the set of extensions is used.

Proposition 6. *For any arg-iAF Δ and argument goal g, we have that:*

- *$PCA_{\sigma}(\Delta, g)$ is false iff $PrA^{\sigma}_{\Delta^p}(g) = 0$;*
- *$NSA_{\sigma}(\Delta, g)$ is true iff $PrA^{\sigma}_{\Delta^p}(g) = 1$.*

The connection between iAF and PrAF is also investigated in [55], where the PrAF associated to an iAF assigns a probability in $(0, 1)$ to each uncertain argument. Moreover, PCA, PSA, NCA, and NSA are related to the concept of *probabilistic credulous/skeptical acceptance* [33], which is different from that of probabilistic acceptance of Definition 3. Indeed, the probability that an argument g is credulously (resp. skeptically) accepted is the sum of the probabilities of the possible worlds where g is credulously (resp. skeptically) accepted, according to a given semantics σ. Hence, the probability of a world Λ is added to the summation iff g belongs to at least one (resp. every) σ-extension of Λ. In contrast, with the aim of offering a more granular approach, Definition 3 uses the probabilities assigned to σ-extensions by the PDF $Pr(\cdot, \Lambda, \sigma)$. For instance, taking the PrAF Δ^p of Example 25 (see Fig. 16), under the complete semantics the probabilistic skeptical (resp. credulous) acceptance of b is 0 (resp. 1/2), while $PrA^{\mathtt{co}}_{\Delta^p}(\mathtt{b}) = 1/6$ as b belongs to one of three extensions of one of the two worlds (cf. Table 5).

Although the conditions of Proposition 6 are similar to those identified in [55], they refer to different notions of probabilistic acceptance. In fact, while probabilistic skeptical and credulous acceptance define an interval, Definition 3 provides a precise value in that interval.

The following proposition considers acyclic arg-iAFs, a subclass of odd-cycle free iAFs.

Proposition 7. *Let $\Delta = \langle A, B, R, \emptyset\rangle$ be an acyclic arg-iAF and g a goal. It holds that:*

- $PSA_\sigma(\Delta, g) \equiv PCA_\sigma(\Delta, g)$ *and* $NSA_\sigma(\Delta, g) \equiv NCA_\sigma(\Delta, g)$*;*
- $PSA_\sigma(\Delta, g)$ *is true iff* $PrA^\sigma_{\Delta^p}(g) \geq \frac{1}{2^{|B|}}$*;*
- $NSA_\sigma(\Delta, g)$ *is false iff* $PrA^\sigma_{\Delta^p}(g) \leq 1 - \frac{1}{2^{|B|}}$.

It is worth noting that, although we transformed an iAF into a PrAF by assigning to each uncertain argument a probability equal to $\frac{1}{2}$ (cf. Definition 16), we could map—with no impact on our complexity results—an iAF into an EPrAF by assigning to each uncertain argument a probability interval $(\frac{1}{2} - \delta, \frac{1}{2} + \delta)$ with $\delta \in [0, \frac{1}{2}]$, that is by assigning to each uncertain argument any probability interval contained in $(0, 1)$.

8. Discussion and conclusions

We have explored the problem of computing the probability of acceptance of a goal argument in probabilistic argumentation frameworks. Our approach stems from the fact that, in our view, probabilistic credulous acceptance may not provide intuitive answers as it generalizes the classical credulous acceptance problem for AFs in one dimension only, that is, via probabilities over possible worlds. Our approach also considers another dimension, i.e. it also assigns probabilities to the extensions of each possible world, by exploiting dependencies among arguments (i.e. dependencies among SCCs).[5] As shown in our running example, this enables more intuitive answers w.r.t. probabilistic credulous or skeptical acceptance (e.g. the probabilities of acceptance of mutually conflicting arguments such as fish and meat sum up to 1, but this is not the case for probabilistic credulous or skeptical acceptance), as well as probabilistic answers under uniform distribution. Thus, we introduced the problem PrA[σ], where a PDF is assumed over the set of extensions, and devised PrEA[σ] as a concrete instance, where the PDF leverages our notion of explanations for extensions.

An alternative definition for explaining complete extensions has been recently proposed in [45]. It exploits the concept of reduct, i.e. a sub-framework obtained by removing true and false arguments w.r.t. a complete extension. Intuitively, given an AF Λ, $(\mathtt{gr}(\Lambda), S_1, S_2)$ is a (successful) *explanation scheme* for a complete extension $E \in \mathtt{co}(\Lambda)$ if $E = \mathtt{gr}(\Lambda) \cup S_1 \cup S_2$, where S_1 is a conflict-free set of arguments in $\widehat{\Lambda} = \langle \widehat{A}, \widehat{\Sigma}\rangle$ appearing in even cycles of $\widehat{\Lambda}$, $S_2 = \mathtt{gr}\big(\widehat{\Lambda}\downarrow_{\widehat{A}\setminus S_1^*}\big)$, and $\mathtt{gr}(\Lambda) \cup S_1 \cup S_2$ defends S_1—we refer the reader to [45] for further details. Although our approach (which is based on the one introduced in [1]) and that proposed in [45] share the same underlying idea of applying the grounded semantics in a step-wise fashion as well as choosing from arguments appearing in even cycles, the two approaches may yield different explanations as illustrated in the following example. Consider the AF Λ obtained from that of Example 5 (i.e. the deterministic version of the PrAF of Example 1) by adding the argument cake and an attack from white to cake. For the complete extension $E = \{\mathtt{meat}, \mathtt{red}, \mathtt{cake}\}$, we have that $X = \langle \mathtt{meat}\rangle$ is the only explanation for E (note that $\langle \mathtt{meat}, \mathtt{red}\rangle$ is not an explanation for E), while $(\emptyset, \{\mathtt{meat}, \mathtt{red}\}, \{\mathtt{cake}\})$ is an explanation scheme for E.

Another important difference with respect to the approach proposed in [45] is that, given a complete extension E for an AF Λ, by applying Definition 4, we are able to build an explanation X for E in polynomial time, while the approach in [45] explicitly requires guessing the set S_1 for the explanation scheme $(\mathtt{gr}(\Lambda), S_1, S_2)$ for E. Moreover, in our context, explanations are sequences of arguments that allow us to assign probabilistic values to extensions.

The concept of *strong explanation* is proposed in [37], inspired by the related notions introduced in [38,63,39]. Intuitively, given an AF $\Lambda = \langle A, \Sigma\rangle$, a set S of arguments is a strong explanation for a set E of arguments if, for each AF $\Lambda' = \Lambda\downarrow_B$ with $S \subseteq B \subseteq A$, it holds that $E \subseteq E'$ with $E' \in \sigma(\Lambda')$. Again, our approach and that in [37] may yield different explanations as illustrated in the following example. Consider the AF $\Lambda = \langle A, \Sigma\rangle$ with $A = \{\mathtt{a}, \mathtt{b}, \mathtt{c}, \mathtt{d}\}$, $\Sigma = \{(\mathtt{a}, \mathtt{b}), (\mathtt{b}, \mathtt{a}), (\mathtt{a}, \mathtt{c}), (\mathtt{b}, \mathtt{c}), (\mathtt{c}, \mathtt{d})\}$. Suppose we want to explain the preferred (and stable) extension $E = \{\mathtt{a}, \mathtt{d}\}$. Then, $S = \{\mathtt{a}\}$ is not a strong explanation for E, as for $\Lambda' = \Lambda\downarrow_{\{\mathtt{a}\}}$ we have a unique preferred extension $E' = \{\mathtt{a}\}$ and $E \not\subseteq E'$. In contrast, in our case, $\langle \mathtt{a}\rangle$ is the only explanation for extension E.

Integrating explanations in argumentation systems is important for enhancing the argumentation and persuasion capabilities of software agents [34,64,65,35]. For these reasons, several researchers explored how to deal with explanations in

[5] SCCs have been extensively investigated in argumentation since they are inherently related to computational aspects of AFs (e.g. the SCC-decomposability principle, allowing to evaluate the status of arguments in any SCC independently of that of the attackers [59–62]).

formal argumentation. Significant work in this field includes [36], where a new argumentation semantics is proposed for capturing explanations in AF, and [66] that focuses on ABA frameworks [66–68]. They treat an explanation as a semantics to answer why an argument is accepted or not. Thus, an explanation is viewed as a set of arguments, instead of a sequence of arguments, needed for explaining such an extension. In [36] an explanation is as a set of arguments justifying a given argument by means of a proponent-opponent dispute-tree [69]. A similar approach based on debate trees as proof procedure for computing grounded, ideal, and preferred semantics, has been proposed in [70]. However, in our perspective, explanations provide a tool to assign probabilities to extensions, and an explanation can be viewed as a sequence of choices to be made to justify how an extension is obtained.

Analogously to several other computational approaches in formal argumentation, our approach suffers from high computational complexity [71–74]. However, after showing that $\mathsf{PrA}[\sigma]$ and $\mathsf{PrEA}[\sigma]$ are $FP^{\#P}$-hard, even for acyclic PrAFs, we investigated the existence of polynomial-time algorithms for $\mathsf{PrEA}[\sigma]$ in terms of approximate computation via FP(A)RASes. This is analogous to what is done in [25,75], where Monte-Carlo techniques are proposed to estimate the probability that a set of arguments is an extension in PrAF as well as in a form of a structured argumentation framework.

We also found that approximate computation via FPARAS is not possible for general PrAFs and for all the considered semantics, besides the grounded. Thus, we proposed an additive error approximation algorithm for solving $\mathsf{PrEA}[\sigma]$ in the cases of probabilistic AFs without odd-length cycles and any semantics σ, and for $\mathsf{PrEA}[\mathtt{gr}]$ in general PrAFs. Our results immediately apply to probabilistic frameworks with uncertain attacks, thanks to the results of [43]. Moreover, we investigated the approximate complexity of credulous and skeptical acceptance problems for PrAFs, and related it to that of explanation-based probabilistic acceptance. Finally, we introduced the concept of explanation-based probabilistic acceptance for iAF, extending our proposal for PrAF and the complexity and approximation results to the case of iAF.

To the best of our knowledge, this is the first piece of work investigating probabilistic acceptance in combination with explanations for probabilistic AFs (and incomplete AFs). As a first direction for future work, we plan to extend our notion of explanation, and investigate the counterparts of our problems, in other contexts such as weighted AF [76,77] and structured argumentation [78–80]. In particular, it would be interesting to consider for instance p-ASPIC [20] and Probabilistic Assumption-Based Argumentation (PABA) [19,66], the probabilistic versions of ASPIC [81] and ABA [67,82], respectively. To this end, exploiting the fact that (flat) ABA admit AF as an instance [83], it would be interesting to investigate how to extend the notion of Explanation-based Probabilistic Acceptance to (fragments of) PABA by elaborating on the ideas underlying our definition of explanation. The complexity of credulous and skeptical acceptance problems in PABA, along with that of a novel decision problem concerning *strong acceptance* with probability 1, has been investigated in [84]. Credulous (resp. skeptical) strong acceptance is the problem of checking whether the probabilistic credulous (resp. skeptical) acceptance of a given goal literal is equal to 1. In our setting, we could define the explanation-based strong acceptance problem as the problem of checking whether the explanation-based probabilistic acceptance of a given goal argument g is equal to 1, that is, given a PrAF Δ, a semantics σ, and an argument g, deciding whether $PrEA^{\sigma}_{\Delta}(g)$ is equal to 1. It is worth noting that deciding whether $PrEA^{\sigma}_{\Delta}(g) = 1$ corresponds to check whether g is skeptically accepted in every possible world of Δ, that is checking whether $PrSA^{\sigma}_{\Delta}(g) = 1$. That is, in our setting, deciding the explanation-based strong acceptance problem is equivalent to deciding the strong skeptical acceptance problem.

As a second direction, we plan to investigate other ways of defining a PDF over the set of extensions that enable other instantiations of $\mathsf{PrA}[\sigma]$, not necessarily defined using our concept of explanations. Given the strict relationship between abstract argumentation semantics and partial stable models of logic programs [85], we believe the concepts in this work could be applicable to similar probabilistic approaches for other KR formalisms, such as the well-know ProbLog System, a probabilistic version of ProLog [86–88]—this is another interesting direction that we plan to address for future work.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

No data was used for the research described in the article.

Acknowledgements

The authors wish to thank the anonymous referees for providing detailed comments and suggestions that helped to substantially improve the paper. Moreover, the authors acknowledge the support of the PNRR project FAIR - Future AI Research (PE00000013), Spoke 9 - Green-aware AI, under the NRRP MUR program funded by the NextGenerationEU. Finally, this work was also funded by the Next Generation EU - Italian NRRP, Mission 4, Component 2, Investment 1.5, call for the creation and strengthening of 'Innovation Ecosystems', building 'Territorial R&D Leaders' (Directorial Decree n. 2021/3277) - project Tech4You - Technologies for climate change adaptation and quality of life improvement, n. ECS0000009. This work reflects only the authors' views and opinions, neither the Ministry for University and Research nor the European Commission can be considered responsible for them.

Appendix A. Proofs

In this appendix we provide the proofs of the results stated in the core of the paper.

To ease readability, we restate the results and organize them in sections by following the same structure as in the paper.

A.1. Explanations

Proposition 1. *Let $\Lambda = \langle A, \Sigma\rangle$ be an AF, $\sigma \in \{\mathtt{co}, \mathtt{gr}, \mathtt{pr}, \mathtt{st}, \mathtt{sst}\}$ a semantics, $E \in \sigma(\Lambda)$ an extension, $a \in E$ an argument and $G_a = \mathtt{gr}(\Lambda_a)$, Then, $E' = E \setminus G_a$ is a σ-extension for Λ^*_a.*

Proof. As the set of complete extensions forms a complete-meet semilattice (cf. Theorem 25 of [13]), w.l.o.g. let $\sigma = \mathtt{co}$. Recall that $E \in \mathtt{co}(\Lambda)$ iff $E = Acc(E)$ and E is conflict-free.

Let $\Lambda' = \langle A', \Sigma'\rangle = \Lambda^*_a = \Lambda\downarrow_{A\setminus G^*_a}$, it is sufficient to prove that $E' = E \setminus G_a = Acc(E')$ as E' is clearly conflict-free as it is a subset of a conflict-free set. Observe that $A' \subseteq A$ and $\Sigma' \subseteq \Sigma$; moreover, $E' \subseteq E$ and $Def(E') = E'^+ \subseteq Def(E) = E^+$. Reasoning by contradiction, suppose that $E' \subset E')$. Thus there exists an argument $z \in Acc(E') \setminus E'$, that is, $z \in \{A' \setminus E' \mid \forall b \in A'.(b,z) \in \Sigma' \implies b \in Def(E')\}$. This implies that $z \in \{A \setminus E \mid \forall (b,z) \in \Sigma \implies b \in Def(E)\}$, that is $z \in Acc(E) \setminus E$, contradiction (E is a complete extension of Λ).

Suppose now that $E' \supset Acc(E')$. Thus there exists $z \in E'$ and $\exists (b,z) \in \Sigma'$ s.t. $b \notin Def(E')$. As $Def(E) \cap A' = Def(E') \cap A'$ it also holds that $b \notin Def(E)$, implying that $z \in E \setminus Acc(E)$, contradiction. $\square$

Proposition 2. *Let Λ be an AF and $\sigma \in \{\mathtt{gr}, \mathtt{co}, \mathtt{pr}, \mathtt{st}, \mathtt{sst}\}$ a semantics. Then:*

i) for every $E \in \sigma(\Lambda)$, $Exp^\sigma_\Lambda(E) \neq \emptyset$;
ii) for every $E_i, E_j \in \sigma(\Lambda)$ with $E_i \neq E_j$, $Exp^\sigma_\Lambda(E_i) \cap Exp^\sigma_\Lambda(E_j) = \emptyset$.

Proof. Item *i*) follows from Definition 4, where an explanation is always determined for each extension.

Item *ii*), reasoning by contradiction, assume that $Exp^\sigma_\Lambda(E_i) \cap Exp^\sigma_\Lambda(E_j) \neq \emptyset$, and let X be an explanation in $Exp^\sigma_\Lambda(E_i) \cap Exp^\sigma_\Lambda(E_j)$, and $\widetilde{X} = set(X) \cap A$, where A is the set of arguments of Λ. Thus, from Theorem 1 (see below), we have that $E_i = \mathtt{gr}(\Lambda_{\widetilde{X}})$ and $E_j = \mathtt{gr}(\Lambda_{\widetilde{X}})$, that implies $E_i = E_j$ (contradiction). $\square$

Theorem 1. *Let $\Lambda = \langle A, \Sigma\rangle$ be an AF, σ a semantics in $\{\mathtt{gr}, \mathtt{co}, \mathtt{pr}, \mathtt{st}, \mathtt{sst}\}$ and E a σ-extension. Then, for any $X \in Exp^\sigma_\Lambda(E)$ and $\widetilde{X} = set(X) \cap A$ we have that $E = \mathtt{gr}(\Lambda_{\widetilde{X}})$ and $\widetilde{X}^- \subseteq E^+$, where $\Lambda_{\widetilde{X}}$ is the AF derived from Λ by deleting attacks to arguments in $\widetilde{X}$.*

Proof. We start by showing that $E = \mathtt{gr}(\Lambda_{\widetilde{X}})$. To prove that $E = \mathtt{gr}(\Lambda_{\widetilde{X}})$, we show that the following two items hold.

1. $E = \mathtt{gr}(\Lambda_E)$. We have that $E \subseteq \mathtt{gr}(\Lambda_E)$ as arguments in E are not attacked in Λ_E and thus they are in $\mathtt{gr}(\Lambda_E)$. To show that $\mathtt{gr}(\Lambda_E) \subseteq E$, it is sufficient to prove that $\mathtt{gr}(\Lambda\downarrow_{A\setminus(E\cup E^+)}) = \emptyset$. Arguments in $\Lambda\downarrow_{A\setminus(E\cup E^+)}$ are such that all attacks towards them are from undecided arguments w.r.t. E (i.e. they are neither accepted nor defeated w.r.t. E), as they are attacked from arguments in E^+ or in $A \setminus (E \cup E^+)$ in Λ. The first kind of attacks can be removed as they are irrelevant [89,90]. As for the second kind of attacks, since each argument a in $\Lambda\downarrow_{A\setminus(E\cup E^+)}$ is undecided w.r.t. E, there is at least an argument b of $\Lambda\downarrow_{A\setminus(E\cup E^+)}$ that is in turn undecided w.r.t. E and that attacks a. Thus, the grounded extension of $\Lambda\downarrow_{A\setminus(E\cup E^+)}$ is empty.
2. $\mathtt{gr}(\Lambda_E) = \mathtt{gr}(\Lambda_{\widetilde{X}})$. We first show that $\mathtt{gr}(\Lambda_{\widetilde{X}}) \subseteq \mathtt{gr}(\Lambda_E)$. Let $\Lambda_{\widetilde{X}} = \langle A_{\widetilde{X}}, \Sigma_{\widetilde{X}}\rangle$ and $\Lambda_E = \langle A_E, \Sigma_E\rangle$. Reasoning by contradiction, assume that $\mathtt{gr}(\Lambda_{\widetilde{X}}) \nsubseteq \mathtt{gr}(\Lambda_E)$, that is there exists an argument $z \in \mathtt{gr}(\Lambda_{\widetilde{X}})$ s.t. $z \notin \mathtt{gr}(\Lambda_E)$. Thus, there is an attack $(z', z) \in \Sigma_E$ s.t. $z' \notin E^+$. As $\Sigma_E \subseteq \Sigma_{\widetilde{X}}$ and $\widetilde{X}^+ \subseteq E^+$, we have that $(z', z) \in \Sigma_{\widetilde{X}}$ and $z' \notin \widetilde{X}^+$, meaning that $z \notin \mathtt{gr}(\Lambda_{\widetilde{X}})$, contradiction. We now show that $\mathtt{gr}(\Lambda_E) \subseteq \mathtt{gr}(\Lambda_{\widetilde{X}})$. Reasoning by contradiction, assume $\mathtt{gr}(\Lambda_E) \nsubseteq \mathtt{gr}(\Lambda_{\widetilde{X}})$, that is there exists an argument $z \in \mathtt{gr}(\Lambda_E)$ s.t. $z \notin \mathtt{gr}(\Lambda_{\widetilde{X}})$. Thus, z is such that $\forall (z', z) \in \Sigma_E \Rightarrow z' \in E^+$. Moreover, we have that $\exists (z'', z) \in \Sigma_{\widetilde{X}}$ s.t. $z'' \notin \widetilde{X}^+$. Since $\Sigma_{\widetilde{X}} \subseteq \Sigma$ and $\widetilde{X}^+ \subseteq E^+$, we have that $\exists (z'', z) \in \Sigma_{\widetilde{X}}\ s.t.\ z'' \notin \widetilde{X}^+ \ \Rightarrow\ \exists (z'', z) \in \Sigma\ s.t.\ z'' \notin E^+$ and thus $z \notin E$. Since we have shown that $E = \mathtt{gr}(\Lambda_E)$, $z \notin \mathtt{gr}(\Lambda_E)$, contradiction.

Finally, we show that $\widetilde{X}^- \subseteq E^+$. Let $X = \langle a_1, \ldots, a_n\rangle$ and $G_{a_i} = \mathtt{gr}(\Lambda_i)$ where Λ_i is the AF obtained at step i of Definition 4. Observe that, arguments in $\widetilde{X}^-$ appear in either $G^+_{a_i}$ or in Ω, that in turn consists of arguments that are 'labeled' as false in the steps $j > i$. As Definition 4 requires that at the last step of the process (i.e. Item 1) it holds that $\Omega = \emptyset$, we obtain that $\widetilde{X}^- \subseteq E^+$. $\square$

Proposition 3. *Let $\Lambda = \langle A, \Sigma\rangle$ be an AF, σ a semantics in $\{\mathtt{gr}, \mathtt{co}, \mathtt{pr}, \mathtt{st}, \mathtt{sst}\}$ and E a σ-extension. Let $\mathcal{C}$ and $\mathcal{D}$ be two linear orderings of the SCCs of Λ (according to the topological ordering of the graph representing the AF Λ). Let $PrE_{\mathcal{O}}(E, \Lambda, \sigma)$ be the probability associated with extension E under a linear ordering $\mathcal{O}$. Then, it holds that $PrE_{\mathcal{C}}(E, \Lambda, \sigma) = PrE_{\mathcal{D}}(E, \Lambda, \sigma)$.*

Proof. Let $\mathcal{C} = \langle \mathcal{C}_1, \dots, \mathcal{C}_n \rangle$ and $\mathcal{D} = \langle \mathcal{D}_1, \dots, \mathcal{D}_n \rangle$ be two linear orderings of the SCCs of Λ. Let $X = \langle x_1^1, \dots, x_1^{k_1}, x_2^1, \dots, x_2^{k_2}, \dots, x_n^1, \dots, x_n^{k_n} \rangle$ be an explanation for E w.r.t. $\mathcal{C}$ such that:

- $\bigcup_{j=1}^{k_j} x_i^j \subseteq \{\varepsilon\} \cup \mathcal{C}_i$ for any $i \in [1, n]$; and
- $\pi(x_i^j) = \pi(\text{parent}(x_i^j)) \cdot \frac{1}{p_i^j}$ where $p_i^j = |\text{children}(\text{parent}(x_i^j))|$.

We now prove that there exists an explanation $Y = \langle y_1^1, \dots, y_1^{l_1}, y_2^1, \dots, y_2^{l_2}, \dots, y_n^1, \dots, y_n^{l_n} \rangle$ for E w.r.t. $\mathcal{D}$ such that Y (interpreted as a word) is an anagram of X and $\pi(X) = \pi(Y)$.

First observe that, the construction of an explanation according to Definition 4 proceeds component by component and it cannot be the case that explanations contain any sequence of elements $\langle x_1, x_2 \rangle$ such that x_1 and x_2 are respectively choices made w.r.t. the components $\mathcal{C}_i$ and $\mathcal{C}_j$ with $j < i$.

For any i-th component ($i \in [1, n]$), if $\mathcal{C}_i = \mathcal{D}_i$ then $X[i] = \langle x_i^1, \dots, x_i^{k_i} \rangle = Y[i] = \langle y_i^1, \dots, y_i^{l_i} \rangle$ and $p(x_i^j) = p(y_i^j)$ for any $j \in [1, k_i]$, as Definition 4 is deterministic, in the sense that all possible explanation choices appear in the trie. Otherwise, $\mathcal{C}_i \neq \mathcal{D}_i$, meaning that there is no topological ordering between $\mathcal{C}_i$ and $\mathcal{D}_i$; in such a case, there must exist $m \neq i$ and $Y[m] = \langle y_m^1, \dots, y_m^{l_m} \rangle$ such that $\mathcal{C}_i = \mathcal{D}_m$, $X[i] = \langle x_i^1, \dots, x_i^{k_i} \rangle = Y[m] = \langle y_m^1, \dots, y_m^{l_m} \rangle$ and $\pi(x_i^j) = \pi(y_m^j)$ for any $j \in [1, k_i]$. We show that this condition holds by reasoning by contradiction. Assume that there exists no such $Y[m]$ such that $X[i] = Y[m]$.

If $m > i$, then there exist either a) at least one x_i^j in $X[i]$ which is not in $Y[m]$, or b) at least one y_m^h in $Y[m]$ which is not in $X[i]$. Let us first focus on case a); as shown below, in case b) we can reason analogously. If x_i^j does not belong to $Y[m]$, then there must exists a previous step in the process defined by Definition 4 that refers to a component $\mathcal{D}_k$ with $k < m$ that computes the status of x_i^j, in the sense that there exists some element y_k^z in $Y[k]$ such that $x_i^j \in G^*(\Lambda_{y_k^z})$. This implies that the status of x_i^j appearing in $\mathcal{C}_i = \mathcal{D}_m$ depends on that of y_k^z. As $\mathcal{D}_k$ precedes $\mathcal{D}_m$ and $\mathcal{D}_m = \mathcal{C}_i$, it follows that there must exist a component $\mathcal{C}_l = \mathcal{D}_k$ with $l < i$, implying that x_i^j cannot be part of $X[i]$; contradiction. In case b), i.e. y_m^h does not belong to $X[i]$, there must exists a previous step in the process underlying Definition 4 that refers to a component $\mathcal{C}_k$ with $k < i$ that computes the status of y_m^h, and thus there exists some element $x_k^z \in X[k]$ such that $y_m^h \in G^*(\Lambda_{x_k^z})$. This implies that the status of y_m^h appearing in $\mathcal{C}_i = \mathcal{D}_m$ depends on that of x_k^z. As $\mathcal{C}_k$ precedes $\mathcal{C}_i$ and $\mathcal{D}_m = \mathcal{C}_i$, we obtain that there must exist a component $\mathcal{D}_l = \mathcal{C}_k$ with $l < m$, implying that y_m^h cannot be part of $Y[m]$, which is a contradiction.

Clearly, if $m < i$, we can reason analogously to what is done above, as it suffice to swap the two components considered and make analogous considerations.

As what is shown above holds for each i-th component (with $i \in [1, n]$), we have that X and Y (interpreted as words) are anagrams. Finally, as this holds for any explanation of E, it follows that $\pi(X) = \pi(Y)$ and thus $PrE_{\mathcal{C}}(E, \Lambda, \sigma) = PrE_{\mathcal{D}}(E, \Lambda, \sigma)$. $\square$

A.2. Exact and approximate complexity

Theorem 9. *For $\sigma \in \{\mathtt{gr}, \mathtt{co}, \mathtt{pr}, \mathtt{st}, \mathtt{sst}\}$, $\mathsf{PrA}[\sigma]$ is $FP^{\#P}$-hard, even for acyclic PrAFs and for any chosen PDF.*

Proof. We show a reduction to our problem from the #P-hard problem #P2CNF [91], that is, the problem of counting the number of satisfying assignments of a CNF formula where each clause consists of exactly 2 positive literals. Since a problem is $\text{FP}^{\#\text{P}}$-hard iff it is #P-hard, this suffices to prove the statement.

Let $\phi = C_1 \wedge C_2 \wedge \dots \wedge C_k$ be a P2CNF, where $X = \{x_1, \dots, x_n\}$ is the set of its propositional variables. We define an (acyclic) PrAF $\Delta = \langle A, \Sigma, P \rangle$ as follows:

- The set A consists of: i) an argument a_i for each propositional variable $x_i \in X$; ii) an argument c_i for each clause C_i appearing in ϕ; and iii) an argument φ;
- Σ contains, for each clause $C_i = x_j \vee x_\ell$ (with $i \in \{1, \dots, k\}$), an attack (c_i, φ), and two attacks (a_j, c_i) and (a_ℓ, c_i).
- Function P assigns probability $\frac{1}{2}$ to every argument corresponding to a propositional variable (i.e. $\forall i \in \{1, \dots, n\}$, $P(a_i) = \frac{1}{2}$), and probability 1 to all the other arguments (i.e. $\forall i \in \{1, \dots, k\}$, $P(c_i) = 1$, and $P(\varphi) = 1$).

Finally, we let the goal argument be φ.

We first show that there is a bijection $\beta : \mathcal{T} \to pw(\Delta)$ between the set $\mathcal{T}$ of truth assignments of P2CNF ϕ and the set $pw(\Delta)$ of possible worlds of Δ. Function β is such that, given a truth assignment τ for the propositional variables of ϕ, the possible world $w = \beta(\tau) = \langle A_w, \Sigma_w \rangle$ is an AF such that $A_w = A \setminus \{a_j \in A \mid \tau(x_j) = false\}$ and $\Sigma_w = \Sigma \setminus \{(a_j, c_i) \in \Sigma \mid \tau(a_j) = false\}$.

We now show that for every truth assignment τ for the propositional variables of ϕ, ϕ evaluates to *true* under τ iff the goal argument φ belongs to the grounded extension of the AF $w = \beta(\tau)$. Indeed, if ϕ evaluates to *true* under τ, then the set $E=\{\varphi\}\cup\{a_j \mid \tau(x_j)=true\}$ is the grounded extension of $w=\beta(\tau)$ because it is conflict-free, it is admissible (for each c_i attacking φ, there is $a_j \in E$ attacking c_i), and contains all the arguments it defends. On the other hand, if φ belongs to the grounded extension of $w=\beta(\tau)$ then for each c_i attacking φ, there must be $a_j \in E$ attacking c_i, meaning that at least one variable per clause is assigned to *true* by τ.

Since the PrAF $\Delta=\langle A,\Sigma,P\rangle$ is acyclic, then every AF $w \in pw(\Delta)$ is acyclic. Thus the grounded extension of $w \in pw(\Delta)$ is the unique extension of w under any semantics in $\sigma \in \{\texttt{gr},\texttt{co},\texttt{pr},\texttt{st},\texttt{sst}\}$. Given $w \in pw(\Delta)$, if extension E is the unique σ-extension of w, then for any PDF $\Pr(\cdot,w,\sigma)$, $Pr(E,w,\sigma)=1$.

Therefore, $PrA^{\sigma}_{\Delta}(\varphi)$ is the sum of the probabilities $\mathcal{I}(w)$ of the possible worlds $w \in pw(\Delta)$ where φ belongs to the unique σ-extension of w, that is the sum of the probabilities of $w \in pw(\Delta)$ whose corresponding truth assignment $\tau=\beta^{-1}(w)$ makes ϕ *true*. Since the probability of every possible world $w \in pw(\Delta)$ is equal to $\frac{1}{2^n}$ (there are n arguments in A having probability equal to $\frac{1}{2}$, and all the other arguments are assigned probability equal to 1), it is the case that $2^n \cdot PrA^{\sigma}_{\Delta}(\varphi)$ is the number of satisfying assignments of ϕ, which suffices to complete the proof. $\square$

Theorem 3. *For $\sigma \in \{\texttt{co},\texttt{pr},\texttt{st},\texttt{sst}\}$,* PrEA[$\sigma$] *is $FP^{\#P}$-hard for AFs (that is, for PrAFs where all probabilities are set to* 1*).*

Proof. We show a reduction to our problem from #P2CNF [91]. Let $\phi=C_1\wedge C_2\wedge\cdots\wedge C_k$ be a P2CNF, where $X=\{x_1,\ldots,x_n\}$ is the set of its propositional variables. We define a PrAF $\Delta=\langle A,\Sigma,P\rangle$ as follows:

- The set A consists of: i) a pair of arguments, namely a_i and $\overline{a}_i$, for each propositional variable $x_i \in X$; ii) an argument c_i for each clause C_i appearing in ϕ; and iii) an argument φ;
- Σ consists of the following attacks. For each clause $C_i=x_j\vee x_\ell$ (with $i\in\{1,\ldots,k\}$), an attack (c_i,φ) as well as attacks (a_j,c_i) and (a_ℓ,c_i) are in Σ; moreover, for each $j\in\{1,\ldots,n\}$, Σ contains a pair of mutual attacks, $(a_j,\overline{a}_j)$ and $(\overline{a}_j,a_j)$.
- Function P assigns probability equal to 1 to every argument in A.

Finally, we let the goal argument be φ.

Observe that Δ is an AF containing n even cycles, one for each propositional variable in X. Moreover, Δ is coherent [48], that is, $\texttt{st}(\Lambda)=\texttt{sst}(\Lambda)=\texttt{pr}(\Lambda)$ where $\Lambda=\langle A,\Sigma\rangle$ is the unique world of Δ. Thus, in the following, we focus on the stable semantics only.

We can show that there is a bijection $\beta:\mathcal{T}\to\texttt{st}(\Lambda)$ between the set $\mathcal{T}$ of truth assignments of ϕ and the set $\texttt{st}(\Lambda)$ of stable extensions of Λ. In particular, function β is such that, given a truth assignment τ for the propositional variables of ϕ, the set $S=\beta(\tau)=\{a_j \mid \tau(x_j)=true\}\cup\{\overline{a}_j \mid \tau(x_j)=false\}\cup\{c_i \mid C_i=x_j\vee x_\ell, \tau(x_j)=true \ \vee\ \tau(x_\ell)=true\}\cup\{\varphi \mid \tau \text{ is a satisfying assignment for } \phi\}$ is a stable extension of Λ. In fact, if ϕ evaluates to *true* under τ, then S contains φ, no c_i is in S as for each c_i (attacking φ), with $C_i=x_j\vee x_\ell$ being a clause of ϕ, there is either a_j or a_ℓ in S attacking c_i; moreover; either a_j or $\overline{a}_j$ in S for each $j\in\{1,\ldots,n\}$. On the other hand, if ϕ evaluates to *false* under τ, then S contains at least a c_i argument as there is a clause C_i whose (positive) variables are assigned *false* by τ. Also in this case, either a_j or $\overline{a}_j$ in S for each $j\in\{1,\ldots,n\}$. Thus, $S=\beta(\tau)$ is conflict-free, admissible, and total, meaning that it is a stable extension for each truth assignment τ for ϕ.

Observe that, for every truth assignment τ for the propositional variables of ϕ, the goal argument φ belongs to the corresponding extension $\beta(\tau)\in\texttt{st}(\Lambda)$ iff ϕ evaluates to *true* under τ.

We now compute $PrEA^{\texttt{st}}_{\Delta}(\varphi)$. Since Δ has a unique world Λ, $PrEA^{\texttt{st}}_{\Delta}(\varphi)=\sum_{E\in\texttt{st}(\Lambda)\wedge\varphi\in E} PrE(E,\Lambda,\texttt{st})$ where $PrE(E,\Lambda,\texttt{st})=\sum_{X\in Exp^{\texttt{st}}_{\Lambda}(E)}\pi(X)$.

Let $E\in\texttt{st}(\Lambda)$ be an extension such that $\varphi\in E$. As said earlier, such kind of extensions is of the form $E=\{\varphi\}\cup\{a_j \mid \tau(x_j)=true\}\cup\{\overline{a}_j \mid \tau(x_j)=false\}$. It can be shown that an explanation X for E is any sequence $X=\langle y_1,\ldots,y_n\rangle$ of arguments in $E\setminus\{\varphi\}$. In fact, $\Lambda^*=\Lambda$ since $\texttt{gr}(\Lambda)=\emptyset$, and Λ has n unattacked SCCs (ordered as $\langle\mathcal{C}_1,\ldots,\mathcal{C}_n\rangle$) consisting of the n cycles corresponding to the pairs of the mutual attacks $(a_j,\overline{a}_j)$ and $(\overline{a}_j,a_j)$ (with $j\in\{1,\ldots,n\}$). Hence, the first argument y_1 in X belongs to the SCC $\mathcal{C}_1$ of $\Lambda_1=\Lambda^*$. Next, the second argument y_2 in X belongs to the SCC $\mathcal{C}_2$ of the AF $\Lambda_2=\Lambda^*_{y_1}$ obtained from Λ_1 by removing the attacker of y_1 (that is $\overline{a}_j$ if $y_1=a_j$; a_j if $y=\overline{a}_j$). In general, the j^{th} argument y_i in X (with $j\in\{2,\ldots,n\}$) belongs to the SCC $\mathcal{C}_j$ of the AF Λ_j obtained from Λ_{j-1} by removing the attackers of y_{j-1} as well as the arguments in the grounded extension of Λ_{j-1} or attacked by them.

Consider now an extension $E\in\texttt{st}(\Lambda)$ such that $\varphi\notin E$. This is an extension corresponding to a truth assignment $\tau=\beta^{-1}(E)$ under which ϕ evaluates to *false*. Such kind of extensions is of the form $E=\{c_i \mid C_i=x_j\vee x_\ell, \tau(x_j)=false, \tau(x_\ell)=false\}\cup\{a_j \mid \tau(x_j)=true\}\cup\{\overline{a}_j \mid \tau(x_j)=false\}$. Reasoning similarly to the previous case, it can be shown that an explanation X for E is any sequence $X=\langle y_1,\ldots,y_n\rangle$ of arguments in $E\setminus\{c_i \mid C_i=x_j\vee x_\ell, \tau(x_j)=false, \tau(x_\ell)=false\}$.

Since an explanation X for any extension $E\in\texttt{st}(\Lambda)$ is a sequence of n arguments, each of them in $\{a_{i_j},\overline{a}_{i_j}\}$ (with $i_j\in\{1,\ldots,n\}$), there is only one explanation for each stable extension E of Λ.

Moreover, Λ has 2^n stable extensions, one for each truth assignment τ of ϕ. This leads to a probabilistic trie $\mathcal{T}^{\texttt{st}}_{\Lambda}$ for Λ consisting of the 2^n sequences in $Exp^{\texttt{st}}(\Lambda)$, and such that $\pi(X)=\frac{1}{2^n}$ since all explanations are of length n.

Therefore, $PrEA^{\mathtt{st}}_{\Delta}(\varphi) = \sum_{X \in Exp^{\mathtt{st}}_{\Lambda}(E) \wedge E \in \mathtt{st}(\Lambda) \wedge \varphi \in E} \pi(X) = \frac{\#\phi}{2^n}$, where $\#\phi$ is the number of satisfying assignments for ϕ. That is, $\#\phi = 2^n \cdot PrEA^{\mathtt{st}}_{\Delta}(\varphi)$.

Finally, as for the considered framework it holds that $\mathtt{st}(\Lambda) = \mathtt{sst}(\Lambda) = \mathtt{pr}(\Lambda)$, the result also holds for the semi-stable and preferred semantics.

We now consider the complete semantics (i.e. $\sigma = \mathtt{co}$). We define a PrAF $\Delta = \langle A, \Sigma, P\rangle$ as follows:

- The set A consists of: i) a triple of arguments, namely a_i, $\overline{a}_i$, and c_{a_i} for each propositional variable $x_i \in X$; ii) an argument c_i for each clause C_i appearing in ϕ; and iii) an argument φ;
- Σ consists of the following attacks. For each clause $C_i = x_j \vee x_\ell$ (with $i \in \{1, \dots, k\}$), an attack (c_i, φ) as well as attacks (a_j, c_i) and (a_ℓ, c_i) are in Σ; moreover, for each $j \in \{1, \dots, n\}$, Σ contains a pair of mutual attacks, $(a_j, \overline{a}_j)$ and $(\overline{a}_j, a_j)$. Finally, we have the attacks (c_{a_i}, φ) for each variable $x_i \in X$ as well as (a_i, c_{a_i}) and $(\overline{a_i}, c_{a_i})$.
- Function P assigns probability equal to 1 to every argument in A.

Finally, we let the goal argument be φ.

We can show that there is a bijection $\beta : \mathcal{T} \to \mathtt{co}(\Lambda)^*$ between the set $\mathcal{T}$ of truth assignments satisfying ϕ and the set $\mathtt{co}(\Lambda)^*$ of complete extensions of Λ containing φ. In particular, function β is such that, given a truth assignment τ for the propositional variables of ϕ s.t. τ is a satisfying assignment for ϕ, the set $S = \beta(\tau) = \{a_j \mid \tau(x_j) = true\} \cup \{\overline{a}_j \mid \tau(x_j) = false\} \cup \{c_i \mid C_i = x_j \vee x_\ell, \tau(x_j) = true \vee \tau(x_\ell) = true\} \cup \{\varphi \mid \tau \text{ is a satisfying assignment for } \phi\}$ is a complete extension of Λ containing φ. In fact, as ϕ evaluates to *true* under τ, then S contains φ, no c_i is in S as for each c_i (attacking φ), with $C_i = x_j \vee x_\ell$ being a clause of ϕ, there is either a_j or a_ℓ in S attacking c_i; moreover; either a_j or $\overline{a}_j$ in S for each $j \in \{1, \dots, n\}$. The same consideration holds for arguments c_{a_i}. On the other hand, for any complete extension $E \in \mathtt{co}(\Lambda)$ s.t. $\varphi \in E$, E is also a stable extension and $\beta^{-1}(E)$ gives a truth assignment τ s.t. $\tau(x_i) = true$ iff $a_i \in E$.

We now compute $PrEA^{\mathtt{co}}_{\Delta}(\varphi)$. Since Δ has a unique world Λ, $PrEA^{\mathtt{co}}_{\Delta}(\varphi) = \sum_{E \in \mathtt{co}(\Lambda) \wedge \varphi \in E} PrE(E, \Lambda, \mathtt{co})$ where $PrE(E, \Lambda, \mathtt{co}) = \sum_{X \in Exp^{\mathtt{co}}_{\Lambda}(E)} \pi(X)$. Let $E \in \mathtt{co}(\Lambda)$ be an extension such that $\varphi \in E$. As said earlier, such kind of extensions is of the form $E = \{\varphi\} \cup \{a_j \mid \tau(x_j) = true\} \cup \{\overline{a}_j \mid \tau(x_j) = false\}$. It can be shown that an explanation X for E is any sequence $X = \langle y_1, \dots, y_n\rangle$ of elements in $E \setminus \{\varphi\} \cup \{\varepsilon\}$. In fact, $\Lambda^* = \Lambda$ since $\mathtt{gr}(\Lambda) = \emptyset$, and Λ has n unattacked SCCs (ordered as $\langle \mathcal{C}_1, \dots, \mathcal{C}_n\rangle$) consisting of the n cycles corresponding to the pairs of the mutual attacks $(a_j, \overline{a}_j)$ and $(\overline{a}_j, a_j)$ (with $j \in \{1, \dots, n\}$). Hence, at each step, either $a_i, \overline{a_i}$, or ε can be chosen.

Moreover, Λ has 3^n complete extensions. This leads to a probabilistic trie $\mathcal{T}^{\mathtt{co}}_{\Lambda}$ for Λ consisting of the 3^n sequences in $Exp^{\mathtt{co}}(\Lambda)$, and such that $\pi(X) = \frac{1}{3^n}$.

Therefore, $PrEA^{\mathtt{co}}_{\Delta}(\varphi) = \sum_{X \in Exp^{\mathtt{co}}_{\Lambda}(E) \wedge E \in \mathtt{co}(\Lambda) \wedge \varphi \in E} \pi(X) = \frac{\#\phi}{3^n}$, where $\#\phi$ is the number of satisfying assignments for ϕ. That is, $\#\phi = 3^n \cdot PrEA^{\mathtt{co}}_{\Delta}(\varphi)$. □

Proposition 4. *For any chosen PDF, i)* PrA[$\mathtt{gr}$] *is in FP for AF, and ii)* PrA[σ] *is in FP for* $\sigma \in \{\mathtt{pr}, \mathtt{co}, \mathtt{st}, \mathtt{sst}\}$ *and acyclic AFs.*

Proof. For AFs the grounded semantics prescribes a unique extension, that can be computed in polynomial time, and for acyclic AFs all semantics collapse to the grounded one. Therefore, in both cases there is only one extension whose probability is 1. □

Theorem 4. *Consider a semantics* $\sigma \in \{\mathtt{gr}, \mathtt{co}, \mathtt{pr}, \mathtt{st}, \mathtt{sst}\}$. *Unless NP ⊆ BPP, there is no FPRAS for* PrA[σ], *even for acyclic PrAFs and for any chosen PDF.*

Proof. The claim follows from the proof of Theorem 2. In particular, in that proof we show a reduction from #P2CNF to PrA[σ], for every semantics $\sigma \in \{\mathtt{gr}, \mathtt{co}, \mathtt{pr}, \mathtt{st}, \mathtt{sst}\}$ and for any chosen PDF. Starting from a P2CNF formula ϕ, the constructed PrAF Δ and argument g are such that Δ is acyclic, and $\#\phi = \frac{PrA^{\sigma}_{\Delta}(g)}{2^n}$, where $\#\phi$ is the number of truth assignments satisfying ϕ, and n is the number of propositional variables of ϕ. It is well-known that #P2CNF does not admit an FPRAS (unless NP ⊆ BPP) [91]. Since $\#\phi$ and $PrA^{\sigma}_{\Delta}(g)$ coincide, up to a multiplicative factor (i.e. $\frac{1}{2^n}$, which can be easily computed in polynomial time), the reduction then shows that PrA[σ] does not admit an FPRAS, for all $\sigma \in \{\mathtt{gr}, \mathtt{co}, \mathtt{pr}, \mathtt{st}, \mathtt{sst}\}$, and for any chosen PDF, even when considering only acyclic PrAFs (unless NP ⊆ BPP). □

Lemma 1. *For each* $\sigma \in \{\mathtt{pr}, \mathtt{st}, \mathtt{sst}\}$, UnCA[$\sigma$] *is NP-hard.*

Proof. We provide a reduction from the NP-complete problem 3SAT, that is, the problem asking whether a given 3CNF formula is satisfiable. We point out that, although the construction is somehow known in the literature, what we are going to prove here is that this construction is also a reduction from 3SAT to the restricted problem UnCA[σ].

Let $\phi = C_1 \wedge C_2 \wedge \cdots \wedge C_k$ be a 3CNF, where $X = \{x_1, \dots, x_n\}$ is the set of its propositional variables. We define the AF $\Lambda = \langle A, \Sigma\rangle$ as follows.

- The set A contains: i) two arguments a_x and $a_{\bar{x}}$, for each variable $x \in X$ and its negation $\bar{x}$; ii) an argument c_i, for each clause C_i of ϕ; iii) an argument φ and an argument ψ;
- The set Σ contains: i) for each variable $x \in X$ and its negation $\bar{x}$, two attacks $(a_x, a_{\bar{x}})$ and $(a_{\bar{x}}, a_x)$; ii) for each clause $C_i = \ell_1 \vee \ell_2 \vee \ell_3$ of ϕ, an attack (c_i, φ) and three attacks (a_{ℓ_1}, c_i), (a_{ℓ_2}, c_i), and (a_{ℓ_3}, c_i); iii) for each $\alpha \in A \setminus \{\varphi\}$, an attack (ψ, α); iv) the attack (φ, ψ).

The goal is to prove that, for each $\sigma \in \{\mathtt{pr}, \mathtt{st}, \mathtt{sst}\}$, when ϕ is unsatisfiable, then no extension $E \in \sigma(\Lambda)$ contains φ, and when ϕ is satisfiable, every extension $E \in \sigma(\Lambda)$ contains φ.

We start with some considerations.

(1) First, note that no preferred (and thus, no (semi)-stable) extension contains the argument ψ. If this would be the case, then, since ψ attacks all arguments in $A \setminus \{\varphi\}$, and φ attacks ψ, such an extension would be the singleton $\{\psi\}$. However, $\{\psi\} \neq Acc(\{\psi\})$ (as φ attacks ψ, but ψ does not attack φ).

(2) Second, note that for any preferred (and hence, any (semi)-stable) extension E, if E is non-empty, it must contain φ (and not contain ψ, from (1)). In fact, if some argument in $A \setminus \{\varphi, \psi\}$ is in E, since ψ attacks every argument, except for φ, φ must be in E, as it is the one attacking ψ.

We first show that when the formula ϕ is unsatisfiable, then no preferred extension (and thus, no (semi)-stable extension) contains φ. Assume, towards a contradiction, that there exists a preferred extension E of Λ, such that $\varphi \in E$. Since $\varphi \in E$, no argument c_i, for $i \in \{1, \dots, k\}$ is in E, as E would not be conflict-free. Moreover, since each c_i attacks φ, but $\psi \notin E$ (from (1)), some conflict-free subset B of $\{a_{x_1}, a_{\bar{x}_1}, \dots, a_{x_n}, a_{\bar{x}_n}\}$ must be contained in E, for E to be acceptable w.r.t. itself, i.e. $E = Acc(E)$. However, since ϕ is unsatisfiable, any such a set B will always leave at least one argument c_i unattacked, and thus $E \neq Acc(E)$, obtaining a contradiction.

We now show that when the formula ϕ is satisfiable, then a (non-empty) stable extension exists. Assume ϕ is satisfiable, and let τ be a truth assignment for the propositional variables of ϕ, such that τ satisfies ϕ. We consider the set

$$E_\tau = \{a_x \in A \mid \tau(x) = true\} \cup \{a_{\bar{x}} \in A \mid \tau(x) = false\} \cup \{\varphi\}.$$

We claim E_τ is a stable extension. Clearly, E_τ is a maximal conflict-free set, since one argument a_x from each pair $\{a_x, a_{\bar{x}}\}$ is in E_τ, no argument c_i can be in E_τ, as τ satisfies ϕ and thus all c_i's are attacked by E_τ. Finally, ψ is not in E_τ from (1). What remains to show is that every argument attacking some argument in E_τ, is attacked by some argument in E_τ. In fact, every argument of the form a_x not in E_τ and attacking $a_{\bar{x}} \in E_\tau$ is attacked by $a_{\bar{x}}$, and vice versa. Moreover, the argument ψ attacking every argument in E_τ is attacked by φ. Finally, the fact that every argument of the form c_i, for $i \in \{1, \dots, k\}$, is attacked by some argument in E_τ, follows by construction of E_τ and from the fact that τ satisfies ϕ. Thus, E_τ is a stable extension of Λ. From consideration (2), any non-empty preferred extension of Λ must contain φ. Thus, since at least a non-empty stable extension exists when ϕ is satisfiable (i.e. E_τ), and thus all stable extensions are non-empty, we conclude that when ϕ is satisfiable, all stable, semi-stable and preferred extensions of Λ contain φ, and the claim follows. □

Theorem 5. *Let $\sigma \in \{\mathtt{pr}, \mathtt{st}, \mathtt{sst}\}$. Unless NP ⊆ BPP, there is no FPARAS for* PrA[σ]*, for any chosen PDF.*

Proof. From Lemma 1, it suffices to prove that the existence of an FPARAS for PrA[σ], with a given fixed PDF, implies that UnCA[σ] is in BPP.

Consider a semantics $\sigma \in \{\mathtt{pr}, \mathtt{st}, \mathtt{sst}\}$, and assume that A is indeed an FPARAS for PrA[σ], for some fixed PDF. We now show that UnCA[σ] is in BPP, by means of the following randomized procedure. Let BPPAlgo be the randomized decision procedure taking as input an σ-uniform pair (Λ, g) of an AF Λ and argument g, and performing the following:

1. Let $\epsilon = 1/3$ and $\delta = 1/3$;
2. Construct the PrAF Δ from Λ by giving probability 1 to all arguments in Λ;
3. Compute $\hat{p} = A(\Delta, g, \epsilon, \delta)$;
4. if $\hat{p} \in [1 - \epsilon, 1 + \epsilon]$, then return yes;
5. else return no.

Since A runs in polynomial time, so does BPPAlgo. Let now (Λ, g) be an σ-uniform pair given as input to BPPAlgo. We consider two cases.

Assume first that an extension $E \in \sigma(\Lambda)$ exists, such that $g \in E$. Then, since (Λ, g) is σ-uniform, and since every argument in the PrAF Δ constructed in line 2 has probability 1, we conclude that $PrA^\sigma_\Delta(g) = 1$ (regardless of the chosen PDF). Thus, since A is an FPARAS, line 4 of algorithm BPPAlgo is executed (i.e. returns the right answer for UnCA[σ] with input (Λ, g)) with probability at least $1 - \delta = 2/3$, and line 5 is executed (i.e. returns the wrong answer) with probability at most $\delta = 1/3$.

Assume now that no extension $E \in \sigma(\Lambda)$ exists such that $g \in E$. Again, since every argument in the constructed PrAF Δ has probability 1, we conclude that $PrA^\sigma_\Delta(g) = 0$, regardless of the chosen PDF. Since A is an FPARAS, $Pr(|A(\Delta, g, \epsilon, \delta) - 0| \leq \epsilon) \geq 1 - \delta$. Moreover, since $\epsilon < 1/2$, the events $|A(\Delta, g, \epsilon, \delta) - 0| \leq \epsilon$ and $|A(\Delta, g, \epsilon, \delta) - 1| \leq \epsilon$ are disjoint, and thus $Pr(|A(\Delta, g, \epsilon, \delta) - 1| \leq \epsilon)$ is at most the probability that the event $|A(\Delta, g, \epsilon, \delta) - 0| \leq \epsilon$ does not hold, i.e. δ. Hence, line

4 of algorithm BPPAlgo is executed (i.e. returns the wrong answer for UnCA[σ] with input (Λ, g)) with probability at most $\delta = 1/3$, and thus executes line 5 (i.e. returns the right answer) with probability at least $1 - \delta = 2/3$. Hence, UnCA[σ] is in BPP, as required. $\square$

Corollary 3. *Let $\sigma \in \{\texttt{pr}, \texttt{st}, \texttt{sst}\}$. There is no FPARAS and no FPRAS for* PrA[σ] *for AF (PrAF with probabilities equal to* 1*) and for any chosen PDF.*

Proof. As for FPARAS, the claim follows by observing that the proof of Theorem 5 works for AF, as a PrAF with probabilities equal to 1 is considered.

Then, the non-existence of an FPARAS implies the non existence of an FPRAS as $f(x) \le 1$ (with function f being the output of problem PrA[σ], cf. Definition 8) and thus $\Pr(|A(x, \epsilon, \delta) - f(x)| \le \epsilon) \ngeq 1 - \delta$ implies that $\Pr(|A(x, \epsilon, \delta) - f(x)| \le \epsilon \cdot f(x)) \ngeq 1 - \delta$. $\square$

Theorem 6. *Algorithm 2 with input an AF Λ and a semantics $\sigma \in \{\texttt{gr}, \texttt{co}, \texttt{pr}, \texttt{st}, \texttt{sst}\}$ is such that:*

- *It outputs an $X \in Exp^{\sigma}(\Lambda)$ with probability $\pi(X)$, and*
- *it runs in polynomial time,*

whenever i) $\sigma = \texttt{gr}$, or ii) Λ has no odd-length cycles.

Proof. If $\sigma = \texttt{gr}$, at each iteration the algorithm *i*) determines the first SCC, *ii*) appends ε to X, *iii*) deletes from the current AF Λ all arguments in the SCC, and *iv*) adds to Γ all arguments attacked by the delete ones. Thus, it returns (with probability 1) a sequence of symbols whose length is bounded by the number of SCCs, that in turns is bounded by the number of arguments.

Consider now the case of $\sigma \in \{\texttt{co}, \texttt{pr}, \texttt{st}, \texttt{sst}\}$. Since we are considering AFs Λ without odd cycles, we have that $\texttt{pr}(\Lambda) = \texttt{st}(\Lambda) = \texttt{sst}(\Lambda)$. We first prove the following auxiliary lemma.

Lemma 3. *Consider an AF $\Lambda = \langle A, \Sigma\rangle$ without odd-length cycles. Then, the following hold:*

- *for every argument a in the first SCC of $\hat{\Lambda}$, there exists a complete and stable (and hence preferred and semi-stable) extension E that contains a.*
- *Every stable (and hence preferred and semi-stable) extension of Λ contains at least one argument in the first SCC of $\hat{\Lambda}$.*

Proof. (Item 1) Let C be the first SCC of $\hat{\Lambda} = \langle \hat{A}, \hat{\Sigma}\rangle$. We first show that for every argument $a \in C$, and every argument $b_1, b_2 \in C$, such that even-length paths π_1, π_2 exist from b_1 to a and from b_2 to a, respectively, it holds that $\{a, b_1, b_2\}$ is conflict-free in $\hat{\Lambda}$. Assume, towards a contradiction, that either *1)* $\{a, b_1\}$ is not conflict-free (the case for $\{a, b_2\}$ is symmetric) or *2)* $\{b_1, b_2\}$ is not conflict-free. If *1)*, then either a attacks b_1 or b_1 attacks a. Note that a cannot attack b_1, otherwise the cycle $\pi_1, (a, b_1)$ is of odd-length. Thus, it must be that b_1 attacks a. However, since a, b_1 are in the same SCC, there is a path π' from a to b_1. Moreover, π' must be of even-length, as otherwise the cycle π', π_1 is of odd-length. However, π' cannot be of even-length, otherwise the cycle $\pi', (b_1, a)$ is of odd-length. Hence $\{a, b_1\}$ is conflict-free. Symmetrically, the same applies to $\{a, b_2\}$. Assume now that *2)* holds. Then, either b_1 attacks b_2, or vice versa. We focus only on the case that b_1 attacks b_2, as the other case is symmetric. As discussed, any path π', from a to b_1 must be of even-length. However, the cycle $\pi', (b_1, b_2), \pi_2$ is of odd-length, because also π_2 is of even-length. Hence $\{b_1, b_2\}$ is conflict-free.

Consider now an argument $a \in C$. By the above property, the set

$$E_a = \{a\} \cup \{b \in C \mid b \text{ reaches } a \text{ with an even-length path}\}$$

is conflict-free in $\hat{\Lambda}$. We now show that E_a is a stable extension of $\hat{\Lambda}{\downarrow_C}$, by showing that every argument in $C \setminus E$ is attacked by some argument of E_a. Assume, towards a contradiction, that there is $d \in C \setminus E_a$ such that no argument in E_a attacks d. Since $d \in C$, there is a path π from d to a, and this path cannot be of even-length, otherwise $d \in E_a$. Moreover, since no argument in E_a attacks d and since d is in a SCC, there must be an argument $d' \in C \setminus E_a$ attacking d. However, the path $(d', d), \pi$ is of even-length, which implies that $d' \in E_a$, obtaining a contradiction. Hence, E_a is a stable extension of C.

Since C is the first SCC of $\hat{\Lambda}$, E_a is a subset of some stable extension of $\hat{\Lambda}$, and the claim follows. In particular, as these extensions are also complete, the statement also holds for complete semantics.

(Item 2) The claim follows from the fact that any stable extension E of $\hat{\Lambda}$ must attack any argument a not in E. So, if $a \in C$, where C is the first SCC of $\hat{\Lambda}$, the only way of attacking a is by means of an argument b in C. Hence, $b \in C$ must be in E. Thus, any stable extension of $\hat{\Lambda}$ must contain at least one argument from the first SCC of $\hat{\Lambda}$. $\square$

With the above lemma in place, it is easy to prove that for every stable extension E of Λ, there is an execution of Algorithm 2 that outputs the explanation X s.t. $ext(X) = E$, or equivalently (cf. Theorem 1) $E = \texttt{gr}(\Lambda_{\tilde{X}})$. In particular, consider an execution of Algorithm 2 with input Λ. Let $\Lambda^1 = \Lambda$, $E^1 = \emptyset$, and for each iteration $i > 0$ of the algorithm, we let a_i be the argument chosen in the first SCC of $\hat{\Lambda}^i$, $\Lambda^{i+1} = \hat{\Lambda}^i_{a_i}$ is the AF considered at the beginning of iteration $i+1$, and $E^{i+1} = E^i \cup G_{a_i}$, for $G_{a_i} = \mathtt{gr}(\Lambda^i_{a_i})$. By Lemma 3, every argument a of the first SCC of $\hat{\Lambda}^i$ for each $i \geq 0$, belongs to some stable extension of $\hat{\Lambda}^i$ and hence of Λ^i. Moreover, since $G_{a_i} = \mathtt{gr}(\Lambda^i_{a_i})$ belongs to any such extension, we conclude that E^i is a subset of a stable extension of Λ, for each $i \geq 0$. Thus, if there is n such that E^n cannot be further extended, which means that no more arguments exist belonging to some stable extension of Λ (by Lemma 3), we conclude that E^n is a stable extension of Λ. Moreover, by construction, the sequence $\langle a_1, \ldots, a_n \rangle$ is an explanation for E^n.

Regarding the addition of new attacks during the process, since they start from arguments whose status w.r.t. E^n is defeated they are irrelevant [90]. Also, by adding these attacks we obtain that the arguments in the first SCC apart those removed (i.e. those in G_{a_i} and $G^+_{a_i}$) still form a SCC.

Let $\mathcal{C}_i$ be the set of arguments of the first SCC obtained at some iteration i. We have that the probability of choosing a_i is $1/|\mathcal{C}_i|$. Therefore, the probability of choosing a particular sequence X of arguments corresponding to an explanation of some E is equal to $p = \prod_{i=1}^{n} 1/|\mathcal{C}_i|$. This value coincides with the value associated with the leaf node corresponding to X, in the probabilistic trie $\mathcal{T}^{\mathtt{pr}}_{\Lambda}$.

The algorithm runs in polynomial time as the number of iterations is linear in the number of arguments and each iteration performs polynomial time operations. Moreover, computing the grounded extensions G_x is polynomial too.

Considering the complete semantics, a first difference w.r.t. the case of stable (preferred and semi-stable) semantics discussed earlier is that now the addition of new attacks can be relevant [90]. However, since the arguments in Ω will be defeated in next steps of the algorithm, in all the complete extension E' explained by the explanations constructed in the subsequent steps, their status is defeated.

That is, the update of line 14 is irrelevant w.r.t. complete semantics, that is, any extension E' continues to be a complete extension of the AF being processed [89].

The second difference concerns the use of the set Γ. As complete extensions may admit arguments that are neither accepted nor defeated, it corresponds to making no choice at some step of the algorithm, i.e. no argument x of the i-th SCC is accepted. Hence, as complete extension enjoys the directionality property [92],[6] the arguments that are attacked from x (corresponding to Γ) cannot be selected from the set $\mathcal{C}$ at any future step of the algorithm. This corresponds to the choice of ε. In this case we remove all the arguments of this SCC. □

Theorem 7. *Problem* PrEA[σ]*, with* $\sigma \in \{\mathtt{gr}, \mathtt{co}, \mathtt{pr}, \mathtt{st}, \mathtt{sst}\}$*, has an FPARAS if either i)* $\sigma = \mathtt{gr}$*, or ii) the input PrAF has no odd cycles.*

Proof. With Theorem 6 in place, it suffices to prove that Algorithm 1 provides the error and probabilistic guarantees of an FPARAS. That is, for some $\sigma \in \{\mathtt{gr}, \mathtt{co}, \mathtt{pr}, \mathtt{st}, \mathtt{sst}\}$, for every PrAF Δ, argument g, and numbers $\epsilon > 0$, $0 < \delta < 1$, it holds that

$$\Pr(\left|\mathsf{Apx}(\Delta, g, \epsilon, \delta) - PrEA^{\sigma}_{\Delta}(g)\right| \leq \epsilon) \geq 1 - \delta.$$

To prove the above, it suffices to note that Apx is a random variable of the form

$$\frac{1}{n} \cdot \sum_{i=1}^{n} X_i,$$

where $n = \lceil \frac{1}{2\epsilon^2} \cdot \ln(2/\delta) \rceil$, and each X_i is a random variable with values in $\{0, 1\}$, such that $\Pr(X_i = 1) = PrEA^{\sigma}_{\Delta}(g)$. Thus, from Hoeffding's inequality [49],

$$\Pr(\left|\mathsf{Apx}(\Delta, g, \epsilon, \delta) - PrEA^{\sigma}_{\Delta}(g)\right| \leq \epsilon) \geq 1 - \delta,$$

for all $n \geq \frac{1}{2\epsilon^2} \cdot \ln(2/\delta)$, and the claim follows. □

A.3. Inapproximability for credulous and skeptical acceptance

Theorem 8. *For* $\sigma \in \{\mathtt{gr}, \mathtt{co}, \mathtt{pr}, \mathtt{st}, \mathtt{sst}\}$*, unless NP* $\subseteq$ *BPP, there is no FPRAS for* PrCA[σ] *and* PrSA[σ]*, even for acyclic PrAFs.*

Proof. This follows by reasoning exactly as in the proof of Theorem 4 where the construction of the proof of Theorem 2 is used. In particular, in the proof of Theorem 2 we show a reduction from #P2CNF to PrA[σ], for every semantics

[6] We recall that the idea underlying directionality is that the status (accepted, defeated, undecided) of an argument a should be affected only by that of its defeaters (which in turn are affected by their defeaters and so on). More formally, an argumentation semantics σ satisfies the directionality property if and only if for any argumentation framework $\Lambda = \langle A, \Sigma \rangle$ and set of arguments $S \subseteq A$ receiving no attack in Λ it holds that $\{E \cap S \mid E \in \sigma(\Lambda)\} = \sigma(\Lambda\downarrow_S)$.

$\sigma \in \{\mathtt{gr}, \mathtt{co}, \mathtt{pr}, \mathtt{st}, \mathtt{sst}\}$ and for any chosen PDF. Since the PrAF Δ considered in that proof is acyclic, every AF $w \in pw(\Delta)$ is acyclic. Thus, the grounded extension of $w \in pw(\Delta)$ is the unique extension of w under any semantics in $\sigma \in \{\mathtt{gr}, \mathtt{co}, \mathtt{pr}, \mathtt{st}, \mathtt{sst}\}$. Thus, PrA[$\sigma$] coincides with PrCA[$\sigma$] and PrSA[$\sigma$], from which the claim follows by reasoning as in the proof of Theorem 4. □

Theorem 9. *Unless $NP \subseteq BPP$, i) there is no FPARAS for* PrCA[σ] *and for* PrSA[σ] *with $\sigma \in \{\mathtt{pr}, \mathtt{st}, \mathtt{sst}\}$, even for PrAFs without odd-length cycles; and ii) there is no FPARAS for* PrCA[co]*, even for AFs (PrAFs with probabilities all equal to 1).*

Proof. We use the fact that checking whether a given argument g is credulously accepted in an AF Λ is NP-hard, when $\sigma \in \{\mathtt{pr}, \mathtt{st}, \mathtt{sst}\}$, even for AFs having no odd-length cycles [3]. With this in place, we can reason similarly to what is done to prove Theorem 5.

Case i), PrCA. The construction is the same as the one given in the proof of Theorem 5. The only difference is that the NP-hard problem of credulous acceptance is solved, via the randomized procedure exploiting an FPARAS for PrCA[σ].

Case i), PrSA. First observe that UnCA[σ] for $\sigma \in \{\mathtt{pr}, \mathtt{st}, \mathtt{sst}\}$ also answers to the problem of checking whether all σ-extensions of the σ-uniform AF Λ contain the goal argument g. Thus, for any semantics $\sigma \in \{\mathtt{pr}, \mathtt{st}, \mathtt{sst}\}$ determining the skeptical acceptance in σ-uniform AFs is still NP-hard. Thus, we can use again the construction given in the proof of Theorem 5. The only difference is that the NP-hard problem of determining the skeptical acceptance in σ-uniform AFs is solved, via the randomized procedure exploiting an FPARAS for PrSA[σ].

Case ii). For the proof, we rely on the next lemma that shows a gap property of the problem of credulously accepting an argument.

We say that a pair (Λ, g) of an AF Λ and argument g is co-*empty*, if either g is credulously accepted under complete semantics or $\mathtt{co}(\Lambda) = \{\emptyset\}$, that is $\emptyset$ is the only complete extensions of Λ.

Let us now consider the following restriction of the classical credulous acceptance problem under complete semantics.

PROBLEM :	EmCA[co]
INPUT :	An co-empty pair (Λ, g).
QUESTION :	Is there $E \in \sigma(\Lambda)$ such that $g \in E$?

Lemma 4. EmCA[co] *is NP-hard.*

Proof. Consider the construction introduced in the proof of Lemma 1, where a reduction from 3SAT to UnCA[σ], with $\sigma \in \{\mathtt{pr}, \mathtt{st}, \mathtt{sst}\}$, is provided. We use the notation introduced in that proof. We show that if $\phi = C_1 \wedge C_2 \wedge \cdots \wedge C_k$ is unsatisfiable, then the only complete extension is the empty set. Moreover, if ϕ is satisfiable, then there is an extension $E \in \sigma(\Lambda)$ containing φ. This is proved in the proof of Lemma 1 for preferred semantics, and thus holds for complete semantics (as preferred extensions are a subset of complete extensions).

We now show that if ϕ is unsatisfiable, then no complete extension contains φ. This suffices to prove that the only complete extension of Λ is the empty set. In fact, assume towards a contradiction that there exists a complete extension E of Λ, such that $\varphi \in E$. Since $\varphi \in E$, no argument c_i, for $i \in \{1, \dots, k\}$ is in E, as E would not be conflict-free. Moreover, since each c_i attacks φ, and $\psi \notin E$, some conflict-free subset B of $\{a_{x_1}, a_{\bar{x}_1}, \dots, a_{x_n}, a_{\bar{x}_n}\}$ must be contained in E, for E to be acceptable w.r.t. itself, i.e. $E = Acc(E)$. However, since ϕ is unsatisfiable, any such a set B leaves at least one argument c_i unattacked, and thus $E \neq Acc(E)$, obtaining a contradiction. □

We can now exploit the above lemma to prove our inapproximability result, similarly to what is done in the proof of Theorem 5.

From Lemma 4, it suffices to prove that the existence of an FPARAS for PrCA[co], implies that EmCA[co] is in BPP. Assume that A is indeed an FPARAS for PrCA[co]. We now show that EmCA[co] is in BPP, by means of the following randomized procedure. Let BPPAlgo be the randomized decision procedure taking as input an co-empty pair (Λ, g) of an AF Λ and argument g, and performing the following:

1. Let $\epsilon = 1/3$ and $\delta = 1/3$;
2. Construct the PrAF Δ from Λ by giving probability 1 to all arguments in Λ;
3. Compute $\hat{p} = A(\Delta, g, \epsilon, \delta)$;
4. if $\hat{p} \in [1 - \epsilon, 1 + \epsilon]$, then return yes;
5. else return no.

Since A runs in polynomial time, so does BPPAlgo. Let now (Λ, g) be an co-empty pair given as input to BPPAlgo. We consider two cases.

Assume first that an extension $E \in \mathtt{co}(\Lambda)$ exists, such that $g \in E$. Then, since (Λ, g) is co-empty, and since every argument in the PrAF Δ constructed in line 2 has probability 1, we conclude that $PrCA^{\mathtt{co}}_{\Delta}(g) = 1$. Thus, since A is an FPARAS, line 4 of algorithm BPPAlgo is executed (i.e. returns the right answer for EmCA[σ] with input (Λ, g)) with probability at least $1 - \delta = 2/3$, and line 5 is executed (i.e. returns the wrong answer) with probability at most $\delta = 1/3$.

Assume now that g is not credulously accepted in Λ under complete semantics. Again, since every argument in the constructed PrAF Δ has probability 1, we conclude that $PrCA^{\mathsf{co}}_{\Delta}(g) = 0$. Since A is an FPARAS, $Pr(|A(\Delta, g, \epsilon, \delta) - 0| \leq \epsilon) \geq 1 - \delta$. Moreover, since $\epsilon < 1/2$, the events $|A(\Delta, g, \epsilon, \delta) - 0| \leq \epsilon$ and $|A(\Delta, g, \epsilon, \delta) - 1| \leq \epsilon$ are disjoint, and thus $Pr(|A(\Delta, g, \epsilon, \delta) - 1| \leq \epsilon)$ is at most the probability that the event $|A(\Delta, g, \epsilon, \delta) - 0| \leq \epsilon$ does not hold, i.e. δ. Hence, line 4 of algorithm BPPAlgo is executed (i.e. returns the wrong answer for EmCA[co] with input (Λ, g)) with probability at most $\delta = 1/3$, and thus executes line 5 (i.e. returns the right answer) with probability at least $1 - \delta = 2/3$. Hence, EmCA[co] is in BPP, as required. $\square$

A.4. Extended PrAFs

Theorem 10. *Given an EPrAF $\Delta = \langle A, \Sigma, P_I\rangle$, for any PrAF $\Delta' = \langle A, \Sigma, P\rangle$ such that $P(a) \in P_I(a)$ for all $a \in A$, it holds that $PrA^{\sigma}_{\Delta'}(g) \in EPrA^{\sigma}_{\Delta}(g)$ for all $g \in A$, independently of the PDF $Pr(\cdot, w, \sigma)$ used.*

Proof. First note that any possible world w' of Δ' is also a possible world of Δ and $\mathcal{I}(w') \in \mathcal{I}_I(w)$ where $w \in pw(\Delta)$. The possible worlds of Δ are the same as those of Δ'. Thus, we have that $Pr(E, w', \sigma)$ gives the same value independently if w is a possible world of Δ or Δ' and thus we have that the following formula holds as $\mathcal{I}(w') \in \mathcal{I}_I(w')$ for any world $w' \in pw(\Delta')$.

$$\sum_{\substack{w' \in pw(\Delta') \wedge \\ E \in \sigma(w) \wedge g \in E}} \mathcal{I}(w') \in \left[\sum_{\substack{w' \in pw(\Delta) \wedge \\ E \in \sigma(w') \wedge g \in E}} \mathcal{I}_I(w')[1], \sum_{\substack{w' \in pw(\Delta) \wedge \\ E \in \sigma(w') \wedge g \in E}} \mathcal{I}_I(w')[2] \right]^* \quad \square$$

A.5. Incomplete argumentation framework

Lemma 2. *For any iAF Δ and $\varphi \in \{arg, att, farg\}$, $af_{\Delta} : comp(\varphi(\Delta)) \rightarrow comp(\Delta)$ is a surjective function.*

Proof. Let $\Delta = \langle A, B, R, T\rangle$ be an iAF, we now show that $\forall \Lambda'' \in comp(\Delta)\ \exists \Lambda' \in comp(\varphi(\Delta))$ s.t. $af_{\Delta}(\Lambda') = \Lambda$.

Let $\Delta' = \varphi(\Delta) = \langle A^*, B^*, R^*, T^*\rangle$, $\Lambda'' \in comp(\Delta)$ be a completion of Δ, and $\Lambda' = \langle A', R'\rangle \in comp(\varphi(\Delta))$ a completion of Δ' such that:

- $(\varphi = arg)$ A' consists of all the arguments in Λ'' plus the certain arguments α_{ab} such that $(a, b) \in T$, and for each attack (a, b) in Λ'' which is uncertain in iAF Δ, we have that $\beta_{ab} \in A'$; moreover $R' = (A' \times A') \cap (R^* \cup T^*)$.
- $(\varphi = att)$ A' consists of all the arguments in $A \cup B$ plus the certain argument α. R' consists of all the attacks in Λ'' and, for each argument b which is uncertain in Δ and does not belong to Λ'', the attack (α, b).
- $(\varphi = farg)$ A' consists of all the arguments in $A \cup B$ plus the arguments $\alpha_{ab}, \beta_{ab}, \beta^c_{ab}$ such that $(a, b) \in T$, and for each argument b which is uncertain in Δ the argument b^c; also, for each argument b (resp. attack (a, b)) which is uncertain in Δ, if b (resp. (a, b)) is in Λ'', the argument b^u (resp. β^u_{ab}). Finally, $R' = (A' \times A') \cap (R^* \cup T^*)$.

By using the definition of function af given in the paper, we can now show that $af_{\Delta}(\Lambda') = \Lambda''$. By following the construction given above, we have the AF $\Lambda' = \langle A', R'\rangle$ where:

- $(\varphi = arg)$
  - $A' = A'' \cup \{\alpha_{ab} \mid (a, b) \in T\} \cup \{\beta_{ab} \mid (a, b) \in (R'' \cap T)\}$; and
  - $R' = (A' \times A') \cap (R^* \cup T^*)$.

  Then, using the definition of function af given in the paper, we can build the AF $af_{\Delta}(\Lambda') = \langle A^{\diamond}, R^{\diamond}\rangle$ as follows:
  - $A^{\diamond} = A \cup ((B \cap A') \setminus \emptyset) = A''$; and
  - $R^{\diamond} = (R \cap (A'' \times A''))\ \cup\ ((T \cap (A'' \times A'')) \setminus \{(a, b) \mid (\beta_{ab} \notin A') \vee (\beta^c_{ab} \in A' \wedge \beta^u_{ab} \notin A')\})$, that can be rewritten as: $R^{\diamond} = (R \cap (A'' \times A''))\ \cup\ ((T \cap (A'' \times A'')) \setminus \{(a, b)\ \mid\ \beta_{ab} \notin \{\beta_{ab} \mid (a, b) \in (R'' \cap T)\}\}$ or, equivalently, $R^{\diamond} = (R \cap (A'' \times A''))\ \cup\ ((T \cap (A'' \times A'')) \cap R'')$, stating that $R^{\diamond}$ contains, for any pair of arguments $\langle a, b\rangle$ of Λ'', an attack (a, b) iff either (a, b) is a certain attack of Δ or (a, b) is an uncertain attack of Δ that has been selected in R''. Thus, $af_{\Delta}(\Lambda') = \langle A^{\diamond}, R^{\diamond}\rangle = \langle A'', R''\rangle$.
- $(\varphi = att)$
  - $A' = A \cup B \cup \{\alpha\}$; and
  - $R' = R'' \cup \{(\alpha, b) \mid b \in (B \setminus A'')\}$.

  Then, using the definition of function af given in the paper, we can build the AF $af_{\Delta}(\Lambda') = \langle A^{\diamond}, R^{\diamond}\rangle$ as follows:
  - $A^{\diamond} = A \cup ((B \cap A') \setminus \{a \mid (\alpha, a) \in R'\})$, that can be rewritten as: $A^{\diamond} = A \cup (B \setminus \{a \mid (\alpha, a) \in \{(\alpha, b) \mid b \in (B \setminus A'')\}\})$ or, equivalently, $A^{\diamond} = A \cup (B \cap A'')$. As $A \subseteq A''$, it holds that $A^{\diamond} = A''$.

 - $R^\diamond = (R \cap (A'' \times A'')) \ \cup \ ((T \cap (A'' \times A'')) \setminus \{(a,b) \mid (\beta_{ab} \notin A') \vee (\beta^c_{ab} \in A' \wedge \beta^u_{ab} \notin A')\})$, that can be rewritten as: $R^\diamond = (R \cap (A'' \times A'')) \ \cup \ ((T \cap (A'' \times A'')) \setminus \emptyset)$, or simply, R''. Thus, $af_\Delta(\Lambda') = \langle A^\diamond, R^\diamond \rangle = \langle A'', R'' \rangle$.
- ($\varphi = farg$)
 - $A' = A \cup B \cup \{\alpha_{ab}, \beta_{ab}, \beta^c_{ab} \mid (a,b) \in T\} \cup \{b^c \mid b \in B\} \cup \{b^u \mid b \in B \cap A''\} \cup \{b^u_{ab} \mid (a,b) \in T \cap R''\}$; and
 - $R' = (A' \times A') \cap (R^* \cap T^*)$.

 Then, using the definition of function af given in the paper, we can build the AF $af_\Delta(\Lambda') = \langle A^\diamond, R^\diamond \rangle$ as follows:
 - $A^\diamond = A \cup ((B \cap A') \setminus \{a \mid a^u \notin A'\})$, that can be rewritten as $A^\diamond = A \cup (B \cap A'')$ (since $a^u \notin A'$ iff $a^u \notin (B \cap A'')$). As $A \subseteq A''$, it holds that $A^\diamond = A''$.
 - $R^\diamond = (R \cap (A'' \times A'')) \ \cup \ ((T \cap (A'' \times A'')) \setminus \{(a,b) \mid (\beta_{ab} \notin A') \vee (\beta^c_{ab} \in A' \wedge \beta^u_{ab} \notin A')\})$. Since $\beta_{ab} \notin A'$ (resp., $\beta^u_{ab} \notin A'$) iff $(a,b) \notin T$ (resp., $(a,b) \notin (T \cap R'')$), it is equivalent to:
 $R^\diamond = (R \cap (A'' \times A'')) \ \cup \ ((T \cap (A'' \times A'')) \setminus \{(a,b) \mid ((a,b) \notin T) \vee ((a,b) \in T \wedge (a,b) \notin (T \cap R''))\})$, that can be rewritten as $R^\diamond = (R \cap (A'' \times A'')) \ \cup \ ((T \cap (A'' \times A'')) \cap R'')\})$, or simply, R''. Thus, $af_\Delta(\Lambda') = \langle A^\diamond, R^\diamond \rangle = \langle A'', R'' \rangle$. $\square$

Theorem 11. *Let $\Delta = \langle A, B, R, T \rangle$ be an iAF, $\sigma \in \{\mathtt{gr}, \mathtt{co}, \mathtt{st}, \mathtt{pr}, \mathtt{sst}\}$, and $\varphi \in \{arg, att, farg\}$ and let $\Lambda \in comp(\Delta)$. Then,*

- *$comp(\Delta) = \{af_\Delta(\Lambda) \mid \Lambda \in comp(\varphi(\Delta))\}$, and*
- *$\sigma(\Lambda) = \{E \cap (A \cup B) \mid \Lambda' \in comp(\varphi(\Delta)) \ \wedge \ \Lambda = af_\Delta(\Lambda') \ \wedge \ E \in \sigma(\Lambda')\}$.*

Proof. The first item follows from the result of Lemma 2 that states that af_Δ is a surjective function, and thus its image coincides with its codomain.

As for the second item, first observe that there exists a surjection from $comp(\varphi(\Delta))$ to $comp(\Delta)$ (Lemma 2). Thus, for any $\Lambda \in comp(\Delta)$ there exists (at least one element) $\Lambda' \in comp(\varphi(\Delta))$, and we now show that they have the same complete extensions (modulo arguments/attacks added in the mapping).

In the following, let $\Lambda' \in comp(\varphi(\Delta))$ and $\Lambda \in comp(\Delta)$ be the AFs described in the proof of Lemma 2 for each mapping φ, respectively. We first show that their extensions are equivalent. Then, we show that any other $\Lambda'' \in comp(\varphi(\Delta))$ such that $af_\Delta(\Lambda'') = \Lambda$ and $\Lambda'' \neq \Lambda'$ the result still holds.

For $\varphi = arg$ the complete extensions of Λ are equal to those of Λ' if considering only the arguments in $A \cup B$. For Λ containing a given set of attacks $(a,b) \in T$, Λ' contains the corresponding set of arguments $\beta_{ab} \in B'$. An argument b is in a complete extension of Λ (resp. Λ') only if a (resp. β_{ab}) is not; we can reason analogously for (a,b) (resp. β_{ab}) not in Λ (resp. Λ'). The status of arguments b in Λ and Λ' is determined by possibly different arguments but having equivalent status w.r.t. a complete extension of Λ and the corresponding one of Λ'.

For $\varphi = att$, any uncertain argument $a \in B$ generates two sets of completions of Λ, S_1 and S_2, including or not including a. Δ' generates S'_1/S'_2, equivalent to S_1/S_2 but excluding or not excluding attack (α, a). The admissible sets of elements of S_1 and S'_1 are equal if ignoring α argument (that is admissible as not attacked). An element of S'_2 is such that a is attacked by α and not defended. As α is in all extensions, then every attack from a is defended too, so S_2 generates the same admissible extensions as that of S'_2 modulo α argument.

Finally, for $\varphi = farg$, AF Λ containing (resp. not containing) a given set of uncertain arguments $b \in B$ is equivalent to Λ' containing (resp. not containing) the corresponding arguments b^u. If b is in a complete extension of Λ, b is in a complete extension of Λ' because it is defended by b^u (that is not attacked by any argument). If b is not in a complete extension of Λ, b is not in a complete extension of Λ' attacked by b_c without being defended, meaning that attacks from b in Λ' can be omitted. Then, extensions are equal modulo meta-arguments added in the transformation.

We now need to prove that for any other $\Lambda'' \in comp(\varphi(\Delta))$ such that $af_\Delta(\Lambda'') = \Lambda$ and $\Lambda'' \neq \Lambda'$ the result continues to hold.

For $\varphi = arg$, observe that Λ'' might only differ from Λ' for β_{xy} arguments. Thus, it might be the case that there exists an argument β_{ab} in Λ'' even if a is not an argument of Λ. However β_{ab} would never appear in any complete extension of Λ'' as it is attacked by α_{ab} which in turn can only be attacked by a. As a consequence of the fact that a is not in Λ'', the status of b does not depend on that of β_{ab}. As observed earlier, an argument b is in a complete extension of Λ (resp. Λ'') only if a (resp. β_{ab}) is not, and thus, the complete extensions of Λ'' are still equal modulo meta-arguments to that of Λ'.

For $\varphi = att$, observe that Λ'' might only differ from Λ' for the presence of uncertain attacks (a,b) of Δ even if a is not in Λ. In this case, note that (i) a would never appear in any complete extension of Λ'' as it is attacked by α by construction of $af_\Delta(\Lambda'')$. As a consequence of the fact that α is not attacked in Λ'', a does never appear in any complete extension of Λ''. Thus, the complete extensions of Λ'' are equal modulo meta-arguments to that of Λ' and the attack from a to b has no effect on the status of b.

In the case of $\varphi = farg$, observe that Λ'' might differ from Λ' for β^u_{xy} arguments. Thus, it might be the case that there exists an argument β^u_{ab} in Λ'' even if a is not an argument of Λ. However, β_{ab} would never appear in any complete extension of Λ'' as it is attacked by α_{ab} which in turn can only be attacked by a. Thus, the complete extensions of Λ'' are equal modulo meta-arguments to that of Λ' (again, attack from a to b has no effect on the status of b).

As complete extensions define a complete-meet semilattice, the results hold for the other semantics obtained by putting restrictions on the sets of complete extensions, that is $\sigma \in \{\mathtt{gr}, \mathtt{st}, \mathtt{pr}, \mathtt{sst}\}$. $\square$

Proposition 5. *For any arg-iAF Δ, $comp(\Delta) = \{\Lambda \mid \Lambda \in pw(\Delta^p) \wedge \mathcal{I}(\Lambda) > 0\}$.*

Proof. Let $\Delta = \langle A, B, R, \emptyset\rangle$. The result follows by observing that the power-set of B corresponds to the sample space of the events associated with Δ^p having probability greater than 0. $\square$

Corollary 5. *For $\sigma \in \{\texttt{gr}, \texttt{co}, \texttt{pr}, \texttt{st}, \texttt{sst}\}$, $\mathsf{PrA}[\sigma]$ is $\mathrm{FP}^{\#\mathrm{P}}$-hard, even for acyclic arg-iAF and for any chosen PDF.*

Proof. The construction given in the proof of Theorem 2 can be used here since that hardness result is shown in the specific case of arguments having probability set to 1/2. Thus, we can show a reduction starting from an acyclic arg-iAF whose derived PrAF is that in the proof of Theorem 2. $\square$

Corollary 6. *Consider a semantics $\sigma \in \{\texttt{gr}, \texttt{co}, \texttt{pr}, \texttt{st}, \texttt{sst}\}$. Unless NP $\subseteq$ BPP, there is no FPRAS for $\mathsf{PrA}[\sigma]$, even for acyclic arg-iAFs and for any chosen PDF.*

Proof. The claim follows from the proofs of Theorem 2 and Theorem 4. In fact, we can show a reduction from #P2CNF to $\mathsf{PrA}[\sigma]$, for every semantics $\sigma \in \{\texttt{gr}, \texttt{co}, \texttt{pr}, \texttt{st}, \texttt{sst}\}$ and for any chosen PDF. Starting from a P2CNF formula, we can define an acyclic arg-iAF Δ such that the derived PrAF Δ^p is that of the proof of Theorem 2 (where uncertain arguments correspond to arguments having probability 1/2). Given this, it suffices to reason as in the proof Theorem 4 where the result is shown for PrAF in order to prove the claim for iAF. $\square$

Corollary 7. *Let $\sigma \in \{\texttt{pr}, \texttt{st}, \texttt{sst}\}$. Unless NP $\subseteq$ BPP, there is no FPARAS for $\mathsf{PrA}[\sigma]$, for any chosen PDF.*

Proof. The claim follows by reasoning as in the proof of Theorem 5, where a PrAF with all probabilities set to 1 is used, that is, it works for arg-iAFs even with no uncertain argument (see also Corollary 3). $\square$

Corollary 8. *Problem $\mathsf{PrEA}[\sigma]$, with $\sigma \in \{\texttt{gr}, \texttt{co}, \texttt{pr}, \texttt{st}, \texttt{sst}\}$, has an FPARAS if either i) $\sigma = \texttt{gr}$, or ii) the input arg-iAF has no odd cycles.*

Proof. The claim follows from Theorem 7 since if conditions i) or ii) hold for arg-iAF Δ, then they hold for the derived PrAF Δ^p used in Definition 17. $\square$

Proposition 6. *For any arg-iAF Δ and argument goal g, we have that:*

- *$PCA_\sigma(\Delta, g)$ is false iff $PrA^\sigma_{\Delta^p}(g) = 0$;*
- *$NSA_\sigma(\Delta, g)$ is true iff $PrA^\sigma_{\Delta^p}(g) = 1$.*

Proof. $PCA_\sigma(\Delta, g)$ is false iff $\forall \Lambda \in comp(\Delta)$, $\nexists E \in \sigma(\Lambda)$ s.t. $g \in E$, that is, using Proposition 5, $\forall \Lambda \in pw(\Delta^p)$, $\nexists E \in \sigma(\Lambda)$ s.t. $g \in E$, meaning that $PrA^\sigma_{\Delta^p}(g) = 0$ (cf. Definition 3). $NSA_\sigma(\Delta, g)$ is true iff $\forall \Lambda \in comp(\Delta)$, $\forall E \in \sigma(\Lambda)$, $g \in E$, that is, $\forall \Lambda \in pw(\Delta^p)$, $\forall E \in \sigma(\Lambda)$, $g \in E$, and thus $PrA^\sigma_{\Delta^p}(g) = 1$ as Pr is a PDF. $\square$

Proposition 7. *Let $\Delta = \langle A, B, R, \emptyset\rangle$ be an acyclic arg-iAF and g a goal. It holds that:*

- *$PSA_\sigma(\Delta, g) \equiv PCA_\sigma(\Delta, g)$ and $NSA_\sigma(\Delta, g) \equiv NCA_\sigma(\Delta, g)$;*
- *$PSA_\sigma(\Delta, g)$ is true iff $PrA^\sigma_{\Delta^p}(g) \geq \frac{1}{2^{|B|}}$;*
- *$NSA_\sigma(\Delta, g)$ is false iff $PrA^\sigma_{\Delta^p}(g) \leq 1 - \frac{1}{2^{|B|}}$.*

Proof. Since Δ is acyclic, each $\Lambda \in comp(\Delta)$ has exactly one complete extension, thus $SA_\sigma(\Lambda, g) \equiv CA_\sigma(\Lambda, g)$, from which it follows that $PSA_\sigma(\Delta, g) \equiv PCA_\sigma(\Delta, g)$ and $NSA_\sigma(\Delta, g) \equiv NCA_\sigma(\Delta, g)$. Moreover, $PSA_\sigma(\Delta, g)$ is true iff there is $\Lambda \in comp(\Delta)$ and its unique extension contains g, or equivalently, there is a non-zero probability world $\Lambda \in pw(\Delta^p)$ and its unique extension E contains g, i.e. $Pr(E, \Lambda, \sigma) = 1$ (cf. Definition 3). Since every non-zero probability world Λ has probability $\frac{1}{2^{|B|}}$, $PrA^\sigma_{\Delta^p}(g) \geq \frac{1}{2^{|B|}}$. Reasoning analogously, $NSA_\sigma(\Delta, g)$ is false iff there is a non-zero probability world $\Lambda \in pw(\Delta^p)$ and its unique extension does not contain g, from which it follows that $PrA^\sigma_{\Delta^p}(g) \leq 1 - \frac{1}{2^{|B|}}$. $\square$

References

[1] G. Alfano, M. Calautti, S. Greco, F. Parisi, I. Trubitsyna, Explainable acceptance in probabilistic abstract argumentation: complexity and approximation, in: Proc. of the 17th Int. Conf. on Principles of Knowledge Representation and Reasoning (KR), 2020, pp. 33–43.
[2] T. Bench-Capon, P.E. Dunne, Argumentation in artificial intelligence, Artif. Intell. 171 (2007) 619–641.
[3] G.R. Simari, I. Rahwan (Eds.), Argumentation in Artificial Intelligence, Springer, 2009.
[4] K. Atkinson, P. Baroni, M. Giacomin, A. Hunter, H. Prakken, C. Reed, G.R. Simari, M. Thimm, S. Villata, Towards artificial argumentation, AI Mag. 38 (3) (2017) 25–36.
[5] K. Atkinson, T.J.M. Bench-Capon, Argumentation schemes in AI and law, Argument Comput. 12 (3) (2021) 417–434.

[6] L. Amgoud, H. Prade, Using arguments for making and explaining decisions, Artif. Intell. 173 (3–4) (2009) 413–436.
[7] T.J.M. Bench-Capon, K. Atkinson, A.Z. Wyner, Using argumentation to structure e-participation in policy making, in: Transactions on Large-Scale Data and Knowledge-Centered Systems, vol. 18, 2015, pp. 1–29.
[8] M. Snaith, R.Ø. Nielsen, S.R. Kotnis, A. Pease, Ethical challenges in argumentation and dialogue in a healthcare context, Argument Comput. 12 (2) (2021) 249–264.
[9] N. Kökciyan, I. Sassoon, E. Sklar, S. Modgil, S. Parsons, Applying metalevel argumentation frameworks to support medical decision making, IEEE Intell. Syst. 36 (2) (2021) 64–71.
[10] A. Pazienza, D. Grossi, F. Grasso, R. Palmieri, M. Zito, S. Ferilli, An abstract argumentation approach for the prediction of analysts' recommendations following earnings conference calls, Intell. Artif. 13 (2) (2019) 173–188.
[11] M.E.B. Brarda, L.H. Tamargo, A.J. García, Using argumentation to obtain and explain results in a decision support system, IEEE Intell. Syst. 36 (2) (2021) 36–42.
[12] N. Kökciyan, N. Yaglikci, P. Yolum, An argumentation approach for resolving privacy disputes in online social networks, ACM Trans. Internet Technol. 17 (3) (2017) 27:1–27:22.
[13] P.M. Dung, On the acceptability of arguments and its fundamental role in nonmonotonic reasoning, logic programming and n-person games, Artif. Intell. 77 (1995) 321–358.
[14] A. Hunter, S. Polberg, N. Potyka, T. Rienstra, M. Thimm, Probabilistic argumentation: a survey, in: Handbook of Formal Argumentation, vol. 2, 2021, pp. 397–441.
[15] G.F. Georgakopoulos, D.J. Kavvadias, C.H. Papadimitriou, Probabilistic satisfiability, J. Complex. 4 (1) (1988) 1–11.
[16] N.J. Nilsson, Probabilistic logic revisited, Artif. Intell. 59 (1–2) (1993) 39–42.
[17] F. Riguzzi, T. Swift, A survey of probabilistic logic programming, in: M. Kifer, Y.A. Liu (Eds.), Declarative Logic Programming: Theory, Systems, and Applications, ACM / Morgan & Claypool, 2018, pp. 185–228.
[18] D. Suciu, D. Olteanu, C. Ré, C. Koch, Probabilistic Databases, Synthesis Lectures on Data Management, Morgan & Claypool Publishers, 2011.
[19] P.M. Dung, P.M. Thang, Towards (probabilistic) argumentation for jury-based dispute resolution, in: Proc. of Int. Conf. on Computational Models of Argument (COMMA), 2010, pp. 171–182.
[20] T. Rienstra, Towards a probabilistic Dung-style argumentation system, in: Proc. of Int. Conf. on Agreement Technologies (AT), 2012, pp. 138–152.
[21] D. Doder, S. Woltran, Probabilistic argumentation frameworks–a logical approach, in: Proc. of Int. Conf. on Scalable Uncertainty Management (SUM), 2014, pp. 134–147.
[22] A. Hunter, Some foundations for probabilistic abstract argumentation, in: Proc. of Int. Conf. on Computational Models of Argument (COMMA), 2012, pp. 117–128.
[23] H. Li, N. Oren, T.J. Norman, Probabilistic argumentation frameworks, in: Proc. of Int. Workshop on Theory and Applications of Formal Argumentation (TAFA), 2011, pp. 1–16.
[24] B. Fazzinga, S. Flesca, F. Parisi, On the complexity of probabilistic abstract argumentation frameworks, ACM Trans. Comput. Log. 16 (3) (2015) 22:1–22:39.
[25] B. Fazzinga, S. Flesca, F. Parisi, On efficiently estimating the probability of extensions in abstract argumentation frameworks, Int. J. Approx. Reason. 69 (2016) 106–132.
[26] B. Fazzinga, S. Flesca, F. Furfaro, Complexity of fundamental problems in probabilistic abstract argumentation: beyond independence, Artif. Intell. 268 (2019) 1–29.
[27] N. Potyka, A polynomial-time fragment of epistemic probabilistic argumentation, Int. J. Approx. Reason. 115 (2019) 265–289.
[28] R. Riveret, N. Oren, G. Sartor, A probabilistic deontic argumentation framework, Int. J. Approx. Reason. 126 (2020) 249–271.
[29] P. Dondio, Toward a computational analysis of probabilistic argumentation frameworks, Cybern. Syst. 45 (3) (2014) 254–278.
[30] A. Hunter, Probabilistic qualification of attack in abstract argumentation, Int. J. Approx. Reason. 55 (2) (2014) 607–638.
[31] S. Polberg, A. Hunter, Empirical evaluation of abstract argumentation: supporting the need for bipolar and probabilistic approaches, Int. J. Approx. Reason. 93 (2018) 487–543.
[32] A. Hunter, A probabilistic approach to modelling uncertain logical arguments, Int. J. Approx. Reason. 54 (1) (2013) 47–81.
[33] B. Fazzinga, S. Flesca, F. Furfaro, Credulous and skeptical acceptability in probabilistic abstract argumentation: complexity results, Intell. Artif. 12 (2) (2018) 181–191.
[34] B. Moulin, H. Irandoust, M. Bélanger, G. Desbordes, Explanation and argumentation capabilities: towards the creation of more persuasive agents, Artif. Intell. Rev. 17 (3) (2002) 169–222.
[35] T. Miller, Explanation in artificial intelligence: insights from the social sciences, Artif. Intell. 267 (2019) 1–38.
[36] X. Fan, F. Toni, On computing explanations in argumentation, in: Proc. of AAAI Conf. on Artificial Intelligence, 2015, pp. 1496–1502.
[37] M. Ulbricht, J.P. Wallner, Strong explanations in abstract argumentation, in: Proc. of Thirty-Fifth AAAI Conf. on Artificial Intelligence, 2021, pp. 6496–6504.
[38] G. Brewka, M. Ulbricht, Strong explanations for nonmonotonic reasoning, in: Description Logic, Theory Combination, and All That - Essays Dedicated to Franz Baader on the Occasion of His 60th Birthday, vol. 11560, 2019, pp. 135–146.
[39] Z.G. Saribatur, J.P. Wallner, S. Woltran, Explaining non-acceptability in abstract argumentation, in: Proc. of the 24th Eur. Conf. on Artificial Intelligence (ECAI), vol. 325, 2020, pp. 881–888.
[40] D. Baumeister, D. Neugebauer, J. Rothe, H. Schadrack, Verification in incomplete argumentation frameworks, Artif. Intell. 264 (2018) 1–26.
[41] M. Caminada, Semi-stable semantics, in: Proc. of Computational Models of Argument (COMMA), vol. 144, 2006, pp. 121–130.
[42] P. Baroni, M. Caminada, M. Giacomin, An introduction to argumentation semantics, Knowl. Eng. Rev. 26 (4) (2011) 365–410.
[43] T. Mantadelis, S. Bistarelli, Probabilistic abstract argumentation frameworks, a possible world view, Int. J. Approx. Reason. 119 (2020) 204–219.
[44] M. Thimm, A probabilistic semantics for abstract argumentation, in: Proc. of Eur. Conf. on Artificial Intelligence (ECAI), 2012, pp. 750–755.
[45] R. Baumann, M. Ulbricht, Choices and their consequences - explaining acceptable sets in abstract argumentation frameworks, in: Proc. of the 18th Int. Conf. on Principles of Knowledge Representation and Reasoning (KR), 2021, pp. 110–119.
[46] S. Arora, B. Barak, Computational Complexity - a Modern Approach, Cambridge University Press, 2009.
[47] K. Ko, Some observations on the probabilistic algorithms and NP-hard problems, Inf. Process. Lett. 14 (1) (1982) 39–43.
[48] P.E. Dunne, T.J.M. Bench-Capon, Coherence in finite argument systems, Artif. Intell. 141 (1/2) (2002) 187–203.
[49] W. Hoeffding, Probability inequalities for sums of bounded random variables, J. Am. Stat. Assoc. 58 (301) (1963) 13–30.
[50] P. Walley, Statistical Reasoning with Imprecise Probabilities, Chapman & Hall, 1991.
[51] R.T. Ng, V.S. Subrahmanian, Probabilistic logic programming, Inf. Comput. 101 (2) (1992) 150–201.
[52] V.S. Subrahmanian, Probabilistic databases and logic programming, in: P. Codognet (Ed.), Logic Programming, 17th International Conference, ICLP 2001, Paphos, Cyprus, November 26 - December 1, 2001, Proceedings, in: Lecture Notes in Computer Science, vol. 2237, Springer, 2001, p. 10.
[53] J. Grant, F. Parisi, A. Parker, V.S. Subrahmanian, An agm-style belief revision mechanism for probabilistic spatio-temporal logics, Artif. Intell. 174 (1) (2010) 72–104.
[54] J. Grant, C. Molinaro, F. Parisi, Probabilistic spatio-temporal knowledge bases: capacity constraints, count queries, and consistency checking, Int. J. Approx. Reason. 100 (2018) 1–28.

[55] D. Baumeister, M. Järvisalo, D. Neugebauer, A. Niskanen, J. Rothe, Acceptance in incomplete argumentation frameworks, Artif. Intell. 295 (2021) 103470.
[56] B. Fazzinga, S. Flesca, F. Furfaro, Revisiting the notion of extension over incomplete abstract argumentation frameworks, in: Proc. of Int. Joint Conf. on Artificial Intelligence (IJCAI), 2020, pp. 1712–1718.
[57] G. Alfano, S. Greco, F. Parisi, I. Trubitsyna, Incomplete argumentation frameworks: properties and complexity, in: Proc. of AAAI Conf. on Artificial Intelligence, 2022, pp. 5451–5460.
[58] D. Baumeister, D. Neugebauer, J. Rothe, Credulous and skeptical acceptance in incomplete argumentation frameworks, in: Proc. of Int. Conf. on Computational Models of Argument (COMMA), 2018, pp. 181–192.
[59] P. Baroni, M. Giacomin, G. Guida, SCC-recursiveness: a general schema for argumentation semantics, Artif. Intell. 168 (1–2) (2005) 162–210.
[60] F. Cerutti, M. Giacomin, M. Vallati, M. Zanella, An SCC recursive meta-algorithm for computing preferred labellings in abstract argumentation, in: Proc. of Int. Conf. on Principles of Knowledge Representation and Reasoning (KR), 2014, pp. 42–51.
[61] P. Baroni, G. Boella, F. Cerutti, M. Giacomin, L.W.N. van der Torre, S. Villata, On the input/output behavior of argumentation frameworks, Artif. Intell. 217 (2014) 144–197.
[62] T. Rienstra, M. Thimm, B. Liao, L.W.N. van der Torre, Probabilistic abstract argumentation based on SCC decomposability, in: Proc. of Int. Conf. on Principles of Knowledge Representation and Reasoning (KR), 2018, pp. 168–177.
[63] G. Brewka, M. Thimm, M. Ulbricht, Strong inconsistency, Artif. Intell. 267 (2019) 78–117.
[64] F. Bex, D. Walton, Combining explanation and argumentation in dialogue, Argument Comput. 7 (1) (2016) 55–68.
[65] K. Cyras, D. Birch, Y. Guo, F. Toni, R. Dulay, S. Turvey, D. Greenberg, T. Hapuarachchi, Explanations by arbitrated argumentative dispute, Expert Syst. Appl. 127 (2019) 141–156.
[66] R. Craven, F. Toni, Argument graphs and assumption-based argumentation, Artif. Intell. 233 (2016) 1–59.
[67] P.M. Dung, R.A. Kowalski, F. Toni, Assumption-based argumentation, in: Argumentation in Artificial Intelligence, Springer, 2009, pp. 199–218.
[68] N.D. Hung, Computing probabilistic assumption-based argumentation, in: Proc. of the Pacific Rim Int. Conf. on Artificial Intelligence, (PRICAI), vol. 9810, Springer, 2016, pp. 152–166.
[69] P.M. Dung, P. Mancarella, F. Toni, Computing ideal sceptical argumentation, Artif. Intell. 171 (10–15) (2007) 642–674.
[70] P.M. Thang, P.M. Dung, N.D. Hung, Towards a common framework for dialectical proof procedures in abstract argumentation, J. Log. Comput. 19 (6) (2009) 1071–1109.
[71] P.E. Dunne, M. Wooldridge, Complexity of abstract argumentation, in: Argumentation in Artificial Intelligence, Springer, 2009, pp. 85–104.
[72] W. Dvorák, M. Järvisalo, J.P. Wallner, S. Woltran, Complexity-sensitive decision procedures for abstract argumentation, Artif. Intell. 206 (2014) 53–78.
[73] M. Kröll, R. Pichler, S. Woltran, On the complexity of enumerating the extensions of abstract argumentation frameworks, in: Proc. of Int. Joint Conf. on Artificial Intelligence (IJCAI), 2017, pp. 1145–1152.
[74] G. Alfano, S. Greco, F. Parisi, An efficient algorithm for skeptical preferred acceptance in dynamic argumentation frameworks, in: Proc. of Int. Joint Conf. on Artificial Intelligence (IJCAI), 2019, pp. 18–24.
[75] B. Fazzinga, S. Flesca, F. Parisi, A. Pietramala, Computing or estimating extensions' probabilities over structured probabilistic argumentation frameworks, FLAP 3 (2) (2016) 177–200.
[76] S. Bistarelli, F. Rossi, F. Santini, A novel weighted defence and its relaxation in abstract argumentation, Int. J. Approx. Reason. 92 (2018) 66–86.
[77] L. Amgoud, D. Doder, S. Vesic, Evaluation of argument strength in attack graphs: foundations and semantics, Artif. Intell. 302 (2022) 103607.
[78] G.I. Simari, P. Shakarian, M.A. Falappa, A quantitative approach to belief revision in structured probabilistic argumentation, Ann. Math. Artif. Intell. 76 (3–4) (2016) 375–408.
[79] P. Shakarian, G.I. Simari, G. Moores, D. Paulo, S. Parsons, M.A. Falappa, A. Aleali, Belief revision in structured probabilistic argumentation - model and application to cyber security, Ann. Math. Artif. Intell. 78 (3–4) (2016) 259–301.
[80] T. Alsinet, R. Béjar, L. Godo, F. Guitart, RP-DeLP: a weighted defeasible argumentation framework based on a recursive semantics, J. Log. Comput. 26 (4) (2016) 1315–1360.
[81] H. Prakken, An abstract framework for argumentation with structured arguments, Argument Comput. 1 (2) (2010) 93–124.
[82] F. Toni, A tutorial on assumption-based argumentation, Argument Comput. 5 (1) (2014) 89–117.
[83] C. Schulz, F. Toni, Labellings for assumption-based and abstract argumentation, Int. J. Approx. Reason. 84 (2017) 110–149.
[84] K. Cyras, Q. Heinrich, F. Toni, Computational complexity of flat and generic assumption-based argumentation, with and without probabilities, Artif. Intell. 293 (2021) 103449.
[85] G. Alfano, S. Greco, F. Parisi, I. Trubitsyna, On the semantics of abstract argumentation frameworks: a logic programming approach, Theory Pract. Log. Program. 20 (5) (2020) 703–718.
[86] L.D. Raedt, A. Kimmig, H. Toivonen, Problog: a probabilistic prolog and its application in link discovery, in: Proceedings of the 20th International Joint Conference on Artificial Intelligence (IJCAI), 2007, pp. 2462–2467.
[87] L.D. Raedt, A. Kimmig, Probabilistic (logic) programming concepts, Mach. Learn. 100 (1) (2015) 5–47.
[88] A. Dries, A. Kimmig, W. Meert, J. Renkens, G.V. den Broeck, J. Vlasselaer, L.D. Raedt, Problog2: probabilistic logic programming, in: Proceeding of the European Conference on Machine Learning and Knowledge Discovery in Databases ECML PKDD, in: Lecture Notes in Computer Science, vol. 9286, 2015, pp. 312–315.
[89] G. Alfano, S. Greco, F. Parisi, Efficient computation of extensions for dynamic abstract argumentation frameworks: an incremental approach, in: Proc. of Int. Joint Conf. on Artificial Intelligence (IJCAI), 2017, pp. 49–55.
[90] T. Rienstra, C. Sakama, L. van der Torre, B. Liao, A principle-based robustness analysis of admissibility-based argumentation semantics, Argument Comput. 11 (3) (2020) 305–339.
[91] F. Welsh, A. Gale, The complexity of counting problems, in: Aspects of Complexity, De Gruyter, 2001.
[92] P. Baroni, M. Giacomin, On principle-based evaluation of extension-based argumentation semantics, Artif. Intell. 171 (10–15) (2007) 675–700.