A Preliminary Report on Probabilistic Attack Normal Form for Constellation Semantics

Theofrastos Mantadellis\textsuperscript{a,1} and Stefano Bistarelli\textsuperscript{a,b}

\textsuperscript{a} Department of Mathematics and Computer Science, University of Perugia, Italy
\textsuperscript{b} Member of the INdAM Research group GNCS

Abstract. After Dung’s founding work in Abstract Argumentation Frameworks there has been a growing interest in extending the Dung’s semantics in order to describe more complex or real life situations. Several of these approaches take the direction of weighted or probabilistic extensions. One of the most prominent probabilistic approaches is that of constellation Probabilistic Abstract Argumentation Frameworks from Li et al.

In this paper, we present a normal form for constellation probabilistic abstract argumentation frameworks. Furthermore, we present a transformation from general constellation probabilistic abstract argumentation frameworks to the presented normal form. In this way we illustrate that the simpler normal form has equal representation power with the general one.

Keywords. Probabilistic Abstract Argumentation Frameworks, Constellation Semantics, Normal Forms

1. Introduction

Argumentation is an everyday method of humanity to discuss and solve myriad different situations where opinions or point of views conflict. Abstract Argumentation Frameworks \cite{9} (AAFs) aim in modeling everyday situations where information is inconsistent or incomplete. Many different extensions of AAFs from Dung’s pioneering work have appeared in order to describe different everyday situations. Sample works includes assumption based argumentation \cite{4}, extending AAFs with support \cite{21}, introducing labels \cite{26}. Other approaches of extending AAFs have focused on introducing weights in elements of the AAF, such as \cite{23}. Such approaches are powerful tools to model voting systems, belief in arguments and argument strength.

Knowledge representation with the use of probabilistic information has been used in many areas of computer science. Probabilistic information, is a powerful medium to represent knowledge. Similarly, many researchers have extended AAFs by adding probabilistic information. These very prominent extensions of AAFs have been categorized in two big groups by Hunter \cite{15}: the \textit{epistemic} approaches and the \textit{constellation} approaches.

\textsuperscript{1}Project: TBD
The epistemic approaches, such as those presented in [24,14] describe probabilistic AAFs that the uncertainty does not alter the structure of the AAFs. These type of AAFs use the probability assignments to quantify the existing uncertainty of arguments in AAFs and not to introduce new uncertainty.

The constellation approaches, such as those presented in [16,10,8,17] introduce probabilistic elements in the AAF in such a way that the structure of the AAF becomes uncertain. The constellation approaches generate a set of AAFs with a probabilistic distribution and as such define a probabilistic distribution over the extensions of those AAFs.

In this paper we focus on the constellation approach from Li et al. [16]. [16] introduced probabilistic elements to the structure of AAFs, resulting to a set of AAFs. This allows for a set of arguments to be an (admissible, stable, ground, etc.) extension in some of the AAFs that are represented by the constellation. This simple but yet powerful representation has the ability to represent naturally many different uncertain scenarios.

The works of [16] and [13] can be considered as the pioneering work on combining probabilities with AAFs. Our motivation for this work is to present a normal form for PrAAFs that while simpler has the same expression power with general PrAAFs; we also present a PrAAF to PrAAF transformation that allows any general PrAAF to be converted to an equivalent PrAAF in normal form.

2. Preliminaries

2.1. Abstract Argumentation

An abstract argumentation framework [9] (AAF) is a tuple \( \text{AAF} = (\text{Args}, \text{Atts}) \) where \( \text{Args} \) is a set of arguments and \( \text{Atts} \) a set of attacks among arguments of the form of a binary relation \( \text{Atts} \subseteq \text{Args} \times \text{Args} \). For arguments \( a, b \in \text{Args} \), we use \( a \rightarrow b \) as a shorthand to indicate \((a, b) \in \text{Atts}\) and we say that argument \( a \) attacks argument \( b \). Figure 1 illustrates an example AAF.

![Figure 1. Example AAF \( \{a, b, c, d\}, \{a \rightarrow c, b \rightarrow c, c \rightarrow d\} \). Arguments are represented as cycles and attacks as arrows. Arguments \( a, b \) are attacking argument \( c \) which attacks argument \( d \).](image)

A set of arguments \( S \subseteq \text{Args} \) is said to be conflict-free iff \( \not\exists a, b \in S \) where \( a \rightarrow b \in \text{Atts} \). An argument \( a \in \text{Args} \) is acceptable with respect to set \( S \subseteq \text{Args} \) if no argument attack \( a \) or if \( \forall b \in \text{Args} \) that \( \exists b \rightarrow a \in \text{Atts} \) then \( \exists c \in S \) where \( c \rightarrow b \in \text{Atts} \).

Given the above [9] gives semantics to AAF by the use of extensions over subsets of arguments. Dung first defines the admissible semantics. A set \( S \subseteq \text{Args} \) is admissible iff \( S \) is conflict free and each \( a \in S \) is acceptable with respect to \( S \). Following our example AAF from Figure 1 the set \( \{a, b, d\} \) is admissible. Over time several different semantics have been discussed such as complete, preferred, grounded, stable [9], semi-stable [6], CF2 [12] etc.
2.2. Constellation based Probabilistic Abstract Argumentation Frameworks

Hunter [13], categorizes probabilistic abstract argumentation frameworks (PrAAFs) in two different categories: the constellation and the epistemic PrAAFs. For this paper we will focus on the constellation approaches and we base our work in the definition of PrAAFs by [16].

A constellation approach to PrAAFs defines probabilities over the structure of the AAF graph. One can assign probabilities to either the arguments or/and attacks of the AAF. We refer to arguments/attacks with assigned probabilities less than 1 as probabilistic arguments/attacks and we refer as probabilistic elements to either probabilistic arguments or probabilistic attacks.

A probabilistic element \( e \) exists in an AAF with probability \( P(e) \). These probabilistic elements correspond to random variables, which are assumed to be mutually independent\(^2\). As such, a PrAAF defines a probability distribution over a set of AAFs.

![Figure 2. Example PrAAF (\( \{a, b, c, d\}, \{1, 1, 0.4, 1\}, \{a \rightarrow c, b \rightarrow c, c \rightarrow d\}, \{0.3, 0.7, 1\} \)). Arguments are represented as cycles and attacks as arrows.](image)

**Definition 1.** Formally, a PrAAF is a tuple \( \text{PrAAF} = (\text{Args}, P_{\text{Args}}, \text{Atts}, P_{\text{Atts}}) \) where \( \text{Args}, \text{Atts} \) define an AAF, \( P_{\text{Args}} \) is a set of probabilities for each \( a \in \text{Args} \) with \( 0 < P_{\text{Args}}(a) \leq 1 \) and \( P_{\text{Atts}} \) is a set of probabilities for each \( \rightarrow \in \text{Atts} \) with \( 0 < P_{\text{Atts}}(\rightarrow) \leq 1 \).

Finally, stating an argument or an attack having probability 0 is redundant. A probabilistic argument or attack with 0 probability is an argument or attack that is not part of any AAF that the constellation represents. Figure 2 illustrates an example PrAAF with 3 different probabilistic elements.

As mentioned a PrAAF defines a probability distribution for all the possible non-probabilistic AAFs it contains. Each single possible set of probabilistic elements (arguments or attacks) of the PrAAF is called a possible world. Table 1 presents all possible worlds for the example PrAAF of Figure 2. One can notice that having only three different probabilistic elements it generates eight possible worlds. The possible worlds of a PrAAF are exponential in the number of probabilistic elements (\( 2^N \) where \( N \) the number of probabilistic elements).

**Definition 2 (Probability of Possible World).** The probability of a possible world equals to the product of the probability of each probabilistic element that is in the possible world with the product of one minus the probability of each probabilistic element that is excluded from the possible world.

\(^2\)As we are going to present later in the paper, the structure of AAF might impose dependencies among otherwise assumed independent probabilistic elements.
more, states that the probabilities Li et al. [16] further restrict the possible worlds to what is called induced AAFs. Further-
ections appear as a separate definition, formally:

\[
P_{world} = \prod_{e_i \in \text{AAF}_{world}} P(e_i) \cdot \prod_{e_j \notin \text{AAF}_{world}} (1 - P(e_j))
\]

There has been extensive research on how to compute the probabilities without explicitly enumerating all possible worlds. We point out some works from Probabilistic Logic Programming that is closely related with PrAAFs. For efficient exact inference we direct the reader at [18, 22, 15, 11, 23], and for efficient approximate inference at [5, 25, 19].

2.3. Imposing Restrictions

Li et al. [15] further restrict the possible worlds to what is called induced AAFs. Furthermore, states that the probabilities \( P_{atts} \) are conditional probabilities instead the likelihood of existence. These restrictions appear as a separate definition, formally:

**Definition 3** (Inducing an AAF from a PrAAF). An AAF \( \text{AAF}_{\text{ind}}, \text{atts}_{\text{ind}} \) is said to be induced from a PrAAF \( \text{AAF} (\text{args}, P_{\text{args}}, \text{atts}, P_{\text{atts}}) \) iff all of the following hold:

1. \( \text{args}_{\text{ind}} \subseteq \text{args} \)
2. \( \text{atts}_{\text{ind}} \subseteq \text{atts} \cap (\text{args}_{\text{ind}} \times \text{args}_{\text{ind}}) \)
3. \( \forall a \in \text{args} \text{ such that } P_{\text{args}}(a) = 1, a \in \text{args}_{\text{ind}} \)
4. \( \forall a_1 \rightarrow a_2 \in \text{atts} \text{ such that } P_{\text{atts}}(a_1 \rightarrow a_2) = 1 \text{ and } P_{\text{args}}(a_1) = P_{\text{args}}(a_2) = 1, a_1 \rightarrow a_2 \in \text{atts}_{\text{ind}} \)

| AAF          | Possible World                                                                 | Prob. | Admissible Sets                                                                 |
|--------------|--------------------------------------------------------------------------------|-------|---------------------------------------------------------------------------------|
| ![AAF Image] | ![Possible Worlds Image]                                                      | 0.126 | \{ \}, \{a\}, \{b\}, \{d\}, \{a, b\}, \{a, d\}, \{b, d\}, \{a, b, d\} |
| ![AAF Image] | ![Possible Worlds Image]                                                      | 0.084 | \{ \}, \{a\}, \{b\}, \{c\}, \{a, b\}, \{a, c\}, \{b, c\}, \{a, b, c\} |
| ![AAF Image] | ![Possible Worlds Image]                                                      | 0.294 | \{ \}, \{a\}, \{b\}, \{d\}, \{a, b\}, \{a, d\}, \{b, d\}, \{a, b, d\} |
| ![AAF Image] | ![Possible Worlds Image]                                                      | 0.196 | \{ \}, \{a\}, \{b\}, \{a, b\}, \{b, d\}, \{a, b, d\} |
| ![AAF Image] | ![Possible Worlds Image]                                                      | 0.054 | \{ \}, \{a\}, \{b\}, \{d\}, \{a, b\}, \{a, d\}, \{b, d\}, \{a, b, d\} |
| ![AAF Image] | ![Possible Worlds Image]                                                      | 0.036 | \{ \}, \{a\}, \{b\}, \{a, b\}, \{a, d\}, \{a, b, d\} |
| ![AAF Image] | ![Possible Worlds Image]                                                      | 0.126 | \{ \}, \{a\}, \{b\}, \{d\}, \{a, b\}, \{a, d\}, \{b, d\}, \{a, b, d\} |
| ![AAF Image] | ![Possible Worlds Image]                                                      | 0.084 | \{ \}, \{a\}, \{b\}, \{a, b\}, \{a, d\}, \{b, d\}, \{a, b, d\} |

Table 1. Possible worlds of our example PrAAF from Figure 2. Shaded rows, illustrate possible worlds that generate an invalid AAF.
Furthermore, $P_{Att}(a1 \rightarrow a2)$ is stated to be the conditional probability of the attack existing when both attacking and attacked argument exist in the AAF ($P_{Att}(a1 \rightarrow a2 | a1, a2 \in Args_{Ind})$).

3. Normalizing Constellation Approaches

3.1. PrAAF Normal Forms

A simpler approach to restrict the possible worlds without introducing extra restrictions is to restrict PrAAFs only to probabilistic attacks and not probabilistic arguments.

**Definition 4** (Probabilistic Attack Normal Form). A PrAAF $P$ is in its probabilistic attack normal form if it contains no probabilistic arguments ($\forall a \in Args, P(a) = 1$).

As we illustrate later in the paper any PrAAF can be transformed to a normal form PrAAF that has equivalent possible worlds and define equivalent probability distribution over their admissible extensions. Thus we illustrate that normal form PrAAFs are both the simplest but yet sufficient forms for representing any PrAAF.

4. Transforming General PrAAFs to Normal Form PrAAFs

In order to motivate our normal form for PrAAFs, we illustrate that any general PrAAF can be represented to an equivalent normal form PrAAF. To do so, we present a transformation from any general PrAAF to an equivalent normal form PrAAF.

Our transformation proves, that we can represent as a normal form PrAAF an equivalent constellation of AAFs with the exact same probabilistic distribution over their extensions that a general PrAAF would represent. This indicates that even with our simpler definition for PrAAFs we can model all the same PrAAFs.

4.1. Transforming Probabilistic Arguments to Probabilistic Attacks

We present a transformation of probabilistic argument to probabilistic attack. We start by defining a special argument called ground truth:

**Definition 5** (Ground Truth). We introduce a special argument called Ground Truth and shorthand it with the letter $\eta$. We say that $\eta$ is undeniably true meaning that $\eta$ is never attacked by any argument and is always included in all extensions regardless the semantics used.

The $\eta$ argument restricts the extensions allowed by all semantics in such a way that $\eta$ is always included. For example, in the admissible semantics of an AAF without $\eta$ a valid extension is the empty set $\{\}$, but in an AAF that includes $\eta$ the empty set is not a valid extension under the admissible semantics and the equivalent extension is $\{\eta\}$.

**Definition 6** (Acceptable Extensions). For AAFs that contain $\eta$, an acceptable extension $E$ is one that includes $\eta$ ($\eta \in E$).
**Transformation 1** (Probabilistic argument to probabilistic attack). Any PrAAF $P$, can be transformed to an equivalent PrAAF $P'$ by removing any probabilistic information attached to an argument $a \in \text{Args}$, with $P(a)$ and introducing a probabilistic attack from the ground truth $\eta$ to argument $a$ with probability $1 - P(a)$.

**Definition 7.** We notate $P \equiv_{\eta} P'$ the standard equivalence [20] of PrAAF $P$ with PrAAF $P'$ under semantics $\sigma$ by ignoring the existence of $\eta$ in the acceptable extensions.

**Theorem 1** (Equivalence of transformed PrAAF). A transformed PrAAF $P'$ has an equivalent distribution in terms of admissible sets containing $\eta$ compared with the admissible sets of the original PrAAF $P$.

**Proof.** We split the proof in two parts. First we show that PrAAF $P$ generates AAFs that have the same admissible sets with the generated AAFs from PrAAF $P'$. We point out that for PrAAF $P'$ acceptable admissible sets are only the ones that contain the ground truth argument which we ignore its existence when comparing admissible sets. For example, the empty admissible set of $P$ is equivalent with the $\{ \eta \}$ admissible set of $P'$. A probabilistic argument $pa$ generates two different sets of AAFs, set $S_1$ where $pa$ exists and $S_2$ where $pa$ does not exist.

PrAAF $P'$ generates $S_1'$ the equivalent sets of $S_1$ when $\eta \rightarrow pa$ does not exist and the equivalent $S_2'$ sets of $S_2$ when $\eta \rightarrow pa$ exists. When comparing an AAF with $pa$ versus an AAF without $\eta \rightarrow pa$ the only difference is the existence of $\eta$ as we only consider admissible sets that contain it and we ignore its existence in the admissible sets the two graphs are equivalent thus the $S_1'$ sets are equivalent with the $S_1$ sets.

For $S_2$ where $pa$ does not exist, the equivalent $S_2'$ contains AAFs where the argument $pa$ is been attacked by $\eta$ and is not defended by any other argument. Clearly, as $\eta$ is included in every extension that we consider then every attack originating from $pa$ is defended; thus, the AAFs of $S_2$ generate the same admissible extensions with the AAFs of $S_2'$.

Next part is to show that the probability of each extension is the same. The probability that a set is an admissible extension is been computed by the summation of the possible worlds where that set is admissible. As $S_1, S_2$ are equivalent with $S_1', S_2'$ and produce equivalent AAFs then the possible worlds are equivalent too. The probability of each possible world is also the same as when $pa$ would exist the possible world probability is multiplied by $P(pa)$. In the equivalent case the attack $\eta \rightarrow pa$ does not exist and the possible world probability is multiplied by $1 - (1 - P(pa)) = pa$. Similarly, for the possible worlds that $pa$ does not exist.

**Corollary 1.** PrAAF $P'$ has equivalent acceptable extensions with PrAAF $P$ for all semantics where acceptability of an argument is necessary for the inclusion of the argument in the extension. Such semantics include: complete, preferred, ground and stable semantics. Similarly, as the probabilistic distributions are equivalent then all acceptable extensions of $P'$ will have equal probability with their equivalent extension from $P$.

We also want to note that our transformation is a one-to-one reversible transformation. By one-to-one we mean that the transformation generates a unique equivalent PrAAF from any PrAAF and vice versa, assuming that $\eta$ is known.

---

Such transformation is categorized as a normal expansion [1] of the original PrAAF.
**Proposition 1** (Reversibility of the transformation). *The probabilistic argument to probabilistic attack transformation is reversible and creates a one-on-one equivalent PrAAF.*

**Proof.** Obvious, one can reverse the transformation by just following the reverse steps of the transformation. □

![Figure 3. Example transformed PrAAF](image)

**Figure 3.** Example transformed PrAAF \( \{a, b, c, d, \eta\} \), \( \{1, 1, 1, 1\} \), \( \{a \rightarrow c, b \rightarrow c, c \rightarrow d, \eta \rightarrow c\} \), \( \{0.3, 0.7, 1, 0.6\} \). By using the probabilistic argument to probabilistic attack transformation on the PrAAF of Figure 2 we get the PrAAF of Figure 3. Table 2 presents the possible worlds of the PrAAF of Figure 3. Now, each possible world represents a valid AAF that generates the equivalent acceptable admissible sets like the original PrAAF. Furthermore, the probabilistic distribution is identical.

**Proposition 2** (Complexity of the Transformation). *The probabilistic argument to probabilistic attack transformation has linear complexity \( O(N) \) to the number of probabilistic arguments \( N \) that the original PrAAF contains. It grows the size of the original PrAAF by one argument and by \( N \) attacks. The transformation does not affect the worst case complexity of computing any extension or the probability that a set is any type of an extension.*

5. Conclusion and Future Work

In this paper we have presented a normal form for PrAAFs that is simpler but equally expressive, and a transformation that converts any general PrAAF to an equivalent PrAAF that is at a normal form. The presented normal form for PrAAFs is simpler than the general PrAAFs and allow easier modeling. With our transformation we have illustrated that the normal form can represent any general PrAAF and reproduce equal probability distributions. Our motivation with this paper is to provide a simpler but equally powerful definition for constellation PrAAFs; furthermore, we give a clear insight in the constellation semantics and its restrictions from the point of view of generating possible worlds.

For future work, we are going take advantage of the existing work from the probabilistic logic programming community to provide an efficient system that implements PrAAFs and also examine how the properties of AAF can be used to simplifying probabilistic inference in PrAAFs.
| AAF | Possible World                                                                 | Prob.  | Acceptable Admissible Sets                                      |
|-----|--------------------------------------------------------------------------------|--------|-----------------------------------------------------------------|
| 1   | ¬(a → c) ∧ ¬(b → c) ∧ (η → c) ≡ ¬(a → c) ∧ ¬(b → c) ∧ c                          | 0.126  | {η}, {η, a}, {η, b}, {η, d}, {η, a, b}, {η, a, d}, {η, b, d}, {η, a, b, d} |
| 2   | (a → c) ∧ (η → c) ≡ ¬(a → c) ∧ (η → c) ≡ ¬(a → c) ∧ ¬(b → c) ∧ ¬c              | 0.084  | {η}, {η, a}, {η, b}, {η, c}, {η, a, b}, {η, a, c}, {η, b, c}, {η, a, b, c} |
| 3   | ¬(a → c) ∧ (b → c) ∧ (η → c) ≡ ¬(a → c) ∧ (η → c) ≡ ¬(a → c) ∧ (b → c) ∧ c    | 0.294  | {η}, {η, a}, {η, b}, {η, d}, {η, a, b}, {η, a, d}, {η, b, d}, {η, a, b, d} |
| 4   | (a → c) ∧ (η → c) ≡ ¬(a → c) ∧ (η → c) ≡ ¬(a → c) ∧ ¬(b → c) ∧ ¬c              | 0.196  | {η}, {η, a}, {η, b}, {η, a, b}, {η, b, d}, {η, a, b, d}               |
| 5   | (a → c) ∧ (η → c) ≡ ¬(a → c) ∧ (η → c) ≡ ¬(a → c) ∧ ¬(b → c) ∧ ¬c              | 0.054  | {η}, {η, a}, {η, b}, {η, d}, {η, a, b}, {η, a, d}, {η, b, d}, {η, a, b, d} |
| 6   | (a → c) ∧ (η → c) ≡ ¬(a → c) ∧ (η → c) ≡ ¬(a → c) ∧ ¬(b → c) ∧ ¬c              | 0.036  | {η}, {η, a}, {η, b}, {η, a, b}, {η, a, d}, {η, a, b, d}               |
| 7   | (a → c) ∧ (η → c) ≡ ¬(a → c) ∧ (η → c) ≡ ¬(a → c) ∧ (b → c) ∧ ¬c              | 0.126  | {η}, {η, a}, {η, b}, {η, d}, {η, a, b}, {η, a, d}, {η, b, d}, {η, a, b, d} |
| 8   | (a → c) ∧ (η → c) ≡ ¬(a → c) ∧ (η → c) ≡ ¬(a → c) ∧ ¬(b → c) ∧ ¬c              | 0.084  | {η}, {η, a}, {η, b}, {η, a, b}, {η, a, d}, {η, b, d}, {η, a, b, d}   |

Table 2. Possible worlds after transforming PrAAF of Figure 2.

References

[1] R. Baumann and G. Brewka. Expanding argumentation frameworks: Enforcing and monotonicity results. In *Computational Models of Argument: Proceedings of COMMA*, pages 75–86, 2010.

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

[3] S. Bistarelli, F. Rossi, and F. Santini. A novel weighted defence and its relaxation in abstract argumentation. *Int. J. Approx. Reasoning*, 92:66–86, 2018.

[4] A. Bondarenko, F. Toni, and R. A. Kowalski. An assumption-based framework for non-monotonic reasoning. In *Proceedings of LPNMR*, pages 171–189, Cambridge, MA, USA, 1993. MIT Press.

[5] S. Bragaglia and F. Riguzzi. Approximate inference for logic programs with annotated disjunctions. In *Inductive Logic Programming: Proceedings of ILP*, pages 30–37, 2011.

[6] M. W. A. Caminada, W. A. Carniel, and P. E. Dunne. Semi-stable semantics. *Journal of Logic and Computation*, 22(5):1207–1254, 2012.
