Probabilistic Deduction as a Probabilistic Extension to Assumption-based Argumentation

Extended Abstract

Xiuyi Fan
Nanyang Technological University
Singapore
xyfan@ntu.edu.sg

ABSTRACT

This paper discusses how Probabilistic Deduction (PD), a probabilistic structured argumentation framework with an epistemic approach to probabilistic argumentation, can be viewed as a probabilistic extension to Assumption-based Argumentation.

KEYWORDS

Probabilistic Deduction; Assumption-based Argumentation

ACM Reference Format:
Xiuyi Fan. 2023. Probabilistic Deduction as a Probabilistic Extension to Assumption-based Argumentation: Extended Abstract. In *Proc. of the 22nd International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2023), London, United Kingdom, May 29 – June 2, 2023*, IFAAMAS, 3 pages.

1 INTRODUCTION

Probabilistic Deduction (PD) [4] is a probabilistic structured argumentation framework taking the epistemic approach [7]. PD frameworks are constructed with probabilistic rules (p-rules) [5], which allow one to construct arguments and attacks between arguments as standard inference rules. At the same time, one can also calculate probabilities of arguments based on probabilities of p-rules, which are read as conditional probabilities. The construct central to this work is a mapping, ABA2PD, from Assumption-based Argumentation (ABA) [2] frameworks to PD frameworks. With ABA2PD, arguments and attacks in an ABA framework are mapped to their counterparts, respectively, in a PD framework. Semantically, under the *probabilistic closed-world assumption (P-CWA)* and with *minimum entropy reasoning*, ABA2PD achieves that (1) every ABA argument in F has a counterpart in F_{PD}; (2) if an argument in F_{PD} has probability 1, then its counterpart in the ABA framework is in a complete extension; (3) if an argument in F is in a stable extension, then its counterpart in the PD framework has probability 1; and (4) F is coherent iff F_{PD} is P-CWA consistent.

2 PRELIMINARY

Given *atoms* $\sigma_0, \ldots, \sigma_n$ forming a language $\mathcal{L} = \{\sigma_0, \ldots, \sigma_n\}$, we let $\mathcal{L}^c$ be the closure of $\mathcal{L}$ under the classical negation $\neg$.

Definition 2.1. [5] Given a language $\mathcal{L}$, a *probabilistic rule (p-rule)* is $\sigma_0 \leftarrow \sigma_1, \ldots, \sigma_k : [\theta]$ for $k \geq 0$, $\sigma_i \in \mathcal{L}^c$, and $0 \leq \theta \leq 1$. θ is referred to as the *probability* of the p-rule.

Definition 2.2. [6] Given a language $\mathcal{L}$ with n atoms, the *Complete Conjunction Set (CC Set)* Ω of $\mathcal{L}$ is the set of 2^n conjunction of literals such that each conjunction contains n distinct atoms.

Definition 2.3. [5] Given a language $\mathcal{L}$ and a set of p-rules $\mathcal{R}$, let Ω be the CC set of $\mathcal{L}$. A function $\pi : \Omega \longrightarrow [0, 1]$ is a *consistent probability distribution* with respect to $\mathcal{R}$ on $\mathcal{L}$ for Ω iff:

(1) For all $\omega_i \in \Omega$,
$$0 \leq \pi(\omega_i) \leq 1; \tag{1}$$
(2) It holds that:
$$\sum_{\omega_i \in \Omega} \pi(\omega_i) = 1. \tag{2}$$
(3) For each p-rule $\sigma_0 \leftarrow: [\theta] \in \mathcal{R}$, it holds that:
$$\theta = \sum_{\omega_i \in \Omega, \omega_i \models \sigma_0} \pi(\omega_i). \tag{3}$$
(4) For each p-rule $\sigma_0 \leftarrow \sigma_1, \ldots, \sigma_k : [\theta] \in \mathcal{R}$, $(k > 0)$, it holds that:[1]
$$\theta = \frac{\sum_{\omega_i \in \Omega, \omega_i \models \sigma_0 \wedge \ldots \wedge \sigma_k} \pi(\omega_i)}{\sum_{\omega_i \in \Omega, \omega_i \models \sigma_1 \wedge \ldots \wedge \sigma_k} \pi(\omega_i)}. \tag{4}$$

$\mathcal{R}$ is *consistent* iff there is a consistent distribution wrt $\mathcal{R}$.

Given a language $\mathcal{L}$ and a set of p-rules $\mathcal{R}$, a *deduction* for $\sigma \in \mathcal{L}^c$ with $\Sigma \subseteq \mathcal{L}^c$, denoted $\Sigma \vdash_{\mathsf{D}} \sigma$, is a finite tree $\mathcal{T}$ as defined in ABA deductions for σ, with $\Sigma = \{\sigma \in \mathcal{L}^c | \sigma \text{ labels a node in } \mathcal{T}\}$. Let $\{\Sigma_1 \vdash_{\mathsf{D}} \sigma, \ldots, \Sigma_m \vdash_{\mathsf{D}} \sigma\}$ be all maximal deductions[2] for σ where $\Sigma_1 = \{\sigma_1^1, \ldots, \sigma_{k1}^1\}, \ldots, \Sigma_m = \{\sigma_1^m, \ldots, \sigma_{km}^m\}$. Let
$$S = \bigwedge_{i=1}^{k1} \sigma_i^1 \vee \ldots \vee \bigwedge_{i=1}^{km} \sigma_i^m. \tag{5}$$

The *Probabilistic Closed World Assumption (P-CWA)* asserts that for each $\sigma \in \mathcal{L}^c$,
$$\sum_{\omega \in \Omega, \omega \models \sigma, \omega \not\models S} \pi(\omega) = 0. \tag{6}$$

A set of p-rules $\mathcal{R}$ is consistent under P-CWA iff $\mathcal{R}$ is consistent when P-CWA is imposed. We also refer to such p-rules as P-CWA consistent. For a set of P-CWA consistent p-rules, the probability of $\sigma \in \mathcal{L}^c$ wrt a consistent probability distribution π is
$$\Pr(\sigma) = \sum_{\omega_i \in \Omega, \omega_i \models \sigma} \pi(\omega_i) = \sum_{\omega_i \in \Omega, \omega_i \models S} \pi(\omega_i). \tag{7}$$

Definition 2.4. [4] A *Probabilistic Deduction (PD)* framework is a pair $\langle \mathcal{L}, \mathcal{R} \rangle$ where $\mathcal{L}$ is the language, $\mathcal{R}$ a set of p-rules such that

[1] If $\sum_{\omega_i \in \Omega, \omega_i \models \sigma_1 \wedge \ldots \wedge \sigma_k} \pi(\omega_i) = 0$, then $\sum_{\omega_i \in \Omega, \omega_i \models \sigma_0 \wedge \ldots \wedge \sigma_k} \pi(\omega_i) = 0$ as $\sigma_0 \wedge \ldots \wedge \sigma_k \models \sigma_1 \wedge \ldots \wedge \sigma_k$. We thus require $\theta = 0$ in such cases for π to be consistent.
[2] A deduction $S \vdash_{\mathsf{D}} \sigma$ is maximal when there is no $S' \vdash_{\mathsf{D}} \sigma$ such that $S \subset S'$.

Proc. of the 22nd International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2023), A. Ricci, W. Yeoh, N. Agmon, B. An (eds.), May 29 – June 2, 2023, London, United Kingdom.

- for all $\rho \in \mathcal{R}$, literals in ρ are in $\mathcal{L}^c$,
- $\mathcal{R}$ is P-CWA consistent.

Definition 2.5. [4] Given a PD framework $\langle \mathcal{L}, \mathcal{R} \rangle$, an argument for $\sigma \in \mathcal{L}$ supported by $S \subseteq \mathcal{L}$, denoted $S \vdash \sigma$ is such that there is a deduction $\mathsf{A} = S \vdash_{\mathsf{D}} \sigma$ in which for each leaf node N in A, either

(1) N is labelled by τ, or
(2) N is labelled by some $\sigma' \in \mathcal{L}, \neg\sigma' \leftarrow \_ : [\cdot] \in \mathcal{R}, |S| > 1$.

Definition 2.6. [4] For two arguments $\mathsf{A} = \_ \vdash \sigma$ and $\mathsf{B} = \Sigma \vdash \_$ in some PD framework, A attacks B if $\neg\sigma \in \Sigma$.

The probability of an argument in PD framework is defined as the probability of the conjunction of all literals in the argument. Formally, given a PD framework $\mathsf{F_{PD}}$ and a P-CWA consistent probability distribution π, for each argument $\{\sigma_0, \dots, \sigma_k\} \vdash \sigma_0$ in $\mathsf{F_{PD}}$,

$$\Pr(\{\sigma_0, \dots, \sigma_k\} \vdash \sigma_0) = \sum_{\omega_i \in \Omega, \omega_i \models \sigma_0 \wedge \dots \wedge \sigma_k} \pi(\omega_i). \quad (8)$$

Minimum Entropy reasoning asserts that when multiple distributions are consistent for assigning probabilities to complete conjunctions, a distribution that minimizes the entropy is selected.

Definition 2.7. Given a PD framework $\langle \mathcal{L}, \mathcal{R} \rangle$, let Π be the set of consistent probability distributions wrt $\mathcal{R}$ on $\mathcal{L}$, a *Minimum Entropy Distribution* $\pi_{01} \in \Pi$ is such that

$$\pi_{01} = \arg\min_{\pi \in \Pi} \left(-\sum_{\omega \in \Omega} \pi(\omega) \log(\pi(\omega)) \right).$$

3 ABA-PD MAPPING

Definition 3.1. The function ABA2PD is a mapping from ABA frameworks to PD frameworks such that for an ABA framework $\mathsf{F} = \langle \mathcal{L}, \mathcal{R}, \mathcal{A}, C \rangle$, $\mathsf{ABA2PD(F)} = \langle \mathcal{L}_{\mathsf{pd}}, \mathcal{R}_{\mathsf{pd}} \rangle$, where:

- $\mathcal{L}_{\mathsf{pd}} = \mathcal{L} \cup \{\alpha_{asm} | \alpha \in \mathcal{A}, \alpha_{asm} \notin \mathcal{L}\}$,
- $\mathcal{R}_{\mathsf{pd}} = \mathcal{R}_r \cup \mathcal{R}_a \cup \mathcal{R}_c \cup \mathcal{R}_e$, in which:
  - $\mathcal{R}_r = \{\rho : [1] | \rho \in \mathcal{R}\}$,
  - $\mathcal{R}_a = \{\alpha \leftarrow \alpha_{asm} : [1] | \alpha \in \mathcal{A}\}$,
  - $\mathcal{R}_c = \{\neg\alpha_{asm} \leftarrow \sigma : [1] | C(\alpha) = \sigma\}$, and
  - $\mathcal{R}_e = \{\neg\sigma \leftarrow : [1] | \sigma \in \mathcal{L} \setminus \mathcal{A} \text{ and } \nexists \sigma \leftarrow \_ \in \mathcal{R}\}$.

Syntactically, for each (ABA) argument A in F, there is a counterpart of A in ABA2PD(F). Moreover, if $\mathsf{A} = \_ \vdash_{\mathsf{A}} \sigma$ attacks another argument $\mathsf{B} = \{\alpha, \dots\} \vdash_{\mathsf{A}} \_$ in the ABA framework, where $C(\alpha) = \sigma$, then there exists $\mathsf{A}^+_{\mathsf{PD}} = \{\neg\alpha, \sigma, \dots\} \vdash \neg\alpha_{asm}$ in the PD framework, such that $\mathsf{A}^+_{\mathsf{PD}}$ contains $\neg\alpha_{asm}$ and all literals in $\mathsf{A_{PD}}$. Attacks between ABA arguments are preserved in their PD counterparts, i.e., given ABA framework $\mathsf{F} = \langle \mathcal{L}, \mathcal{R}, \mathcal{A}, C \rangle$, for each ABA argument $\mathsf{A} = \_ \vdash_{\mathsf{A}} \sigma$ in F, if there exists $\alpha \in \mathcal{A}$, such that $C(\alpha) = \sigma$, then there is a PD argument of the form $\mathsf{A}^+_{\mathsf{PD}} = \{\neg\alpha_{asm}, \sigma, \dots\} \vdash \neg\alpha_{asm}$ in ABA2PD(F) such that all literals in $\mathsf{A_{PD}}$ are in $\mathsf{A}^+_{\mathsf{PD}}$.

To establish semantics connections between ABA and PD, we see that PD's probability semantics can be viewed as a complete labelling in ABA-PD frameworks. To this end, we observe that: (1) unattacked arguments have probability 1; (2) an attacked argument has probability 1 iff all its attackers have probability 0; and (3) with the minimum entropy reasoning, an argument has probability 0 iff it has an attacker with probability 1. With these, we have our theorems that connect argument probability with labelling.

Theorem 3.1. Given an ABA-PD framework $\mathsf{F_{PD}}$, let As be the set of arguments in F, with minimum entropy reasoning, the *Probabilistic Labelling* function $\Xi_c : \mathsf{As} \mapsto \{\mathtt{in}, \mathtt{out}, \mathtt{undec}\}$:

$$\Xi_c(\mathsf{A}) = \begin{cases} \mathtt{in} & \text{if } \Pr(\mathsf{A}) = 1, \\ \mathtt{out} & \text{if } \Pr(\mathsf{A}) = 0, \\ \mathtt{undec} & \text{otherwise.} \end{cases}$$

in which $\mathsf{A} \in \mathsf{As}$, is a complete labelling.

Key results that lead to the proof of Theorem 3.1 are (1) Proposition 2 in [1], which states that "a labelling is *complete* iff for all arguments it holds that: an argument is in iff all of its attackers are out, and an argument is out iff it has at least one attacker that is in" and the observation that if a set of p-rules is consistent, then there is a consistent distribution that assigns 1 to a single complete conjunction while assigning 0 to all other complete conjunctions. Such a distribution is a minimum entropy distribution.

Corollary 3.1. Given an ABA framework F, consider the PABA-PD framework ABA2PD(F). For each argument A in F, with a minimum entropy distribution π_{01}, if $\Pr(\mathsf{A_{pd}}) = 1$, then A is in a complete extension.

The other direction of Corollary 3.1, if A in F is complete then $\Pr(\mathsf{A_{pd}}) = 1$, does not hold in general as F with arguments in a three-cycle and an unattacked argument has a complete labelling such that the three arguments in the three-cycle are labelled undec and the unattacked argument labelled in. However, ABA2PD(F) is not P-CWA consistent. Thus, we consider the stable labelling, which is a complete labelling without undec in the next theorem.

Theorem 3.2. Given an ABA framework F with arguments As, let Ξ_s be a stable labelling function on As and $\mathsf{F_{PD}} = \mathsf{ABA2PD(F)}$, there exists a consistent distribution π, such that for each $\mathsf{A} \in \mathsf{As}$,

$$\Pr(\mathsf{A_{pd}}) = \begin{cases} 1 & \text{if } \Xi_s(\mathsf{A}) = \mathtt{in}, \\ 0 & \text{if } \Xi_s(\mathsf{A}) = \mathtt{out}. \end{cases}$$

An interesting observation from Corollary 3.1 and Theorem 3.2 is that these results can be expressed with the notion of *coherence* introduced in Dung's original paper [3]. There, the concept of *controversial arguments* has been defined as arguments that both (indirectly) attack and defend the same argument (e.g., arguments in an odd-length-cycle can be controversial). Dung has shown that argumentation frameworks without any controversial argument are coherent, in the sense that each maximal (wrt set inclusion) admission extension is also a stable extension in such argumentation frameworks (Theorem 33, [3]). With the notion of coherent argumentation framework, we can express our findings as follows[3].

Theorem 3.3. For an ABA framework F, let $\mathsf{F_{PD}} = \mathsf{ABA2PD(F)}$, F is coherent iff $\mathsf{F_{PD}}$ is P-CWA consistent.

An interesting remark is that in Dung's seminal paper, the stable semantics is at the core of the discussion, as which is the notion "underlining exactly the way the notions of stable models in logical programming, extensions in Reiter's default logic, and stable expansion in Moore's autoepistemic logic" [3]. It is interesting to see that PD's probability semantics is related to that as well.

[3]Although Dung's work in [3] is about abstract argumentation, it is clear that the same definition can be applied to ABA as well.

ACKNOWLEDGMENTS

This research is supported by the Ministry of Education, Singapore. (Grant ID: RG17/22)

REFERENCES

[1] P. Baroni, M. Caminada, and M. Giacomin. 2011. An introduction to argumentation semantics. *Knowl. Eng. Rev.* 26, 4 (2011), 365–410. https://doi.org/10.1017/S0269888911000166

[2] K. Čyras, X. Fan, C. Schulz, and F. Toni. 2017. Assumption-Based Argumentation: Disputes, Explanations, Preferences. *IfCoLog JLTA* 4, 8 (2017). https://doi.org/10.1126/scirobotics.aan6080

[3] P. M. Dung. 1995. On the acceptability of arguments and its fundamental role in nonmonotonic reasoning, logic programming and n-person games. *Artificial Intelligence* 77, 2 (1995), 321–357. https://doi.org/10.1016/0004-3702(94)00041-X

[4] Xiuyi Fan. 2022. Probabilistic Deduction: an Approach to Probabilistic Structured Argumentation. https://doi.org/10.48550/ARXIV.2209.00210

[5] Xiuyi Fan. 2022. Rule-PSAT: Relaxing Rule Constraints in Probabilistic Assumption-based Argumentation. In *Proc. of COMMA*.

[6] T. C. Henderson, R. Simmons, B. Serbinowski, M. Cline, D. Sacharny, X. Fan, and A. Mitiche. 2020. Probabilistic sentence satisfiability: An approach to PSAT. *Artificial Intelligence* 278 (2020). https://doi.org/10.1016/j.artint.2019.103199

[7] A. Hunter. 2013. A Probabilistic Approach to Modelling Uncertain Logical Arguments. *International Journal Approximate Reasoning* (2013).