

Admissibility in Probabilistic Argumentation

Nikolai Käfer
Christel Baier
Martin Diller
Clemens Dubsloff
Sarah Alice Gaggl
Faculty of Computer Science
Technische Universität Dresden
Dresden, Germany

NIKOLAI.KAEFER@TU-DRESDEN.DE
CHRISTEL.BAIER@TU-DRESDEN.DE
MARTIN.DILLER@TU-DRESDEN.DE
CLEMENS.DUBSLAFF@TU-DRESDEN.DE
SARAH.GAGGL@TU-DRESDEN.DE

Holger Hermanns
Saarland Informatics Campus
Saarland University
Saarbrücken, Germany

Institute of Intelligent Software
Guangzhou, China

HERMANN@CS.UNI-SAARLAND.DE

Abstract

Abstract argumentation is a prominent reasoning framework. It comes with a variety of semantics and has lately been enhanced by probabilities to enable a quantitative treatment of argumentation. While admissibility is a fundamental notion for classical reasoning in abstract argumentation frameworks, it has barely been reflected so far in the probabilistic setting. In this paper, we address the quantitative treatment of abstract argumentation based on probabilistic notions of admissibility. Our approach follows the natural idea of defining probabilistic semantics for abstract argumentation by systematically imposing constraints on the joint probability distribution on the sets of arguments, rather than on probabilities of single arguments. As a result, there might be either a uniquely defined distribution satisfying the constraints, but also none, many, or even an infinite number of satisfying distributions are possible. We provide probabilistic semantics corresponding to the classical complete and stable semantics and show how labeling schemes provide a bridge from distributions back to argument labelings. In relation to existing work on probabilistic argumentation, we present a taxonomy of semantic notions. Enabled by the constraint-based approach, standard reasoning problems for probabilistic semantics can be tackled by SMT solvers, as we demonstrate by a proof-of-concept implementation.

1. Introduction

In its basic form, an *abstract argumentation framework* (AF) (Dung, 1995) consists of a set of abstract arguments together with a binary relation that represent conflicts between arguments, the so-called *attack relation*. AFs are popular to describe contentious information and draw conclusions from it using formalized arguments. The popularity of the AF concept has led to a variety of extensions like notions to handle preferences and values on arguments (Amgoud & Cayrol, 2002; Bench-Capon, 2003), weights (Dunne et al., 2011), probabilities (Li, Oren, & Norman, 2011; Thimm, 2012; Hunter, 2013) or introducing a positive influence relation between arguments, so-called *supports* (Amgoud et al., 2008;

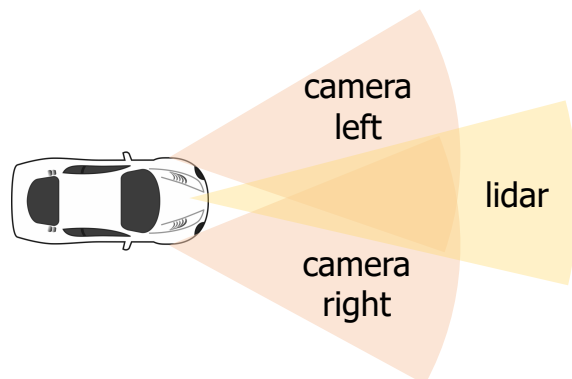


Figure 1: Exemplary sensor layout of a semi-autonomous vehicle

Nouioua & Risch, 2011). Furthermore, *abstract dialectical frameworks* (ADFs) as a powerful generalization of Dung’s framework have been introduced (Brewka & Woltran, 2010; Brewka et al., 2013; Straß & Wallner, 2015; Gaggl et al., 2021), which also allow to handle probabilities (Polberg & Doder, 2014).

In this paper, we focus on the emerging field of AFs in the probabilistic setting. As a concrete example to ground our discussion, we consider an argumentation-based decision framework for a semi-autonomous vehicle as depicted in Figure 1. Here, a central decision entity (the “supervisor”, see Faqeh et al., 2020) has access to possibly conflicting information from several sensors (left/right camera, lidar sensor) with overlapping sensing areas in front of the vehicle. The sensor values are assumed to be of a Boolean nature and indicate whether an obstacle is detected. They together induce reasons to assume that there is – or is not – an obstacle in some specific area in front of the vehicle. The supervisor aggregates this information in order to decide whether to continue moving forward.

We translate this scenario into an argumentation setting as follows. We use literals cl and \bar{cl} to denote arguments expressing that the left camera has detected an obstacle or not, and similar for the right camera (cr and \bar{cr}) and for the lidar sensor (ld and \bar{ld}). Slightly more complex arguments represent reasons for and against an obstacle being either on the left (l and \bar{l}), the right (r and \bar{r}), or in the middle (m and \bar{m}) of the area ahead of the vehicle. For example, the argument $\bar{cl} \Rightarrow \bar{l}$ expresses that a “silent” left camera sensor backs the conclusion that there is no obstacle on the left. The literals ct and st indicate that the vehicle should continue ahead or stop respectively.

The resulting argument graph is depicted in Figure 2, with nodes representing arguments and directed edges representing attacks. In particular, we see first of all that contrary sensor readings attack each other (e.g., \bar{cr} and cr). Also, arguments about sensors can “undermine” arguments about the location of obstacles (e.g., \bar{cl} undermines $cl \Rightarrow l$) while arguments about the location can “rebut” each other when they make contradictory claims (e.g., $cl \Rightarrow m$ and $\bar{ld} \Rightarrow \bar{m}$ rebut each other). Moreover, arguments for an obstacle being in the middle can counter the decision to continue (e.g., $ld \Rightarrow m$ attacks ct), while the decision to continue invalidates the decision to stop (ct attacks st). This reflects that if there is no reason to

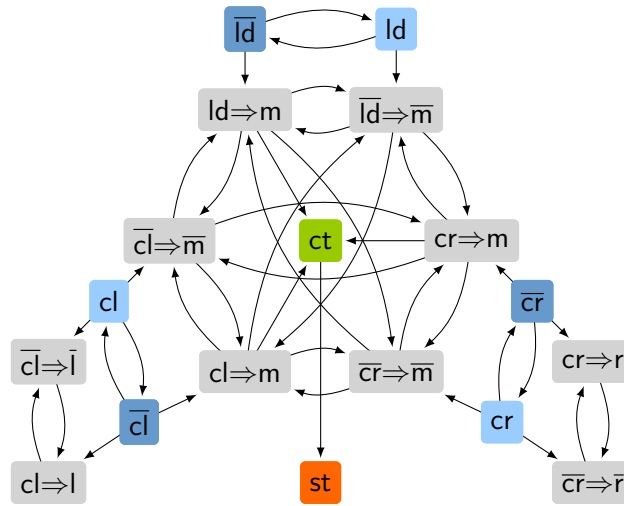


Figure 2: Argument graph for the vehicle example

assume that there is an obstacle in the middle, the vehicle will continue driving. Otherwise, it will need to stop¹.

We are now interested in the situation where the degree of acceptance of arguments lies on a continuum. For instance, object detection by the right camera might fail with some probability (false negatives), spurious detections may be possible (false positives), and the spatial layout of the overlapping camera views may give quantitative information where to expect an object, albeit the information provided may not be clear-cut. This implies that there will be different degrees of uncertainty whether to continue moving forward or not.

This motivation helps to understand why there is a growing body of inspiring research on *probabilistic abstract argumentation frameworks* (PrAFs). In this paper, we contribute to this research spectrum taking a probability-theoretic perspective based on the *epistemic approach* (see Section 4 for an overview). We start off from the idea that a PrAF induces probability distributions over various arguments being accepted or not. Just like classical AFs can induce potentially multiple valid interpretations (called extensions² in the literature), a PrAF can induce multiple such distributions. We embark on extending the basic notions of classical argumentation theory to the probabilistic setting in a conservative manner. In this, we take the fairly natural view that no argument and its attackers may be accepted at the same time. This does not mean that both views on an argument and its attacker cannot have non-zero probability, but simply that the probability of an argument

1. To model the arguments and obtain the argument graph we base ourselves on the ASPIC+ framework (Modgil & Prakken, 2018). Specifically, we use ASPIC+ with possibly non-symmetric negation. Then the ASPIC+ framework underlying our example consists of the ordinary premises $\{cl, cr, ld, \bar{cl}, \bar{cr}, \bar{ld}, ct, st\}$ (i.e., these are uncertain and can be attacked) and defeasible rules $\{cl \Rightarrow l, cl \Rightarrow m, \bar{cl} \Rightarrow \bar{l}, \bar{cl} \Rightarrow \bar{m}, cr \Rightarrow r, cr \Rightarrow m, \bar{cr} \Rightarrow \bar{r}, \bar{cr} \Rightarrow \bar{m}, ld \Rightarrow m, \bar{ld} \Rightarrow \bar{m}\}$. Note that in our example we use the same notation for arguments as for the rules underlying them. There are no axiom premises nor are there strict rules in our example. Dual literals (e.g., cl and \bar{cl} as well as m and \bar{m}) are contradictories of each other while m is the contrary of ct and ct is the contrary of st .
2. In this paper, we avoid using “extension” in this sense, and instead use this term to identify probabilistic semantics that are *conservative extensions* of classical AF semantics, see Definition 5.

and its attacker being simultaneously accepted is zero by construction. Similarly, we would like to ensure that accepted arguments are defended, meaning that each of their attackers are attacked with a probability of one. This is indeed needed for making sense of scenarios like the vehicle example above, but is in contrast to earlier work on PrAFs (Hunter & Thimm, 2017) where probabilities of single arguments were put in focus. When lifting the classical concepts from AFs to PrAFs, especially the notion of admissibility gives rise to a hierarchy of different interpretations and leads to an entire taxonomy of semantics. Along this discussion, it becomes apparent that each lifted semantics concept imposes a set of constraints on the joint probability distributions of arguments to be accepted or not. We will show in this paper that, with one exception, all these sets of constraints we consider are linear, i.e., they are expressible in the linear arithmetic theory of the reals. Among others, this gives us decidability and bridges to the world of SMT solvers, which are nowadays well capable of handling large sets of linear and even non-linear constraints. Indeed, we present a prototypical tool for studying a variety of questions arising in a PrAF. Specifically, we apply the tool to the above vehicle example to pinpoint some fine details that help to understand our contribution as well as getting an impression of practical relevance. For example, the tool can compute a distribution maximizing the value of ct assuming that cl and ld are accepted almost surely with additional constraints on the rate of false positives for the sensors, while satisfying a particular semantics (or even sets thereof). The tool and all experimental data are publicly available at

<https://www.perspicuous-computing.science/cpraa>

In summary, our contribution is fourfold:

- (i) We provide a profound study of admissibility as well as complete and stable semantics in a probability-theoretic approach to abstract argumentation,
- (ii) discuss a hierarchy of resulting semantics in the context of earlier work,
- (iii) investigate the complexity of established computational problems lifted to the probabilistic setting, and
- (iv) present prototypical tool support for experimenting with these semantics and further context-specific constraints.

We structured the paper as follows. In Section 2 we introduce the necessary background on abstract argumentation, as well as notations needed for the probabilistic setting such as assignments and distributions. Our probabilistic argumentation framework is presented in Section 3, where we lift classical argumentation semantics to the probabilistic setting and discuss our probabilistic notions of admissibility, complete semantics, saturation, and labeling schemes. We relate our semantics to the existing (probabilistic) semantics and further related work in Section 4 before we discuss the complexities of various probabilistic argumentation decision problems in Section 5. Details about our implementation and experimental studies on the vehicle example are provided in Section 6. We close the paper with concluding remarks and further work (Section 7).

This paper is the journal version of the conference publication (Baier, Diller, Dubslaff, Gaggl, Hermanns, & Käfer, 2021), containing full proofs and enhancing our framework by saturation semantics, labeling schemes, and complexity-theoretic considerations.

2. Preliminaries

In this section, we introduce the basics of abstract argumentation along with classical argumentation semantics on which we base our probabilistic abstract argumentation framework. We start with a definition of abstract argumentation frameworks following (Dung, 1995). For a detailed discussion on argumentation semantics we refer to (Baroni, Caminada, & Giacomin, 2011).

Definition 1. An abstract argumentation framework (AF) is a pair $\mathcal{F} = \langle Arg, Att \rangle$ where Arg is a finite set of arguments and $Att \subseteq Arg \times Arg$ an attack relation.

A pair $(A, B) \in Att$ means that argument A attacks argument B . We denote the set of attackers of A by $\neg A := \{B \in Arg : (B, A) \in Att\}$ and the set of A 's attackees by $A^\rightarrow := \{B \in Arg : (A, B) \in Att\}$. A is called *initial* if it has no attackers, i.e., if $\neg A$ is empty. These notions naturally extend to sets of arguments $\mathcal{S} \subseteq Arg$ by $\neg \mathcal{S} := \bigcup_{A \in \mathcal{S}} \neg A$ and $\mathcal{S}^\rightarrow := \bigcup_{A \in \mathcal{S}} A^\rightarrow$. Two arguments A and B are in *conflict* if $(A, B) \in Att$ or $(B, A) \in Att$, i.e., either one is attacking the other. \mathcal{S} *defeats* an argument B if $B \in \mathcal{S}^\rightarrow$, i.e., at least one argument in \mathcal{S} attacks B . An argument C is *defended by* \mathcal{S} if \mathcal{S} defeats all attackers $B \in \neg C$. The set of all arguments defended by \mathcal{S} is thus given by

$$Defend(\mathcal{S}) := \{C \in Arg : B \in \mathcal{S}^\rightarrow \text{ for each } B \in \neg C\}.$$

2.1 Classical Argumentation Semantics

Semantics for AFs are given by collections of argument sets that do not exhibit conflicts.

Definition 2. For an AF $\mathcal{F} = \langle Arg, Att \rangle$, a set $\mathcal{S} \subseteq Arg$ is said to be *conflict-free* (CF) if $(A, B) \notin Att$ for all arguments $A, B \in \mathcal{S}$. A *conflict-free argument set* \mathcal{S} is

(ST) *stable* if $\mathcal{S} \cup \mathcal{S}^\rightarrow = Arg$,

(ADM) *admissible* if $\mathcal{S} \subseteq Defend(\mathcal{S})$,

(CMP) *complete* if $\mathcal{S} = Defend(\mathcal{S})$,

(GR) *grounded* if there is no complete $\mathcal{T} \subseteq Arg$ with $\mathcal{T} \subset \mathcal{S}$, and

(PRF) *preferred* if there is no complete $\mathcal{T} \subseteq Arg$ with $\mathcal{T} \supset \mathcal{S}$.

The classical argumentation semantics σ for \mathcal{F} with $\sigma \in \{CF, ST, ADM, CMP, GR, PRF\}$ is the set $[\mathcal{F}]_\sigma$ of all argument sets $\mathcal{S} \subseteq Arg$ where condition σ as above holds.

Classical argumentation semantics thus require the absence of conflicts in their contained argument sets, a property that takes over to the set of defended arguments.

Lemma 3 (Dung, 1995). Let $\mathcal{F} = \langle Arg, Att \rangle$ be an AF. If a set $\mathcal{S} \subseteq Arg$ is *conflict-free*, then so is $Defend(\mathcal{S})$.

2.2 Assignments

In the probabilistic setting, each argument A of a given AF $\mathcal{F} = \langle Arg, Att \rangle$ is treated as a Boolean random variable with the same name. An *assignment* is a function $\beta: Arg \rightarrow \{\top, \text{F}\}$ that determines for each variable, whether the argument is accepted (\top) or not (F). The set of all assignments for Arg is denoted by $\text{Asg}(Arg)$. An *event* ϕ is a set of assignments, i.e., $\phi \subseteq \text{Asg}(Arg)$. There exists a straight-forward one-to-one correspondence between argument sets and assignments: Given an argument set $\mathcal{S} \subseteq Arg$, the corresponding assignment is given by its *characteristic function* $\text{id}_{\mathcal{S}}$ defined by $\text{id}_{\mathcal{S}}(A) := \top$ if $A \in \mathcal{S}$ and $\text{id}_{\mathcal{S}}(A) := \text{F}$ otherwise. Conversely, each assignment $\beta \in \text{Asg}(Arg)$ naturally induces an argument set $Arg_{\beta} := \{A \in Arg : \beta(A) = \top\}$. The switch between argument sets and assignments is thus merely of syntactic nature, and we may use both representations interchangeably. In the same vein an assignment is admissible (or complete, stable, ...) if the corresponding argument set is admissible (or complete, stable, ...). Further, we use propositional logic formulas over arguments to specify sets of assignments, e.g., we write $A \wedge \neg B$ with $A, B \in Arg$ for the set of assignments $\phi \subseteq \text{Asg}(Arg)$ where $\phi = \{\beta \in \text{Asg}(Arg) : \beta(A)=\top, \beta(B)=\text{F}\}$. Using such kind of notation, we define the set of assignments whose corresponding argument sets defend an argument $C \in Arg$ by

$$\Delta(C) := \bigwedge_{B \in \neg C} \bigvee_{A \in \neg B} A$$

Note that as usual, empty conjunctions stand for \top and empty disjunctions for F . To this end, e.g., in case $\neg B$ is empty, $\Delta(C)$ necessarily is the empty set.

2.3 Distributions

A probability distribution over a set X is a function $\mu: X \rightarrow [0, 1]$ where $\sum_{\beta \in X} \mu(\beta) = 1$. The set of all distributions over X is denoted by $\text{Distr}(X)$. The *support* of a distribution μ is defined by $\text{Supp}(\mu) := \{\beta : \mu(\beta) > 0\}$. μ is a *Dirac distribution* if $\mu(\beta) \in \{0, 1\}$ for all $\beta \in \text{Asg}(Arg)$, or, equivalently, if $\text{Supp}(\mu)$ is a singleton. For a fixed $\beta \in X$, we write Dirac_{β} for the uniquely defined Dirac distribution where $\text{Dirac}_{\beta}(\beta) = 1$.

In the following, we are only concerned with distributions over the set of assignments $\text{Asg}(Arg)$. For brevity, we write $\text{Distr}(\mathcal{F})$ for $\text{Distr}(\text{Asg}(Arg))$, the set of all possible distributions for an AF \mathcal{F} . Each distribution $\mu \in \text{Distr}(\mathcal{F})$ induces a probability measure over events, denoted by μ as well, i.e., a function $\mu: 2^{\text{Asg}(Arg)} \rightarrow [0, 1]$ with $\mu(\phi) := \sum_{\beta \in \phi} \mu(\beta)$ for events $\phi \subseteq \text{Asg}(Arg)$. Note that $\Delta(C)$ is an event for any argument $C \in Arg$. As usual, we say that μ has outcome ϕ *almost surely* in case $\mu(\phi) = 1$. For argument sets $\mathcal{S} \subseteq Arg$, we use $\text{Dirac}_{\mathcal{S}}$ as a short form for $\text{Dirac}_{\text{id}_{\mathcal{S}}}$. Finally, for two events $\phi, \psi \subseteq \text{Asg}(Arg)$ with $\mu(\psi) > 0$, the *conditional probability* of ϕ given ψ is defined as

$$\mu(\phi \mid \psi) := \frac{\mu(\phi \wedge \psi)}{\mu(\psi)}.$$

3. Probabilistic Argumentation Semantics

A *probabilistic argumentation semantics* ρ assigns to each AF $\mathcal{F} = \langle Arg, Att \rangle$ a subset $\llbracket \mathcal{F} \rrbracket_\rho$ of $Distr(\mathcal{F})$. While for any classical argumentation semantics σ the argument sets in $[\mathcal{F}]_\sigma$ only specify which arguments can be jointly accepted, distributions in $\llbracket \mathcal{F} \rrbracket_\rho$ give rise to a probabilistic interpretation of the arguments, e.g., to specify the belief in, the acceptance of, or the frequency of arguments to be accepted.

Definition 4 (Likelihood of arguments). *For an AF $\mathcal{F} = \langle Arg, Att \rangle$, a distribution $\mu \in Distr(\mathcal{F})$, and an argument $A \in Arg$, we refer to $\mu(A)$ as the μ -likelihood of A .*

We denote by $Arg_\mu := \{A \in Arg : \mu(A) = 1\}$ the set of almost-surely accepted arguments, i.e., arguments with a μ -likelihood of one. Note that here, A stands for a (basic) propositional logic formula over arguments that specifies the set of all events β where $\beta(A) = \top$.

Classical semantics for AFs can be lifted naturally to the probabilistic setting by Dirac distributions corresponding to the semantics' argument sets. Likewise, probabilistic semantics can be restricted to the classical case by considering the argument sets corresponding to the Dirac distributions induced by the semantics. Under this view, a probabilistic semantics where all induced Dirac distributions agree with the Dirac distributions of the argument sets of a classical semantics is called a *conservative extension*.

Definition 5 (Conservative extension). *A probabilistic argumentation semantics ρ is a conservative extension of a classical argumentation semantics σ if for all AFs $\mathcal{F} = \langle Arg, Att \rangle$ and argument sets $\mathcal{S} \subseteq Arg$, we have*

$$\mathcal{S} \in [\mathcal{F}]_\sigma \text{ iff } Dirac_{\mathcal{S}} \in \llbracket \mathcal{F} \rrbracket_\rho.$$

When looking for a probabilistic semantics that reflects one of the classical notions, constituting a conservative extension is by itself clearly not sufficient. However, we deem it a necessary criterion, and will show that all probabilistic semantics introduced later on are conservative extensions of their corresponding classical semantics.

As a basic instance of a conservative extension, we present the element-wise lifting of classical argumentation semantics, first introduced by (Thimm, Baroni, Giacomin, & Vicig, 2017).

Definition 6 (Element-wise lifting). *Let $\mathcal{F} = \langle Arg, Att \rangle$ be an AF. Then, for a classical argumentation semantics σ , the element-wise- σ semantics ELM- σ is defined by*

$$\llbracket \mathcal{F} \rrbracket_{\text{ELM-}\sigma} := \{\mu \in Distr(\mathcal{F}) : Arg_\beta \in [\mathcal{F}]_\sigma \text{ for all } \beta \in \text{Supp}(\mu)\}.$$

For instance, $\llbracket \mathcal{F} \rrbracket_{\text{ELM-ADM}}$ comprises exactly those distributions where only the assignments corresponding to the admissible sets of \mathcal{F} may have positive probability. Conversely, ELM- σ semantics enforce $\mu(\beta_{\mathcal{S}}) = 0$ for all argument sets $\mathcal{S} \in 2^{Arg} \setminus [\mathcal{F}]_\sigma$ that are not part of the classical argumentation semantics σ .

Lemma 7. *Element-wise lifting of classical semantics yields conservative extensions.*

Proof. Let σ be a classical semantics and $\mathcal{S} \subseteq \text{Arg}$. According to Definition 5, we have to show that $\mathcal{S} \in [\mathcal{F}]_\sigma$ iff $\text{Dirac}_{\mathcal{S}} \in \llbracket \mathcal{F} \rrbracket_{\text{ELM-}\sigma}$.

(\Rightarrow): Let $\mathcal{S} \in [\mathcal{F}]_\sigma$. We have to show $\text{Arg}_\beta \in [\mathcal{F}]_\sigma$ for all assignments $\beta \in \text{Supp}(\text{Dirac}_{\mathcal{S}})$. This holds as the support of $\text{Dirac}_{\mathcal{S}}$ only contains the characteristic function $\text{id}_{\mathcal{S}}$ of \mathcal{S} with $\text{Arg}_{\text{id}_{\mathcal{S}}} = \mathcal{S}$ and thus $\text{Arg}_{\text{id}_{\mathcal{S}}} \in [\mathcal{F}]_\sigma$ by assumption.

(\Leftarrow): Let $\text{Dirac}_{\mathcal{S}} \in \llbracket \mathcal{F} \rrbracket_{\text{ELM-}\sigma}$. Then, $\text{Arg}_{\text{id}_{\mathcal{S}}} \in [\mathcal{F}]_\sigma$ since $\text{Supp}(\text{Dirac}_{\mathcal{S}}) = \{\text{id}_{\mathcal{S}}\}$ and thus, by $\text{Arg}_{\text{id}_{\mathcal{S}}} = \mathcal{S}$ we obtain $\mathcal{S} \in [\mathcal{F}]_\sigma$. \square

The converse of Lemma 7 does not hold, i.e., the notion of conservative extension is more liberal than element-wise lifting.

3.1 Assignment Distribution Properties

Before we discuss probabilistic extensions of admissible and complete semantics, we define the notions of conflict-freeness and defense in the probabilistic setting.

3.1.1 ALMOST-SURELY CONFLICT-FREE

Recall that a set of arguments \mathcal{S} is said to be conflict-free if there is no attack $(A, B) \in \text{Att}$ for all $A, B \in \mathcal{S}$. By the one-to-one correspondence of argument sets to assignments, an assignment β is thus conflict-free if there is no $(A, B) \in \text{Att}$ with $\beta(A) = \beta(B) = \top$. This notion of conflict-freeness naturally transfers to distributions over assignments.

Definition 8 (Almost-surely conflict-free). *Let $\mathcal{F} = \langle \text{Arg}, \text{Att} \rangle$ be an AF. A distribution $\mu \in \text{Distr}(\mathcal{F})$ is called almost-surely conflict-free iff*

$$\mu(A \wedge B) = 0 \quad \text{for all } (A, B) \in \text{Att}. \quad (\text{ASCF})$$

Notably, ASCF does not imply that attacker and attackee cannot both have nonzero μ -likelihood. However, the probability of both being jointly accepted (constituting a classical conflict) is zero. As it turns out, this definition of almost-sure conflict-freeness coincides with the element-wise lifting of conflict-free argument sets.

Lemma 9 (ASCF \Leftrightarrow ELM-CF). *For an AF $\mathcal{F} = \langle \text{Arg}, \text{Att} \rangle$, a distribution $\mu \in \text{Distr}(\mathcal{F})$ is almost-surely conflict-free iff the set Arg_β is conflict-free for each assignment $\beta \in \text{Supp}(\mu)$.*

Proof. (\Rightarrow): Suppose μ is almost-surely conflict-free and assume by contradiction that there is some assignment $\beta \in \text{Supp}(\mu)$ that is not conflict-free. Then, there are arguments $A, B \in \text{Arg}_\beta$ such that A attacks B and $\beta(A) = \beta(B) = \top$. As μ satisfies ASCF, we have $\mu(A \wedge B) = 0$. The assignment β satisfies $A \wedge B$, so $\mu(A \wedge B) \geq \mu(\beta)$ holds. However, by $\beta \in \text{Supp}(\mu)$, we have $\mu(\beta) > 0$, which directly leads to a contradiction:

$$0 = \mu(A \wedge B) \geq \mu(\beta) > 0.$$

(\Leftarrow): We now assume that each $\beta \in \text{Supp}(\mu)$ is conflict-free but μ is not almost-surely conflict-free. Then there is some pair $(A, B) \in \text{Att}$ with $\mu(A \wedge B) > 0$, and thus an assignment β' with $\beta'(A) = \beta'(B) = \top$ and $\mu(\beta') > 0$. However, then $\beta' \in \text{Supp}(\mu)$ but β' is not conflict-free, a contradiction. \square

By Lemma 7 and 9, the set of almost-surely conflict-free distributions conservatively extends the set of conflict-free argument sets.

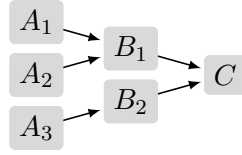


Figure 3: AF fragment for an argument C with attackers B_i and defenders A_i

3.1.2 ALMOST-SURE DEFENSE

In the classical setting, an argument is defended by defeating each of its attackers. Likewise, we say that a distribution almost-surely defends an argument if all attackers are in turn attacked with probability one.

Definition 10 (Almost-sure defense). *For $\mathcal{F} = \langle \text{Arg}, \text{Att} \rangle$ an AF, a distribution $\mu \in \text{Distr}(\mathcal{F})$ almost-surely defends an argument $C \in \text{Arg}$ iff*

$$\mu(\Delta(C)) = 1.$$

The set of all arguments defended almost surely by μ is denoted by $\text{asDefend}(\mu)$.

Example 11. *In the AF from Figure 3, argument C is almost-surely defended by a distribution μ if the following constraint holds:*

$$\mu(\Delta(C)) = \mu\left(\bigwedge_{B \in \neg C} \bigvee_{A \in \rightarrow B} A\right) = \mu((A_1 \vee A_2) \wedge A_3) = 1.$$

That is, the event where argument A_3 as well as arguments A_1 or A_2 are accepted needs to occur almost-surely. This entails that $\mu(A_3) = 1$ is required for C to be almost-surely defended by μ . Argument B_1 is attacked but has no defenders. Recalling that an empty disjunction stands for F , this yields

$$\mu(\Delta(B_1)) = \mu\left(\bigwedge_{A \in \rightarrow B_1} \bigvee_{D \in \neg A} D\right) = \mu(\text{F} \wedge \text{F}) = 0 \neq 1.$$

Thus, B_1 (and likewise B_2) cannot be almost-surely defended by any distribution. Finally, for unattacked arguments like A_1 , we always get

$$\mu(\Delta(A_1)) = \mu(\top) = 1.$$

Initial arguments are thus almost-surely defended by any distribution.

Equivalently, the set $\text{asDefend}(\mu)$ is also characterized by the arguments that are defended by each assignment in the support of μ .

Lemma 12. *Let $\mathcal{F} = \langle \text{Arg}, \text{Att} \rangle$ be an AF and $\mu \in \text{Distr}(\mathcal{F})$. Then:*

$$\text{asDefend}(\mu) = \bigcap_{\beta \in \text{Supp}(\mu)} \text{Defend}(\text{Arg}_\beta).$$

Proof. (\subseteq): Let $C \in \text{asDefend}(\mu)$ and $\beta \in \text{Supp}(\mu)$. To prove $C \in \text{Defend}(\text{Arg}_\beta)$, we pick an attacker B of C and show $B \in \text{Arg}_\beta^{\rightarrow}$. By definition of the almost-surely defended sets, we have $\mu(\Delta(C)) = 1$. This implies $\beta(A) = \tau$ for at least one attacker A of B . Thus, $A \in \text{Arg}_\beta$ and $B \in \text{Arg}_\beta^{\rightarrow}$.

(\supseteq): Let C be an argument such that $C \in \text{Defend}(\text{Arg}_\beta)$ for all $\beta \in \text{Supp}(\mu)$. Then for each attacker $B \in \neg C$ and each $\beta \in \text{Supp}(\mu)$ there is at least one argument $A \in \neg B$ such that $\beta(A) = \tau$. This yields $\mu(\Delta(C)) = 1$ for each $B \in \neg C$ and thus, $C \in \text{asDefend}(\mu)$. \square

Corollary 13. *For any AF $\mathcal{F} = \langle \text{Arg}, \text{Att} \rangle$ and distribution $\mu \in \text{Distr}(\mathcal{F})$:*

(a) *asDefend(μ) is conflict-free if μ is almost-surely conflict-free;*

(b) *asDefend(Dirac \mathcal{S}) = Defend(\mathcal{S}) for each $\mathcal{S} \subseteq \text{Arg}$.*

Proof. (a): If μ is almost-surely conflict-free, then each assignment in $\text{Supp}(\mu)$ is conflict-free (Lemma 9). But then the claim follows from Lemma 12 and the facts that (1) $\text{Defend}(\text{Arg}_\beta)$ is conflict-free if β is conflict-free (cf. Lemma 3) and (2) the intersection of conflict-free argument sets is conflict-free.

(b): Let $\mu = \text{Dirac}_{\mathcal{S}}$ and $\beta = \text{id}_{\mathcal{S}}$. Thus, β is the unique assignment with $\text{Arg}_\beta = \mathcal{S}$ and μ is the (unique) Dirac distribution with $\text{Supp}(\mu) = \{\beta\}$. But then Lemma 12 directly yields $\text{asDefend}(\mu) = \text{Defend}(\mathcal{S})$. \square

In other words, an argument set \mathcal{S} defends an argument C if and only if $\text{Dirac}_{\mathcal{S}}$ almost-surely defends C .

3.1.3 RELATIVE DEFENSE CONSTRAINT

Several of the semantics proposed in the next sections will impose constraints relating the probability of an argument to the probability of the argument being almost-surely defended. For a distribution μ and a comparison relation $\sim \in \{\leq, =, \geq\}$, we define the *relative defense constraint* ($\sim_\mu \Delta$) as the set of constraints

$$\mu(C) \sim \mu(\Delta(C))$$

for all arguments $C \in \text{Arg}$. For example, the constraint ($\leq_\mu \Delta$) is satisfied if for all arguments the probability that the argument is defended is at least as high as the likelihood of the argument itself.

3.2 Admissibility

Recall that a set of arguments \mathcal{S} is called admissible if and only if \mathcal{S} is conflict-free and defends all its elements. In this section, we provide several notions of admissibility in the probabilistic setting. They all have in common to require almost-surely conflict-free distributions but have different approaches to capture the “self-defense” aspect of admissibility.

The first one directly reflects the classical definition and states that almost-surely accepted arguments need to be defended almost-surely.

Definition 14 (WADM). *Let $\mathcal{F} = \langle \text{Arg}, \text{Att} \rangle$ be an AF and $\mu \in \text{Distr}(\mathcal{F})$ an almost-surely conflict-free distribution. Then μ is called weakly admissible iff*

$$\text{Arg}_\mu \subseteq \text{asDefend}(\mu).$$

In other words, an almost-surely conflict-free distribution μ is weakly admissible iff for all arguments C , we have that

$$\mu(C) = 1 \quad \text{implies} \quad \mu(\Delta(C)) = 1.$$

The semantics is *weak* in the sense that all distributions where no argument has a likelihood of one are trivially weakly admissible. As such, weak admissibility is more of a basic requirement like almost-sure conflict-freeness and also *weak* in the sense that it is implied by all our other probabilistic notions of admissibility (which we show in the following). Our notion of weak admissibility should hence not be confused with the recent notion of weak admissibility in the non-probabilistic setting (Baumann, Brewka, & Ulbricht, 2020). This non-probabilistic notion focuses on relaxing requirements on defense, e.g., to not require self-attacking arguments to be defended in case they also attack the set of arguments under consideration.

Towards a stronger probabilistic notion of admissibility, we relate the likelihood of each argument to the probability of its strongest attacker, or, equivalently, the attacker against which the defense is weakest.

Definition 15 (MINADM). *Let $\mathcal{F} = \langle \text{Arg}, \text{Att} \rangle$ be an AF and $\mu \in \text{Distr}(\mathcal{F})$ an almost-surely conflict-free distribution. Then μ is called min-admissible iff for all $C \in \text{Arg}$, we have*

$$\mu(C) \leq \min_{B \in \neg C} \mu\left(\bigvee_{A \in \neg B} A\right).$$

That is, for every argument C the probability of being accepted is bounded from above by the lowest probability of C being defended against any of its attackers. Equivalently,

$$\mu(C) \leq \mu\left(\bigvee_{A \in \neg B} A\right)$$

needs to hold for each attacker $B \in \neg C$.

Example 16. *Consider the AF from Figure 3 and assume we have an almost-surely conflict-free distribution μ with $\mu(A_1 \vee A_2) = 0.7$ and $\mu(A_3) = 0.4$. Then min-admissibility requires $\mu(C) \leq 0.4$ as this is the probability with which C is defended against its strongest attacker, B_2 . This entails $\mu(C) \leq 0.7 = \mu(A_1 \vee A_2)$ also for the remaining attacker B_1 .*

Next, instead of considering each attacker separately, we require that the probability of defense against the joint attacks is an upper bound on each argument's likelihood. This is a first instance of the relative defense constraint introduced in Section 3.1.3.

Definition 17 (JNTADM). *Let $\mathcal{F} = \langle \text{Arg}, \text{Att} \rangle$ be an AF. A distribution $\mu \in \text{Distr}(\mathcal{F})$ is called joint-attack admissible iff μ is almost-surely conflict-free and $(\leq_{\mu} \Delta)$ holds, i.e., for all $C \in \text{Arg}$, we have*

$$\mu(C) \leq \mu(\Delta(C)).$$

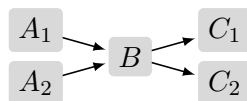


Figure 4: Under PRADM, $\mu(C_1 \vee C_2) \leq \mu(A_1 \vee A_2)$ needs to hold

Example 18. *Continuing Example 11, joint-admissibility requires for argument C that*

$$\mu(C) \leq \mu((A_1 \vee A_2) \wedge A_3).$$

For arguments B_1 and B_2 , we have $\mu(\Delta(B_1)) = \mu(\Delta(B_2)) = 0$, so $\mu(B_1) = \mu(B_2) = 0$ needs to hold for all joint-admissible distributions. Note that for initial arguments like A_1 , the resulting constraint

$$\mu(A_1) \leq \mu(\Delta(A_1)) = 1$$

is trivially fulfilled by every distribution.

For our last notion of admissibility we shift the view from arguments C that need to be defended to arguments B which are attackers of such arguments.

Definition 19 (PRADM). *Let $\mathcal{F} = \langle Arg, Att \rangle$ be an AF. A distribution $\mu \in \text{Distr}(\mathcal{F})$ is called probabilistically admissible iff μ is almost-surely conflict-free and for all $B \in Arg$, we have*

$$\mu\left(\bigvee_{C \in B \rightarrow} C\right) \leq \mu\left(\bigvee_{A \in \rightarrow B} A\right).$$

Recall that if an argument set \mathcal{S} is admissible, then each argument B attacking arguments in \mathcal{S} is itself attacked by some argument in \mathcal{S} . Likewise, if B attacks arguments that are almost-surely accepted, we would expect that B is attacked with likelihood one, which is implied by weak admissibility. Now PRADM further generalizes this relation and requires for each argument B that the probability of accepting arguments attacked by B is bounded by the probability of accepting the arguments which themselves attack B .

Example 20. *On the AF from Figure 3, PRADM yields the same constraints as JNTADM (see Example 18). This changes when considering the AF from Figure 4 where argument B attacks two arguments C_1 and C_2 . For probabilistically admissible distributions μ , we get for B the constraint*

$$\mu(C_1 \vee C_2) \leq \mu(A_1 \vee A_2),$$

whereas joint-attack admissible distributions need to satisfy both $\mu(C_1) \leq \mu(A_1 \vee A_2)$ and $\mu(C_2) \leq \mu(A_1 \vee A_2)$.

Finally, Definition 6 provides the notion of *element-wise admissibility*, which requires that all assignments in the support of a distribution μ are admissible. This can be rephrased as a conditional probability constraint: For each argument C with positive μ -likelihood, the probability of assignments defending C must equal one when conditioned on the event that C is accepted.

Lemma 21. *Let $\mathcal{F} = \langle Arg, Att \rangle$ and $\mu \in \text{Distr}(\mathcal{F})$. Then μ is element-wise admissible iff μ is almost-surely conflict-free and for each argument $C \in Arg$ with $\mu(C) > 0$, it holds*

$$\mu(\Delta(C) \mid C) = 1.$$

Proof. As $\mu(C) > 0$, the constraint $\mu(\Delta(C) \mid C) = 1$ is equivalent to $\mu(\Delta(C) \wedge C) = \mu(C)$, so it suffices to show that $\Delta(C) \wedge C$ and C describe the same event iff μ is element-wise admissible.

(\Rightarrow): Let μ be element-wise admissible and $C \in Arg$. Then for all assignments $\beta \in \text{Supp}(\mu)$ we have $Arg_\beta \subseteq \text{Defend}(Arg_\beta)$. For those β with $\beta(C) = \top$, we thus have $C \in \text{Defend}(Arg_\beta)$. Recall that $\Delta(C)$ is the set of all assignments whose argument sets defend C , so we have $\beta \in \Delta(C)$.

(\Leftarrow): Let $\mu(\Delta(C) \wedge C) = \mu(C)$ for all $C \in Arg$ and assume by contradiction there exists a $\beta \in \text{Supp}(\mu)$ with Arg_β being not admissible. Then either Arg_β is not conflict-free (contradiction) or $Arg_\beta \not\subseteq \text{Defend}(Arg_\beta)$, i.e., there is a $C' \in Arg_\beta$ that is not defended by Arg_β . But then $\beta \notin \Delta(C')$, which contradicts $\mu(\Delta(C') \wedge C') = \mu(C')$. \square

All five presented admissibility notions for distributions are conservative extensions of the non-probabilistic notion of admissibility, as stated in the following lemma.

Lemma 22 (Conservative extension ADM). *Let $\mathcal{F} = \langle Arg, Att \rangle$ be an AF and $\mathcal{S} \subseteq Arg$ be conflict-free. Then, \mathcal{S} is admissible iff $\text{Dirac}_{\mathcal{S}}$ is κ admissible for $\kappa \in \{\text{element-wise, probabilistically, joint-attack, min-, weakly}\}$.*

Proof. As weak admissibility is the weakest notion and element-wise admissibility the strongest one (in anticipation of Lemma 23, 24, and 25), it suffices to show the implications $\text{ADM} \Rightarrow \text{ELM-ADM}$ and $\text{WADM} \Rightarrow \text{ADM}$. The first implication is obvious by Lemma 7. For the second implication we suppose that $\text{Dirac}_{\mathcal{S}}$ is weakly admissible. Hence:

$$\mathcal{S} = Arg_{\text{Dirac}_{\mathcal{S}}} \subseteq \text{asDefend}(\text{Dirac}_{\mathcal{S}}).$$

Corollary 13 yields:

$$\text{asDefend}(\text{Dirac}_{\mathcal{S}}) = \text{Defend}(\mathcal{S}).$$

Thus, $\mathcal{S} \subseteq \text{Defend}(\mathcal{S})$ holds and \mathcal{S} is admissible. \square

3.2.1 RELATIONSHIPS

We now investigate how our probabilistic notions of admissibility relate to each other. First, we provide implications that hold between the different semantics and then give concrete examples that show these implications to be strict.

Element-wise admissibility is the most restricted variant, i.e., each distribution that satisfies ELM-ADM also satisfies the other notions of admissibility. Exemplarily, for an attack $(B, C) \in Att$, whenever $\beta(C) = \top$ for some admissible assignment β , there must be some attacker A of B with $\beta(A) = \top$ as well. Thus, all admissible assignments that contribute to the probability on the left-hand side of the PRADM constraint also appear on the right, so all element-wise admissible distributions are also probabilistically admissible.

Lemma 23 (ELM-ADM \Rightarrow PRADM, ELM-ADM \Rightarrow JNTADM). *Let \mathcal{F} be an AF and $\mu \in \text{Distr}(\mathcal{F})$. If μ is element-wise admissible, then μ is probabilistically and joint-attack admissible.*

Proof. (ELM-ADM \Rightarrow PRADM): Let μ be element-wise admissible. Then each assignment $\beta \in \text{Supp}(\mu)$ is admissible, so given an attack $(B, C) \in \text{Att}$, we have:

$$\text{If } \beta(C) = \top \text{ then } \beta(A) = \top \text{ for some attacker } A \text{ of } B.$$

Thus, if β is an assignment satisfying $\bigvee_{C \in B \rightarrow} C$, then β also satisfies $\bigvee_{A \in \rightarrow B} A$. Hence, PRADM holds, which shows that μ is probabilistically admissible.

(ELM-ADM \Rightarrow JNTADM): Since all assignments $\beta \in \text{Supp}(\mu)$ are admissible, it suffices to show that for each admissible β and $C \in \text{Arg}$, we have:

$$\text{If } \beta(C) = \top \text{ then for each } B \in \rightarrow C \text{ there is some } A \in \rightarrow B \text{ s.t. } \beta(A) = \top.$$

This, however, is clear as the argument set $\text{Arg}_\beta = \{C \in \text{Arg} : \beta(C) = \top\}$ defends all its arguments. So, $\beta(C) = \top$ implies that for each attacker B of C (i.e., each argument $B \in \rightarrow C$) there is an attacker of B in Arg_β (i.e., there is some argument $A \in \rightarrow B$ with $\beta(A) = \top$). \square

Lemma 24 (PRADM \Rightarrow MINADM, JNTADM \Rightarrow MINADM). *Let \mathcal{F} be an AF and $\mu \in \text{Distr}(\mathcal{F})$. If μ is probabilistically or joint-attack admissible, then μ is min-admissible.*

Proof. (PRADM \Rightarrow MINADM): Let μ be probabilistically admissible. For any argument C , we have to show

$$\mu(C) \leq \min_{B \in \rightarrow C} \mu\left(\bigvee_{A \in \rightarrow B} A\right).$$

For this it suffices to pick any attacker $B \in \rightarrow C$ and show that $\mu(C) \leq \mu(\bigvee_{A \in \rightarrow B} A)$. But this follows directly from PRADM as $C \in B \rightarrow$, which yields:

$$\mu(C) \leq \mu\left(\bigvee_{C \in B \rightarrow} C\right) \leq \mu\left(\bigvee_{A \in \rightarrow B} A\right).$$

(JNTADM \Rightarrow MINADM): Let μ be joint-attack admissible, i.e., the relative defense constraint ($\leq_\mu \Delta$) is satisfied. Then we have

$$\mu(C) \leq \mu(\Delta(C)) = \mu\left(\bigwedge_{B \in \rightarrow C} \bigvee_{A \in \rightarrow B} A\right) \leq \min_{B \in \rightarrow C} \mu\left(\bigvee_{A \in \rightarrow B} A\right)$$

because each assignment β which satisfies $\Delta(C)$ also satisfies $\bigvee_{A \in \rightarrow B} A$ for all $B \in \rightarrow C$. \square

Note that the notions joint-attack and min-admissibility collapse in AFs where each argument has at most one attacker.

As weak admissibility only imposes constraints on the likelihood of arguments belonging to Arg_μ , it is strictly weaker than min-admissibility, and hence the most liberal of the considered admissibility notions.

Lemma 25 (MINADM \Rightarrow WADM). *Let $\mathcal{F} = \langle \text{Arg}, \text{Att} \rangle$ be an AF and $\mu \in \text{Distr}(\mathcal{F})$. If μ is min-admissible, then μ almost-surely defends all arguments in Arg_μ .*

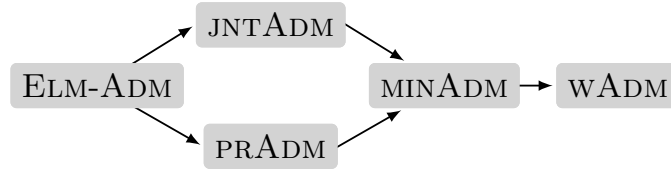


Figure 5: Hierarchy of the admissible semantics from Definitions 14, 15, 17, and 19

Proof. Suppose $C \in \text{Arg}_\mu$. Then, $\mu(C) = 1$. The task is to show that each attacker of C is attacked almost surely. Pick an attacker B of C . As $\mu(C) = 1$, the constraint MINADM directly yields

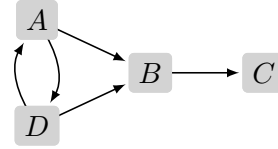
$$\mu\left(\bigvee_{A \in \neg B} A\right) = 1.$$

Thus, μ almost-surely defends C . □

The results so far establish a hierarchy of the five notions of admissibility for distributions as illustrated in Figure 5. The following examples show the inclusions within this hierarchy to be strict.

Example 26 (MINADM $\not\equiv$ PRADM, MINADM $\not\equiv$ JNTADM). *As an example for a min-admissible distribution that is not element-wise admissible and neither joint-attack nor probabilistically admissible, let μ be the distribution for the AF depicted on the right with the probabilities given on the left:*

$$\begin{aligned} \mu(\neg A \wedge \neg B \wedge C \wedge D) &= 1/3 \\ \mu(A \wedge \neg B \wedge \neg C \wedge \neg D) &= 1/4 \\ \mu(\neg A \wedge B \wedge \neg C \wedge \neg D) &= 1/6 \\ \mu(\neg A \wedge \neg B \wedge C \wedge \neg D) &= 1/4 \end{aligned}$$



That is, $\mu(\beta) = 0$ for all other assignments β . Then μ is almost-surely conflict-free (since all assignments in its support are conflict-free, cf. Definition 8) and satisfies MINADM:

- Argument C has likelihood $\mu(C) = 1/3 + 1/4 = 7/12$ and a single attacker B , which is attacked by A and D with probability $1/4 + 1/3 = 7/12$.
- Argument B has likelihood $\mu(B) = 1/6$. Its attackers A and D are attacked with probability $1/3$ resp. $1/4$.
- Argument D has likelihood $\mu(D) = 1/3$, and so is the probability for its attacker A to be attacked. The analogous statement holds for A .

So, μ is min-admissible, but μ is not element-wise admissible as $\neg A \wedge \neg B \wedge C \wedge \neg D$ and $\neg A \wedge B \wedge \neg C \wedge \neg D$ induce the non-admissible argument sets $\{C\}$ and $\{B\}$, respectively, that do not defend their arguments.

Further, μ is not joint-attack admissible, since the argument B has no joint attack of its attackers A and D :

$$\mu(B) = \frac{1}{6} \not\equiv 0 = \mu(A \wedge D) = \mu(\Delta(B)).$$

Indeed, in this example there is no admissible argument set containing B , since B 's attackers A and D attack each other.

To see why μ is not probabilistically admissible, consider argument D . We have $D^\rightarrow = \{A, B\}$ and ${}^\rightarrow D = \{A\}$. But then

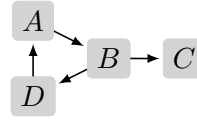
$$\mu\left(\bigvee_{C' \in D^\rightarrow} C'\right) = \mu(A \vee B) = \frac{1}{4} + \frac{1}{6} \not\leq \frac{1}{4} = \mu(A) = \mu\left(\bigvee_{A \in {}^\rightarrow D} A\right).$$

Thus, the constraint PRADM is violated for argument D . Analogously, PRADM does not hold for A .

The following two examples illustrate that probabilistic admissibility and joint-attack admissibility are incomparable.

Example 27 (JNTADM $\not\equiv$ PRADM). Consider the AF depicted on the right below and the distribution μ with the following probabilities for assignments in its support:

$$\begin{aligned} \mu(A \wedge \neg B \wedge C \wedge \neg D) &= 1/3 \\ \mu(\neg A \wedge B \wedge \neg C \wedge \neg D) &= 1/3 \\ \mu(\neg A \wedge \neg B \wedge \neg C \wedge D) &= 1/3 \end{aligned}$$



μ obviously is almost-surely conflict-free and the μ -likelihood of all four arguments is $1/3$. As each argument has exactly one attacker, μ is joint-attack admissible: For example, the JNTADM constraint is satisfied for argument C since ${}^\rightarrow C = \{B\}$, ${}^\rightarrow B = \{A\}$, and

$$\mu(C) = \frac{1}{3} \leq \frac{1}{3} = \mu(A) = \mu\left(\bigvee_{A' \in {}^\rightarrow B} A'\right) = \mu(\Delta(C)).$$

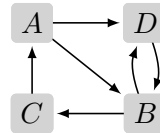
However, μ does not satisfy the constraint PRADM for argument B :

$$\mu\left(\bigvee_{C' \in B^\rightarrow} C'\right) = \mu(C \vee D) = \frac{1}{3} + \frac{1}{3} \not\leq \frac{1}{3} = \mu(A) = \mu\left(\bigvee_{A' \in {}^\rightarrow B} A'\right).$$

Hence, μ is not probabilistically admissible.

Example 28 (PRADM $\not\equiv$ JNTADM). Consider the following almost-surely conflict-free distribution μ for the AF depicted on the right:

$$\begin{aligned} \mu(A \wedge \neg B \wedge \neg C \wedge \neg D) &= 1/4 \\ \mu(\neg A \wedge B \wedge \neg C \wedge \neg D) &= 1/4 \\ \mu(\neg A \wedge \neg B \wedge C \wedge D) &= 1/4 \\ \mu(\neg A \wedge \neg B \wedge C \wedge \neg D) &= 1/4 \end{aligned}$$



We have $\mu(A) = \mu(B) = \mu(D) = 1/4$, and $\mu(C) = 1/2$. This implies that Arg_μ is empty, such we can immediately infer that μ is weakly admissible. To see why μ is probabilistically admissible, we have to check that μ satisfies PRADM for all arguments:

- For argument A , we have $A^\rightarrow = \{B, D\}$ and $\neg A = \{C\}$, such that $\mu(B \vee D) \leq \mu(C)$ has to hold. This is in fact the case due to $\mu(B \vee D) = \mu(C) = 1/2$.
- For argument B , we have $B^\rightarrow = \{C, D\}$, $\neg B = \{A, D\}$, and $\mu(C \vee D) = 1/2 \leq 1/2 = \mu(A \vee D)$.
- For argument C , we have $C^\rightarrow = \{A\}$, $\neg C = \{B\}$, and $\mu(A) = 1/4 \leq 1/4 = \mu(B)$.
- For argument D , we have $D^\rightarrow = \{B\}$, $\neg D = \{A, B\}$, and $\mu(B) = 1/4 \leq 1/2 = \mu(A \vee B)$.

Thus, μ is probabilistically admissible.

To see why μ is not joint-attack admissible, we observe that the μ -likelihood of argument B is positive, but there is no joint attack on B 's attackers A and D . More precisely:

$$\mu(B) = \frac{1}{4} \not\leq 0 = \mu(C \wedge (A \vee B)) = \mu(\Delta(B)).$$

Here, C stands for the attacks on A , and $A \vee B$ for the attacks on D . This shows that μ violates the constraint JNTADM for argument B .

3.3 Complete Semantics

In the non-probabilistic setting, complete semantics imposes stronger constraints than admissibility, as it additionally requires that all arguments defended by a set $\mathcal{S} \subseteq \text{Arg}$ are contained in \mathcal{S} , i.e., $\mathcal{S} = \text{Defend}(\mathcal{S})$. Based on the notions of admissibility for distributions from the previous section, we now extend complete semantics towards five notions in the probabilistic setting. Again, ELM-CMP is an instantiation of Definition 6.

Definition 29. Let $\mathcal{F} = \langle \text{Arg}, \text{Att} \rangle$ be an AF and $\mu \in \text{Distr}(\mathcal{F})$ an almost-surely conflict-free distribution. Then μ is called

(WCMP) weakly complete iff $\text{Arg}_\mu = \text{asDefend}(\mu)$,

(MINCMP) min-complete iff μ satisfies MINADM and $(\geq_\mu \Delta)$,

(JNTCMP) joint-attack complete iff μ satisfies $(=_\mu \Delta)$,

(PRCMP) probabilistically complete iff μ satisfies PRADM and $(\geq_\mu \Delta)$, and

(ELM-CMP) element-wise complete iff $\mu \in \llbracket \mathcal{F} \rrbracket_{\text{ELM-CMP}}$.

It is clear that each of the five notions of complete semantics imposes stricter constraints than the corresponding notion of admissibility, so each κ complete distribution is also κ admissible for $\kappa \in \{\text{element-wise, weakly, probabilistically, min-, joint-attack}\}$. Before we establish the relationships between the five semantics, we remark on the particular choice of constraints for min- and probabilistically complete semantics.

Remark 30. Joint-attack and weakly complete semantics have been defined by requiring, respectively, the joint-attack admissibility constraint $(\leq_\mu \Delta)$ and the weak admissibility constraint $\text{Arg}_\mu \subseteq \text{asDefend}(\mu)$, plus the analogous “reverse condition”, namely $(\geq_\mu \Delta)$ and

$Arg_\mu \supseteq \text{asDefend}(\mu)$, resulting in the equality constraints as stated in Definition 29. One might wonder why min- and probabilistically complete semantics are not defined in an analogous way. We argue that the resulting notions would be too strong. For probabilistically complete semantics, the requirement

$$\mu\left(\bigvee_{C \in B^\rightarrow} C\right) = \mu\left(\bigvee_{A \in \rightarrow B} A\right)$$

would imply for arguments B which do not attack other arguments (i.e., $B^\rightarrow = \emptyset$) that all their attackers $C \in \rightarrow B$ are unacceptable under μ , as for those $\mu(\bigvee_{A \in \rightarrow B} A) = 0$ would need to hold. This condition appears very strong and lacks an intuitive meaning in the spirit of the non-probabilistic complete semantics. Likewise, for min-complete semantics the condition

$$\mu(C) = \min_{B \in \rightarrow C} \mu\left(\bigvee_{A \in \rightarrow B} A\right)$$

also appears questionable as a criterion for complete semantics as it implies that each argument C with likelihood 0 has at least one attacker B that is not attacked under μ (in the sense that $\mu(\bigvee_{A \in \rightarrow B} A) = 0$ holds). This needlessly excludes cases where more than one argument contributes to the attack on C , but none of them is completely unattacked under μ .

Analogous to Lemma 21, element-wise complete semantics can be characterized by element-wise admissibility and the constraint that any defended argument has a conditional likelihood of one.

Lemma 31. *Let $\mathcal{F} = \langle Arg, Att \rangle$ and $\mu \in \text{Distr}(\mathcal{F})$. Then μ is element-wise complete iff μ is element-wise admissible and for all arguments $C \in Arg$ where event $\Delta(C)$ has positive probability under μ we have*

$$\mu(C \mid \Delta(C)) = 1.$$

Proof. As $\mu(\Delta(C)) > 0$, the conditional probability constraint $\mu(C \mid \Delta(C)) = 1$ is equivalent to $\mu(C \wedge \Delta(C)) = \mu(\Delta(C))$. Thus it suffices to show that $C \wedge \Delta(C)$ and $\Delta(C)$ describe the same assignments iff μ is element-wise complete.

(\Rightarrow): Let μ be element-wise complete and $C \in Arg$. Then for all assignments $\beta \in \text{Supp}(\mu)$, we have $Arg_\beta = \text{Defend}(Arg_\beta)$. Now if $\beta \in \Delta(C)$, then $C \in \text{Defend}(Arg_\beta)$ and hence $C \in Arg_\beta$, leading to β also satisfying $C \wedge \Delta(C)$.

(\Leftarrow): Let μ be element-wise admissible, $\mu(C \wedge \Delta(C)) = \mu(\Delta(C))$ for all $C \in Arg$, and assume by contradiction there exists a $\beta \in \text{Supp}(\mu)$ such that Arg_β is not complete, i.e., $Arg_\beta \neq \text{Defend}(Arg_\beta)$. Since β is admissible, we have $Arg_\beta \not\supseteq \text{Defend}(Arg_\beta)$, i.e., there is an argument C' which is defended by Arg_β but not element of the set. But then β only satisfies $\Delta(C')$ but not $C' \wedge \Delta(C')$, which contradicts $\mu(C' \wedge \Delta(C')) = \mu(\Delta(C'))$ because $\mu(\beta) > 0$ due to $\beta \in \text{Supp}(\mu)$. \square

All five complete semantics from Definition 29 are conservative extensions of the non-probabilistic complete semantics for argument sets.

Lemma 32 (Conservative extension CMP). *Let $\mathcal{F} = \langle Arg, Att \rangle$ be an AF and $\mathcal{S} \subseteq Arg$. Then, \mathcal{S} is complete iff $Dirac_{\mathcal{S}}$ is κ complete for $\kappa \in \{\text{element-wise, weakly, probabilistically, min-, joint-attack}\}$.*

Proof. By the relationship results between the different complete semantics established in Section 3.3.1, it suffices to prove the following two implications:

\mathcal{S} is complete $\Rightarrow Dirac_{\mathcal{S}}$ is element-wise complete and

$Dirac_{\mathcal{S}}$ is weakly complete $\Rightarrow \mathcal{S}$ is complete.

The first implication is obvious as $Dirac_{\mathcal{S}}$ is element-wise complete iff $id_{\mathcal{S}}$ is a complete assignment, which again is equivalent to \mathcal{S} being a complete argument set. For the second implication, suppose that $Dirac_{\mathcal{S}}$ is weakly complete. Using Corollary 13 we obtain:

$$Defend(\mathcal{S}) = asDefend(Dirac_{\mathcal{S}}) = Arg_{Dirac_{\mathcal{S}}} = \mathcal{S}.$$

This directly yields that \mathcal{S} is complete. □

Recall that Arg_{μ} denotes the set of arguments that are accepted almost-surely under a distribution μ . Given a probabilistic semantics ρ , one may ask if there are arguments that are almost-surely accepted under all ρ -distributions. For the various admissibility semantics, the answer is negative as the Dirac distribution corresponding to the empty argument set is always part of the semantics and no argument is almost-surely accepted under it. This is no longer the case when turning to the complete semantics from Definition 29: Here, it turns out that the arguments almost-surely accepted under all distributions from one of the semantics are exactly those arguments in the *grounded* argument set of the AF.

Lemma 33. *Let $\mathcal{F} = \langle Arg, Att \rangle$ be an AF with grounded argument set \mathcal{S}_{GR} and let ρ be any of the complete semantics from Definition 29. Then:*

$$\bigcap_{\mu \in \llbracket \mathcal{F} \rrbracket_{\rho}} Arg_{\mu} = \mathcal{S}_{GR}.$$

Proof. (\subseteq): If $A \notin \mathcal{S}_{GR}$, then $Dirac_{\mathcal{S}_{GR}}(A) = 0$. As \mathcal{S}_{GR} is complete, $Dirac_{\mathcal{S}_{GR}}$ is contained in $\llbracket \mathcal{F} \rrbracket_{\rho}$ for every semantics ρ by Lemma 32. Thus, A is not in $\bigcap_{\mu \in \llbracket \mathcal{F} \rrbracket_{\rho}} Arg_{\mu}$.

(\supseteq): Now let $A \in \mathcal{S}_{GR}$. We consider the constructive characterization of the grounded semantics from (Dung, 1995) as the least fixed-point of the $Defend(\cdot)$ function, i.e., $\mathcal{S}_{GR}^0 = Defend(\emptyset)$, $\mathcal{S}_{GR}^i = Defend(\mathcal{S}_{GR}^{i-1})$ for $i > 0$, and $\mathcal{S}_{GR} = \mathcal{S}_{GR}^j$ s.t. $\mathcal{S}_{GR}^j = \mathcal{S}_{GR}^{j+1}$. We show inductively that for every weakly complete distribution μ , we have $\mu(A) = 1$ for each $A \in \mathcal{S}_{GR}^i$.

Base case: Let $A \in \mathcal{S}_{GR}^0 = Defend(\emptyset)$. As A is defended by the empty set, it must be initial. Then A is also almost-surely defended by μ , which implies $\mu(A) = 1$ as μ is weakly complete.

Step of induction: Let $A \in \mathcal{S}_{GR}^i = Defend(\mathcal{S}_{GR}^{i-1})$ for $i > 0$. We can assume $\mu(C) = 1$ for all $C \in \mathcal{S}_{GR}^{i-1}$. Now consider $\mu(\Delta(A)) = \mu(\bigwedge_{B \in \rightarrow A} \bigvee_{C \in \rightarrow B} C)$. As $A \in Defend(\mathcal{S}_{GR}^{i-1})$, we know that for each attacker B of A , there is a defender $C \in \mathcal{S}_{GR}^{i-1}$ with $\mu(C) = 1$. Therefore, $\mu(\Delta(A)) = 1$ which again implies $\mu(A) = 1$.

Thus, for $A \in \mathcal{S}_{GR}$, we have $\mu(A) = 1$ for all weakly complete distributions. As we will see all other notions of completeness imply wCMP, which concludes the proof. □

3.3.1 RELATIONSHIPS

Similar as for the notions of admissibility, we draw relationships between the five complete semantics introduced in Definition 29. This yields connections between them analogously to the case of admissibility (cf. Figure 5).

Lemma 34 (ELM-CMP \Rightarrow JNTCMP, ELM-CMP \Rightarrow PRCMP). *Let \mathcal{F} be an AF and $\mu \in \text{Distr}(\mathcal{F})$. If μ is element-wise complete, then μ is joint-attack and probabilistically complete.*

Proof. Let us first observe that each conflict-free assignment $\beta \in \text{Asg}(\text{Arg})$ is complete iff

$$\forall C \in \text{Arg}. (\beta(C) = \top \Leftrightarrow \forall B \in \neg C. \exists A \in \neg B. \beta(A) = \top).$$

In other words, C is assigned value \top under β iff C is defended in β : $\beta(C) = \top$ iff $\beta \in \Delta(C)$. (ELM-CMP \Rightarrow JNTCMP): Now let μ be element-wise conflict free and $C \in \text{Arg}$. Then all assignments in the support of μ are complete, so we have

$$\mu(C) = \sum_{\beta(C)=\top} \mu(\beta) = \sum_{\beta \in \Delta(C)} \mu(\beta) = \mu(\Delta(C)).$$

Thus, all arguments C satisfy the constraint $(=_{\mu}\Delta)$ and hence μ is joint-attack complete. (ELM-CMP \Rightarrow PRCMP): We have to show that μ is probabilistically admissible and satisfies $(\geq_{\mu}\Delta)$. The latter is clearly implied by μ being joint-attack complete as established above. For the former, recall that element-wise complete semantics implies element-wise admissibility. By Lemma 23, this implies probabilistic admissibility, so μ is probabilistically complete. \square

Lemma 35 (PRCMP \Rightarrow MINCMP, JNTCMP \Rightarrow MINCMP). *Let \mathcal{F} be an AF and $\mu \in \text{Distr}(\mathcal{F})$. If μ is probabilistically or joint-attack complete, then μ is min-complete.*

Proof. Recall that min-complete semantics requires min-admissibility and that the constraint $(\geq_{\mu}\Delta)$ holds. For the case where μ is joint-attack complete, the statement is obvious, since JNTCMP directly implies MINADM and $(\geq_{\mu}\Delta)$. Let us suppose now that μ is probabilistically complete. Then, μ is probabilistically admissible and therefore min-admissible by Lemma 24. Furthermore, by definition of PRCMP, μ also satisfies $(\geq_{\mu}\Delta)$. Thus, μ is min-complete. \square

Lemma 36 (MINCMP \Rightarrow WCMP). *Let \mathcal{F} be an AF and $\mu \in \text{Distr}(\mathcal{F})$. If μ is min-complete, then μ is weakly complete.*

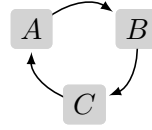
Proof. Let μ be min-complete, i.e., μ satisfies MINADM and $(\geq_{\mu}\Delta)$. Lemma 25 yields $\text{Arg}_{\mu} \subseteq \text{asDefend}(\mu)$ and it remains to show that $\text{asDefend}(\mu) \subseteq \text{Arg}_{\mu}$. Let $C \in \text{asDefend}(\mu)$. Then μ attacks each attacker B of C with probability 1, which is equivalent to $\mu(\Delta(C)) = 1$. Then $(\geq_{\mu}\Delta)$ yields $\mu(C) = 1$ and hence $C \in \text{Arg}_{\mu}$. \square

We now provide an examples illustrating that the implications between the different complete semantics are strict.

Example 37. An example for a distribution that is both joint-attack and probabilistically complete, but without being element-wise complete, is a distribution μ that assigns probability $\frac{1}{3}$ to the assignments $\text{id}_{\{A\}}$, $\text{id}_{\{B\}}$, and $\text{id}_{\{C\}}$ for the simple odd-length cycle AF on the right below. Neither $\{A\}$, $\{B\}$, nor $\{C\}$ are complete sets, so μ is not element-wise complete. Because A , B , and C are isomorphic in the AF under μ , for all constraints imposed by PRCMP and JNTCMP the left-hand side equals the right-hand side, so μ is probabilistically and joint-attack complete.

For an example of a weakly complete distribution that is not complete w.r.t. any of the other four complete semantics, consider the distribution ν provided through the following probabilities of the assignments in its support:

$$\begin{aligned}\nu(A \wedge \neg B \wedge \neg C) &= 1/2 \\ \nu(\neg A \wedge B \wedge \neg C) &= 1/3 \\ \nu(\neg A \wedge \neg B \wedge C) &= 1/6\end{aligned}$$



Then $\text{Arg}_\nu = \emptyset = \text{asDefend}(\nu)$ and hence ν is weakly complete. However,

$$\nu(A) = \frac{1}{2} \not\leq \frac{1}{6} = \nu(B) = \min_{C' \in \neg A} \nu\left(\bigvee_{B' \in \neg C'} B'\right)$$

and thus the MINADM constraint is violated for argument A . To this end, ν is not min-complete and thereby not probabilistically, joint-attack, or element-wise complete.

The distribution of Example 27 is joint-attack complete, but not probabilistically admissible, and therefore not probabilistically complete. Vice versa, the distribution μ of Example 28 is probabilistically complete, but not joint-attack admissible, and therefore not joint-attack complete. Note that μ satisfies the JNTCMP constraint for the arguments A , C , and D but not for B , since JNTADM is violated due to $\mu(B) = 1/4 > 0 = \mu(\Delta(B))$. However, the latter inequation also implies that μ satisfies PRCMP for B and together with PRADM holding for A , C , and D , μ satisfies PRCMP for all arguments.

3.4 Saturation

We recall that in the non-probabilistic setting an argument set $\mathcal{S} \subseteq \text{Arg}$ is called stable if \mathcal{S} is conflict-free and $\mathcal{S} \cup \mathcal{S}^\rightarrow = \text{Arg}$. Equivalently, \mathcal{S} is conflict-free and each $B \in \text{Arg} \setminus \mathcal{S}$ is attacked by \mathcal{S} : $\mathcal{S} \cap \neg B \neq \emptyset$. Stable sets \mathcal{S} can thus be thought of as “perfectly saturated” in the sense that an argument is either in \mathcal{S} or attacked by \mathcal{S} . This is a quite strong constraint, and the existence of stable sets and thus also element-wise stable distributions is not guaranteed. In particular, this is the case for AFs where the underlying graph structure constitutes a simple cycle of odd length, e.g., as in the graph of Example 37. In the classical setting, the *semi-stable* semantics (Caminada, Carnielli, & Dunne, 2006) addresses this issue, being defined as all complete argument sets \mathcal{S} where $\mathcal{S} \cup \mathcal{S}^\rightarrow$ is maximal w.r.t. set inclusion. This entails that if stable argument sets exist, they coincide with the semi-stable sets, and that semi-stable sets always exist even when there is no stable set. However, for distributions such a maximization criterion is not readily expressible as a polynomial constraint. In the following, we provide a different approach of carefully weakening the strong notion of “perfect saturation”.

The idea is to require probabilistic justifications for $\mu(B) = 0$ or $\mu(B) < 1$. *0-saturation* requires that if argument B is accepted with likelihood 0, then B is attacked almost surely, i.e., $\mu(\bigvee_{A \in \rightarrow B} A) = 1$. Likewise, *1-saturation* imposes the constraint that all arguments B where $\mu(B) < 1$ are attacked with positive probability.

Definition 38 (0- and 1-saturated distributions). *Let $\mathcal{F} = \langle \text{Arg}, \text{Att} \rangle$ be an AF and $\mu \in \text{Distr}(\mathcal{F})$ be almost-surely conflict-free. Then μ is called*

(ELM-ST) element-wise stable iff $\mu \in \llbracket \mathcal{F} \rrbracket_{\text{ELM-ST}}$,

(0-SAT) 0-saturated iff $\mu(\bigvee_{A \in \rightarrow B} A) = 1$ for each $B \in \text{Arg}$ with $\mu(B) = 0$,

(1-SAT) 1-saturated iff $\mu(\bigvee_{A \in \rightarrow B} A) > 0$ for each $B \in \text{Arg}$ with $\mu(B) < 1$, and

(0/1-SAT) 0/1-saturated iff μ is 0-saturated and 1-saturated.

We first note that element-wise stable distributions are both 0- and 1-saturated.

Lemma 39 (ELM-ST \Rightarrow 0/1-SAT). *Let $\mathcal{F} = \langle \text{Arg}, \text{Att} \rangle$ be an AF and $\mu \in \text{Distr}(\mathcal{F})$. If μ is element-wise stable then μ is 0/1-saturated.*

Proof. Let μ be element-wise stable. This implies that μ is almost-surely conflict-free. By contradiction, assume μ is not 0/1-saturated. Then there is a $B \in \text{Arg}$ such that either $\mu(B) = 0$ but $\mu(\bigvee_{A \in \rightarrow B} A) \neq 1$, or $\mu(B) < 1$ but $\mu(\bigvee_{A \in \rightarrow B} A) \not> 0$. In both cases, there is an assignment $\beta \in \text{Supp}(\mu)$ with $\beta(B) = \text{F}$ and $\beta(A) = \text{F}$ for all attackers $A \in \rightarrow B$. But then $\text{Arg}_\beta \cap \rightarrow B = \emptyset$, which contradicts β being a stable assignment. \square

On the other hand, 0/1-saturated distributions are not necessarily element-wise stable as the following example shows.

Example 40 (0/1-SAT $\not\Leftarrow$ ELM-ST). *As mentioned before any AF \mathcal{F} where the underlying graph constitutes a simple cycle of odd length $n \geq 3$ (for instance the AF from Example 37) does not have stable argument sets and hence no element-wise stable distributions. However, the distribution that assigns probability $1/n$ to each assignment id_A for every $A \in \text{Arg}$ is 0/1-saturated because the likelihood of each argument is then strictly between 0 and 1.*

Nevertheless, the definitions of element-wise stable, 0-saturated, and 1-saturated distributions agree for Dirac distributions. This implies, in particular, that 0-saturation as well as 1-saturation are conservative extensions of the non-probabilistic stable semantics.

Lemma 41 (Conservative extension ST). *If $\mathcal{S} \subseteq \text{Arg}$ is conflict-free then the following statements are equivalent:*

- (a) \mathcal{S} is stable.
- (b) $\text{Dirac}_{\mathcal{S}}$ is element-wise stable.
- (c) $\text{Dirac}_{\mathcal{S}}$ is 0-saturated.
- (d) $\text{Dirac}_{\mathcal{S}}$ is 1-saturated.
- (e) $\text{Dirac}_{\mathcal{S}}$ is 0/1-saturated.

Proof. The equivalence of statements (a) and (b) is trivial. The conditions required for 0-saturation and 1-saturation of the Dirac distribution $\mu = \text{Diracs}_{\mathcal{S}}$ collapse to the following requirement:

$$\text{If } B \notin \mathcal{S} \text{ then } \mathcal{S} \cap \neg B = \emptyset.$$

This yields the equivalence of (a) to statements (c), (d), and (e). \square

3.4.1 EXISTENCE OF 0/1-SATURATED DISTRIBUTIONS

As mentioned above there are AFs that do not have stable argument sets and thus no element-wise stable distributions, so no guarantee can be given for the existence of 0- or 1-saturated Dirac distributions. Nevertheless, 0/1-saturated distributions always exist provided there are no self-attacks in the AF. Specifically, we show how in those cases the grounded argument set can be extended to a 0/1-saturated distribution.

Lemma 42 (Existence of 0/1-saturated distributions). *For every argumentation framework $\mathcal{F} = \langle \text{Arg}, \text{Att} \rangle$ where Att is irreflexive, there is at least one 0/1-saturated distribution.*

Proof. We partition the set Arg of all arguments as follows: $\mathcal{S}_{\text{GR}} \subseteq \text{Arg}$ is the unique grounded argument set, i.e., $[\mathcal{F}]_{\text{GR}} = \{\mathcal{S}_{\text{GR}}\}$, $\mathcal{S}_{\text{GR}}^{\rightarrow}$ is the set of all arguments attacked by the grounded set, and $\mathcal{H} = \text{Arg} \setminus (\mathcal{S}_{\text{GR}} \cup \mathcal{S}_{\text{GR}}^{\rightarrow})$ contains the remaining arguments. Note that if \mathcal{H} is not empty, then it contains at least two arguments: A single argument in \mathcal{H} would either be initial or attacked from within $\mathcal{S}_{\text{GR}}^{\rightarrow}$, but in both cases it would be defended by \mathcal{S}_{GR} . However, the grounded set is complete, so the argument would need to be contained in \mathcal{S}_{GR} . By the same reasoning we get that all arguments in \mathcal{H} are non-initial and have at least one attacker that is also in \mathcal{H} .

The assumption that \mathcal{F} has no self-attacks then yields the existence of a subset \mathcal{C} of $2^{\mathcal{H}}$ such that (1) all argument sets in \mathcal{C} are conflict-free, (2) each argument $B \in \mathcal{H}$ is contained in at least one $\mathcal{S} \in \mathcal{C}$, (3) for each argument $B \in \mathcal{H}$ there is at least one set $\mathcal{S} \in \mathcal{C}$ with $B \notin \mathcal{S}$. For example, the set $\mathcal{C} = \{\emptyset\} \cup \{\{B\} : B \in \mathcal{H}\}$ satisfies (1)–(3).

We now consider an arbitrary distribution μ such that $\text{Supp}(\mu) = \{\text{id}_{\mathcal{S} \cup \mathcal{S}_{\text{GR}}} : \mathcal{S} \in \mathcal{C}\}$. Each assignment $\text{id}_{\mathcal{S} \cup \mathcal{S}_{\text{GR}}}$ is conflict-free as both \mathcal{S} and \mathcal{S}_{GR} are conflict-free, $\mathcal{S} \cap \mathcal{S}_{\text{GR}}^{\rightarrow} = \emptyset$ holds as $\mathcal{S} \subseteq \mathcal{H}$, and $\mathcal{S}^{\rightarrow} \cap \mathcal{S}_{\text{GR}} = \emptyset$ as \mathcal{S}_{GR} is admissible. Thus, μ is almost-surely conflict-free by Lemma 9. We conclude the proof by checking the constraints imposed by 0/1-saturation for each argument $B \in \text{Arg}$:

- $B \in \mathcal{S}_{\text{GR}}$: Then by the construction of μ , we have $\mu(B) = 1$ and are done.
- $B \in \mathcal{S}_{\text{GR}}^{\rightarrow}$: Then $\mu(B) = 0$, but also $\mu(\bigvee_{A \in \neg B} A) = 1$.
- $B \in \mathcal{H}$: Then $0 < \mu(B) < 1$ by (2) and (3). As noted before, B is non-initial and has at least one attacker $A \in \mathcal{H}$, so $\mu(\bigvee_{A \in \neg B} A) > 0$. \square

Example 43 (AFs with self-attacks). *To illustrate why the restriction to AFs without self-attacks can be necessary to ensure the existence of 0/1-saturated distributions, consider an AF containing an argument B with $\neg B = \{B\}$, i.e., B is self-attacking and has no other attackers. Then B is not contained in any conflict-free argument set, so $\mu(B) = 0$ holds for each almost-surely conflict-free distribution μ . Subsequently, $\mu(\bigvee_{A \in \neg B} A) = \mu(B) = 0$*

and thus no 0-saturated distribution exists. For the same reason there is no 1-saturated distribution since $\mu(B) < 1$ holds but the attack probability of B is not positive.

However, the existence of a self-attacking argument is not sufficient to prevent the existence of 0/1-saturated distribution. Consider the following AF:



Then the unique distribution μ with $\mu(A) = 1$ and $\mu(B) = 0$ is almost-surely conflict-free. As $\mu(\bigvee_{A \in \rightarrow B} A) = 1$ and $\mu(\bigvee_{B \in \rightarrow A} B) = 0$, μ is 0/1-saturated. In fact, μ is also element-wise stable because $\{A\}$ is a stable argument set. An example for an AF with a self-loop that has no stable argument set but still a 0/1-saturated distribution is readily constructed by adding a three-node cycle to the example above.

An algorithm to decide whether a 1-saturated distribution exists for a given AF is detailed in Section 5.3.3.

3.4.2 RELATIONSHIPS

We now consider the relationships between the saturation semantics from Definition 38 and the admissible and complete semantics from Sections 3.2 and 3.3. As shown by Example 43, existence of saturated distributions is not guaranteed. This indicates that none of the admissible and complete semantics are subsumed by any of the saturation semantics. In the other direction, we obtain the following lemma.

Lemma 44 (0-SAT \Rightarrow wADM, 0/1-SAT \Rightarrow wCMP). *Let $\mathcal{F} = \langle Arg, Att \rangle$ be an AF and $\mu \in \text{Distr}(\mathcal{F})$ be conflict-free. Then:*

- (a) *If μ is 0-saturated then μ is weakly admissible, i.e., $Arg_\mu \subseteq \text{asDefend}(\mu)$.*
- (b) *If μ is 1-saturated then $\text{asDefend}(\mu) \subseteq Arg_\mu$.*
- (c) *If μ is 0/1-saturated then μ is weakly complete, i.e., $Arg_\mu = \text{asDefend}(\mu)$.*

Proof. To prove (a), we suppose that μ is 0-saturated and pick an argument $C \in Arg_\mu$. The task is to show that μ almost-surely defends C , i.e., the attack probability for each attacker of C is 1. Let $B \in \rightarrow C$. As $C \in Arg_\mu$ we have $\mu(C) = 1$. As μ is conflict-free we obtain $\mu(B) = 0$. By the 0-saturation of μ we obtain:

$$\mu\left(\bigvee_{A \in \rightarrow B} A\right) = 1.$$

Hence, B is attacked with probability 1. This shows $Arg_\mu \subseteq \text{asDefend}(\mu)$.

To prove (b), we suppose that μ is 1-saturated and aim to show $\text{asDefend}(\mu) \subseteq Arg_\mu$. For this, we pick an argument $C \in \text{asDefend}(\mu)$ and show $\mu(C) = 1$. Suppose by contradiction that $\mu(C) < 1$. But then 1-saturation yields

$$\mu\left(\bigvee_{B \in \rightarrow C} B\right) > 0.$$

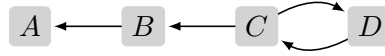
So, there is some attacker B of C with $\mu(B) > 0$. On the other hand, as $C \in asDefend(\mu)$, the attack probability of B under μ equals 1. That is, for each assignment $\beta \in Supp(\mu)$ there exists an argument A with $\beta(A) = \top$ and $(A, B) \in Att$. But then $\beta(B) = \top$ as otherwise β would not be conflict-free. This shows $\beta(B) = \top$ for all $\beta \in Supp(\mu)$. This yields $\mu(\neg B) = 1$ and therefore $\mu(B) = 0$. Contradiction.

Statement (c) is a direct consequence of (a) and (b). □

From Lemma 44 and the fact that weakly complete semantics are also weakly admissible, it is clear that also 0/1-saturated argument sets are weakly admissible. The following examples suffice to show that there are no further inclusion relationships between the saturation, admissible, and complete semantics.

Example 45 (1-SAT $\not\Rightarrow$ WADM). Consider the following almost-surely conflict-free distribution for the AF on the right.

$$\begin{aligned}\mu(A \wedge \neg B \wedge \neg C \wedge \neg D) &= 8/10 \\ \mu(A \wedge \neg B \wedge C \wedge \neg D) &= 1/10 \\ \mu(A \wedge \neg B \wedge \neg C \wedge D) &= 1/10\end{aligned}$$

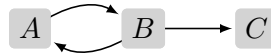


Then μ is 1-saturated as $\mu(A) = 1$, while B , C , and D are attacked with positive probability. But μ is not weakly-admissible as $\mu(A) = 1$ but $\mu(\Delta(A)) = \mu(C) = 1/10 \neq 1$.

Example 46 (0-SAT $\not\Rightarrow$ WCOMP). Consider the AF with only a single argument and no attacks $\mathcal{F} = \langle \{A\}, \{\} \rangle$ and the almost-surely conflict-free distribution μ with support $\mu(A) = 9/10$ and $\mu(\perp) = 1/10$. Then μ is 0-saturated trivially. On the other hand, $\mu(\Delta(A)) = 1$ but $\mu(A) \neq 1$.

Example 47 (0/1-SAT $\not\Rightarrow$ MINADM). Consider the following almost-surely conflict-free distribution for the AF on the right.

$$\begin{aligned}\mu(A \wedge \neg B \wedge \neg C) &= 1/10 \\ \mu(\neg A \wedge B \wedge \neg C) &= 1/10 \\ \mu(\neg A \wedge \neg B \wedge C) &= 8/10\end{aligned}$$



Then μ is 0-saturated trivially and 1-saturated as all arguments are attacked with positive probability. But μ is not min-admissible as $\mu(C) = 8/10 > \min_{X \in \neg \mathcal{C}} \mu(\bigvee_{Y \in \neg X} Y) = \mu(A) = 1/10$.

3.5 Labelings

We established that our probabilistic notions of admissible and complete semantics conservatively extend classical admissibility and complete semantics for *argument sets*. For most classical argumentation semantics, an alternative formulation in terms of *labelings* is available (see Baroni et al. (2011) for an overview). In this section, we investigate ways to get from distributions to labelings, so-called *labeling schemes*, and present a novel approach under which conservative extensions in the argument set view carry over to the labeling-based semantics.

A *labeling* is a function $\mathcal{L}: Arg \rightarrow \{\text{in}, \text{out}, \text{undec}\}$ that assigns to each argument in an AF one of the three statuses in, out, or undec (undecided). A labeling \mathcal{L} is conflict-free iff no attack between any two arguments that are both labeled with in exists. For admissible labelings, the in and out labels need to be justified; complete labelings additionally also require a justification for the undec labels.

Definition 48 (Admissible and complete labelings ADM- \mathcal{L} and CMP- \mathcal{L}). *For an AF $\mathcal{F} = \langle Arg, Att \rangle$, a labeling $\mathcal{L}: Arg \rightarrow \{\text{in}, \text{out}, \text{undec}\}$ is admissible if for each $A \in Arg$ the following two conditions hold:*

1. *If $\mathcal{L}(A) = \text{in}$, then $\mathcal{L}(B) = \text{out}$ for all attackers $B \in \rightarrow A$.*
2. *If $\mathcal{L}(A) = \text{out}$, then $\mathcal{L}(B) = \text{in}$ for at least one attacker $B \in \rightarrow A$.*

\mathcal{L} is complete if it is admissible and the following additional condition holds:

3. *If $\mathcal{L}(A) = \text{undec}$, then $\mathcal{L}(B) \neq \text{in}$ for all attackers $B \in \rightarrow A$ and $\mathcal{L}(B) \neq \text{out}$ for at least one attacker $B \in \rightarrow A$.*

Given a conflict-free argument set \mathcal{S} , a corresponding labeling $\mathcal{L}_{\mathcal{S}}$ can be constructed by assigning the label in to all arguments in \mathcal{S} , out to all arguments attacked by \mathcal{S} , and undec to the remaining arguments in $Arg \setminus (\mathcal{S} \cup \mathcal{S}^{\rightarrow})$:

$$\mathcal{L}_{\mathcal{S}}(A) = \begin{cases} \text{in} & \text{if } A \in \mathcal{S} \\ \text{out} & \text{if } A \in \mathcal{S}^{\rightarrow} \\ \text{undec} & \text{otherwise.} \end{cases}$$

Conversely, given a labeling \mathcal{L} , the corresponding argument set $Arg_{\mathcal{L}}$ is given by all arguments labeled in by \mathcal{L} . The following lemma by (Caminada & Gabbay, 2009) connects the two approaches to classical argumentation for admissibility and complete semantics.

Lemma 49. *Let $\mathcal{F} = \langle Arg, Att \rangle$ be an AF, $\mathcal{S} \subseteq Arg$ a conflict-free argument set, and \mathcal{L} a labeling on \mathcal{F} .*

1. *If \mathcal{S} is admissible, then the labeling $\mathcal{L}_{\mathcal{S}}$ is admissible.*
2. *If \mathcal{L} is admissible, then the set $Arg_{\mathcal{L}}$ is admissible.*
3. *\mathcal{L} is complete iff there is a complete set $\mathcal{S}' \subseteq Arg$ with $\mathcal{L} = \mathcal{L}_{\mathcal{S}'}$.*

The last statement of Lemma 49 implies that for each AF, a bijection between the complete argument sets and the complete labelings exists. This is not the case for admissibility, as several admissible labelings may correspond to the same admissible argument set.

3.5.1 LABELING SCHEMES

In the probabilistic setting, we are interested in labelings induced by distributions. To this end, various *labeling schemes* are imaginable, each specifying how to get from a distribution $\mu \in Distr(\mathcal{F})$ to a labeling $\mathcal{L}_{\mu}: Arg \rightarrow \{\text{in}, \text{out}, \text{undec}\}$. Arguably the most intuitive approach is to choose the label for each argument A based on its likelihood $\mu(A)$. All labeling schemes implementing this approach by a three-way partitioning of the unit interval are instances of the following *interval labeling scheme*.

Definition 50 (Interval labeling scheme). *Let $\mathcal{F} = \langle \text{Arg}, \text{Att} \rangle$ be an AF and $\mu \in \text{Distr}(\mathcal{F})$. Given three disjoint intervals i , o , and u that form a partition of the unit interval, i.e., $i \cup o \cup u = [0, 1]$, the interval labeling $\mathcal{L}_\mu^{i,o,u}$ is defined for each $A \in \text{Arg}$ as*

$$\mathcal{L}_\mu^{i,o,u}(A) = \begin{cases} \text{in} & \text{if } \mu(A) \in i \\ \text{out} & \text{if } \mu(A) \in o \\ \text{undec} & \text{if } \mu(A) \in u. \end{cases}$$

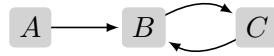
An instance of the interval labeling scheme with $i = (0.5, 1]$, $o = [0, 0.5)$, and $u = \{0.5\}$ can be found in (Hunter & Thimm, 2017), which we will call the *Hunter-Thimm-labeling scheme* \mathcal{L}^{HT} . We will also consider an instance called the *cautious labeling scheme* \mathcal{L}^{cts} where only arguments with likelihood 1 or 0 are, respectively, labeled in or out, and all other arguments are labeled undec. That is, $\mathcal{L}_\mu^{cts} = \mathcal{L}_\mu^{i,o,u}$ with $i = \{1\}$, $o = \{0\}$, and $u = (0, 1)$.

Given the correspondence between argument sets and their induced Dirac distributions, we can check if the labeling for the argument set agrees with the labeling yielded by a labeling scheme for the respective Dirac distribution. If these two labelings coincide for every argument set, we say that the labeling scheme is a conservative extension of argument set induced labelings.

Definition 51 (Conservative extension of labelings). *A labeling scheme \mathcal{L}^x conservatively extends the labelings induced by argument sets iff for all conflict-free sets $\mathcal{S} \in \text{Arg}$:*

$$\mathcal{L}_\mathcal{S} = \mathcal{L}_{\text{Dirac}_\mathcal{S}}^x.$$

Example 52. *Consider the following AF and the admissible argument set $\mathcal{S} = \{A\}$.*



Then the induced labeling $\mathcal{L}_\mathcal{S}$ is given by $\mathcal{L}_\mathcal{S}(A) = \text{in}$, $\mathcal{L}_\mathcal{S}(B) = \text{out}$, and $\mathcal{L}_\mathcal{S}(C) = \text{undec}$. The distribution $\text{Dirac}_\mathcal{S}$ yields the likelihoods $\text{Dirac}_\mathcal{S}(A) = 1$ and $\text{Dirac}_\mathcal{S}(B) = \text{Dirac}_\mathcal{S}(C) = 0$. Now both under the cautious and the Hunter-Thimm-labeling scheme, the labeling induced by $\text{Dirac}_\mathcal{S}$ assigns argument C the label out. Thus, neither $\mathcal{L}_{\text{Dirac}_\mathcal{S}}^{cts}$ nor $\mathcal{L}_{\text{Dirac}_\mathcal{S}}^{HT}$ equals $\mathcal{L}_\mathcal{S}$. In fact, no instance of the interval labeling scheme is a conservative extension: As $\text{Dirac}_\mathcal{S}(B) = \text{Dirac}_\mathcal{S}(C)$, we also have $\mathcal{L}_{\text{Dirac}_\mathcal{S}}^{i,o,u}(B) = \mathcal{L}_{\text{Dirac}_\mathcal{S}}^{i,o,u}(C)$ regardless of the chosen intervals i , o , and u .

Nevertheless, for each argument set \mathcal{S} the labeling $\mathcal{L}_\mathcal{S}$ is contained in the set of all cautious labelings \mathcal{L}_μ^{cts} induced by distributions over the assignments for Arg , under the condition that the given AF has no self-attacks.

Lemma 53. *Let $\mathcal{F} = \langle \text{Arg}, \text{Att} \rangle$ be an AF where the attack relation Att is irreflexive. Then for each conflict-free argument set $\mathcal{S} \subseteq \text{Arg}$ there is a conflict-free distribution $\mu \in \text{Distr}(\mathcal{F})$ such that $\mathcal{L}_\mathcal{S} = \mathcal{L}_\mu^{cts}$.*

Proof. Let $\beta = \text{id}_\mathcal{S}$ be the assignment corresponding to \mathcal{S} and let \mathcal{Y} denote the set of all arguments $C \in \text{Arg} \setminus \mathcal{S}$ that do not have an attacker in \mathcal{S} , i.e.,

$$\mathcal{Y} = \{ C \in \text{Arg} \setminus \mathcal{S} : \mathcal{S} \cap \neg C = \emptyset \}.$$

As $\mathcal{Y} \cap \mathcal{S} = \emptyset$, we have $\beta(C) = \mathbb{F}$ for all $C \in \mathcal{Y}$. For each $C \in \mathcal{Y}$, let β_C denote the assignment that agrees with β for all arguments in $Arg \setminus \{C\}$ but assigns $\beta_C(C) = \mathbb{T}$. Using the assumption that $(C, C) \notin Att$, the assignments β_C are conflict-free. Let now μ be an arbitrary distribution with $Supp(\mu) = \{\beta\} \cup \{\beta_C : C \in \mathcal{Y}\}$. Then each argument in Arg is either in \mathcal{S} , \mathcal{Y} , or $Arg \setminus (\mathcal{S} \cup \mathcal{Y})$:

- For $A \in \mathcal{S}$, we have $\beta(A) = \beta_C(A) = \mathbb{T}$ for all $C \in \mathcal{Y}$. Hence, $\mu(A) = 1$ and $\mathcal{L}_\mu^{cts}(A) = \text{in} = \mathcal{L}_\mathcal{S}(A)$.
- For $C \in \mathcal{Y}$, we have $\mu(C) = \mu(\beta_C)$. By $\beta_C \in Supp(\mu)$, we know $\mu(\beta_C) > 0$, and further $\mu(\beta_C) < 1$ as there are other assignments in $Supp(\mu)$. Hence, $\mathcal{L}_\mu^{cts}(C) = \text{undec} = \mathcal{L}_\mathcal{S}(C)$.
- For $B \in Arg \setminus (\mathcal{S} \cup \mathcal{Y})$, we have $\beta(B) = \beta_C(B) = \mathbb{F}$ for all $C \in \mathcal{Y}$. Hence, $\mu(B) = 0$ and $\mathcal{L}_\mu^{cts}(B) = \text{out} = \mathcal{L}_\mathcal{S}(B)$.

This shows $\mathcal{L}_\mu^{cts} = \mathcal{L}_\mathcal{S}$. From Lemma 9 we obtain that μ is also conflict-free. □

Lemma 53 excludes the case of self-attacks. Indeed, the statement of Lemma 53 does not hold, e.g., for the empty argument set $\mathcal{S} = \emptyset$ in an argumentation framework where all arguments are self-attacking. In this case, the only conflict-free distribution is the Dirac distribution μ assigning 1 to the assignment $A \mapsto \mathbb{F}$ for all $A \in Arg$. But then \mathcal{L}_μ^{cts} labels all arguments out. On the other hand, the labeling induced by $\mathcal{S} = \emptyset$ labels all arguments undec. The same phenomenon would occur for the Hunter-Thimm-labeling scheme.

3.5.2 CONSERVATIVE LABELING SCHEME

We now introduce a labeling scheme that is a conservative extension of labelings induced by argument sets.

Definition 54 (Conservative labeling scheme). *Let $\mathcal{F} = \langle Arg, Att \rangle$ be an AF and $\mu \in Distr(\mathcal{F})$. The conservative labeling \mathcal{L}_μ^{cons} is defined as follows for each argument $A \in Arg$:*

$$\mathcal{L}_\mu^{cons}(A) = \begin{cases} \text{in} & \text{if } \mu(A) = 1 \\ \text{out} & \text{if } \mu(\bigvee_{B \in \neg A} B) = 1 \\ \text{undec} & \text{otherwise.} \end{cases}$$

Lemma 55. *For each conflict-free argument set $\mathcal{S} \subseteq Arg$, the labeling $\mathcal{L}_\mathcal{S}$ agrees with the conservative labeling of the induced Dirac distribution. That is, $\mathcal{L}_\mathcal{S} = \mathcal{L}_{Dirac_\mathcal{S}}^{cons}$.*

Proof. For any argument $A \in Arg$, we have

$$\mathcal{L}_\mathcal{S}(A) = \text{in} \quad \text{iff} \quad A \in \mathcal{S} \quad \text{iff} \quad Dirac_\mathcal{S}(A) = 1 \quad \text{iff} \quad \mathcal{L}_{Dirac_\mathcal{S}}^{cons}(A) = \text{in}.$$

For the out label, we have

$$\mathcal{L}_\mathcal{S}(A) = \text{out} \quad \text{iff} \quad \neg A \cap \mathcal{S} \neq \emptyset \quad \text{iff} \quad Dirac_\mathcal{S}(\bigvee_{B \in \neg A} B) = 1 \quad \text{iff} \quad \mathcal{L}_{Dirac_\mathcal{S}}^{cons}(A) = \text{out}.$$

This yields the claim. □

We can observe that the condition for the out label, $\mu(\bigvee_{B \in \rightarrow A} B) = 1$, implies $\mu(A) = 0$ for each conflict-free distribution μ . This yields the following connection between the cautious and the conservative labeling scheme:

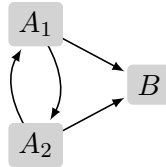
$$\begin{aligned} \mathcal{L}_\mu^{cts}(A) = \text{in} &\quad \Leftrightarrow \quad \mathcal{L}_\mu^{cons}(A) = \text{in} \\ \mathcal{L}_\mu^{cts}(A) = \text{out} &\quad \Leftarrow \quad \mathcal{L}_\mu^{cons}(A) = \text{out} \\ \mathcal{L}_\mu^{cts}(A) = \text{undec} &\quad \Rightarrow \quad \mathcal{L}_\mu^{cons}(A) = \text{undec} \end{aligned}$$

From this relation it is easy to see that for any distribution μ where $\mu(A) = 0$ implies $\mu(\bigvee_{B \in \rightarrow A} B) = 1$, both labeling schemes induce the same labeling. This exact condition is enforced by the notion of 0-saturation from Definition 38. Hence, we obtain:

Lemma 56. *If μ is a 0-saturated distribution then $\mathcal{L}_\mu^{cts} = \mathcal{L}_\mu^{cons}$.*

We conclude the excursus on labelings with an example from Thimm (2012) that highlights how probabilistic semantics may induce labelings that do not appear in the corresponding classical semantics but are still sensible.

Example 57. *Consider the following AF and the two admissible argument sets $\mathcal{S}_1 = \{A_1\}$ and $\mathcal{S}_2 = \{A_2\}$:*



Both induce a labeling function $\mathcal{L}_{\mathcal{S}_i}$ where B is labeled out as B 's attacker A_i is labeled in with respect to \mathcal{S}_i . The distribution μ assigning probability $\frac{1}{2}$ to the corresponding assignments $\text{id}_{\mathcal{S}_1}$ and $\text{id}_{\mathcal{S}_2}$ is element-wise admissible, and the arguments' likelihoods are $\mu(A_1) = \mu(A_2) = \frac{1}{2}$ and $\mu(B) = 0$. Further, we have $\mu(A_1 \vee A_2) = 1$, so μ is 0-saturated and by Lemma 56 the labeling \mathcal{L}_μ^{cts} thus coincides with the labeling \mathcal{L}_μ^{cons} : B is labeled out whereas A_1 and A_2 are labeled undec. This labeling is not admissible as no attacker of B is labeled in, and likewise no argument set exists which could induce it.

Nevertheless, we argue that the classification of B as out is rather natural given the probabilistic background: B is attacked with probability 1, and μ describes a situation where either A_1 or A_2 is accepted. Both are equally likely, but regardless of the outcome B is attacked in either case. This is reflected in the labeling which is undecided about A_1 and A_2 but certain in its judgment of B as out.

4. Taxonomy and Related Work

Most closely related to our approach is the work by Hunter and Thimm (2017). Based on earlier works (Thimm, 2012; Hunter, 2013), the authors extend AFs towards a probabilistic setting by attributing a degree of belief to arguments. Inspired by notions from classical argumentation, they introduce several probabilistic semantics featuring constraints on the likelihood of single arguments that can be compared to our semantics. Adapted to our notation, we recall these semantics below for the sake of self-containedness.

Definition 58. Let $\mathcal{F} = \langle \text{Arg}, \text{Att} \rangle$ be an AF and $\mu \in \text{Distr}(\mathcal{F})$. Then μ is called

- (FOU) founded iff $\mu(A) = 1$ for all initial $A \in \text{Arg}$,
- (sFOU) semi-founded iff $\mu(A) \geq 1/2$ for all initial $A \in \text{Arg}$,
- (OPT) optimistic iff $\mu(\neg A) \leq \sum_{B \in \neg A} \mu(B)$ holds for all $A \in \text{Arg}$,
- (sOPT) semi-optimistic iff (OPT) holds for all non-initial $A \in \text{Arg}$,
- (COH) coherent iff $\mu(A) \leq \mu(\neg B)$ for all $(A, B) \in \text{Att}$,
- (INV) involutory iff $\mu(A) = \mu(\neg B)$ for all $(A, B) \in \text{Att}$,
- (JUS) justifiable iff μ is coherent and optimistic,
- (RAT) rational iff $\mu(A) > 1/2$ implies $\mu(B) \leq 1/2$ for all $(A, B) \in \text{Att}$,
- (MIN) minimal iff $\mu(A) = 0$ for all $A \in \text{Arg}$,
- (NEU) neutral iff $\mu(A) = 1/2$ for all $A \in \text{Arg}$, and
- (MAX) maximal iff $\mu(A) = 1$ for all $A \in \text{Arg}$.

Figure 6 gives an overview of all semantics introduced in this paper in perspective to the semantic notions from (Hunter & Thimm, 2017) and the element-wise lifted classical semantics by Thimm et al. (2017) as given in Definition 6. An arrow from one semantics to another, e.g., $\text{wCMP} \Rightarrow \text{JUS}$, indicates that $\llbracket \mathcal{F} \rrbracket_{\text{wCMP}} \subseteq \llbracket \mathcal{F} \rrbracket_{\text{JUS}}$ holds for arbitrary AFs \mathcal{F} . In this context, BOT denotes the empty semantics and TOP the universal semantics, with $\llbracket \mathcal{F} \rrbracket_{\text{BOT}} = \emptyset$ and $\llbracket \mathcal{F} \rrbracket_{\text{TOP}} = \text{Distr}(\mathcal{F})$. Further, there is at least one AF for each arrow such that the set inclusion is strict, and no other arrows (except for the transitive closure) exist. Proofs for $\text{ELM-ST} \Rightarrow \text{JUS}$ and $\text{ELM-CF} \Rightarrow \text{COH}$ (and thus $\text{ASCF} \Rightarrow \text{COH}$ by Lemma 9) are given in (Thimm et al., 2017).

All probabilistic semantics introduced in this paper entail or require that the distributions are almost-surely conflict-free, so they all imply coherency. However, apart from ELM-ST , no other semantics entails JUS . We show this by the following examples. First, Example 59 covers the admissible and complete semantics (which all subsume ELM-GR and ELM-PRF), followed by Example 60 considering the saturation semantics.

Example 59 ($\text{ELM-GR} \not\Rightarrow \text{sOPT}$, $\text{ELM-PRF} \not\Rightarrow \text{sOPT}$). Consider again the odd cycle AF \mathcal{F} appearing in Example 37 and the assignment $\beta = \{A=F, B=F, C=F\}$. The distribution Dirac_β is element-wise preferred and grounded as the corresponding empty argument set is the only element in both $\llbracket \mathcal{F} \rrbracket_{\text{ST}}$ and $\llbracket \mathcal{F} \rrbracket_{\text{GR}}$. However, Dirac_β is not optimistic (and thus not justifiable) as, e.g., $\text{Dirac}_\beta(\neg A) = 1 \not\leq 0 = \text{Dirac}_\beta(C)$.

Example 60 ($0/1\text{-SAT} \not\Rightarrow \text{sOPT}$). Recall from Example 40 that the distribution μ assigning probability $1/3$ to the assignments $\text{id}_{\{A\}}$, $\text{id}_{\{B\}}$, and $\text{id}_{\{C\}}$ for the odd cycle AF in Example 37 is $0/1$ -saturated. But μ does not satisfy sOPT : for instance, $\mu(\neg A) = 2/3 \not\leq 1/3 = \mu(C)$.

In turn, all our notions of complete semantics are founded: They all imply weak completeness, and initial arguments have maximal likelihood in weakly complete distributions since they are always almost-surely defended.

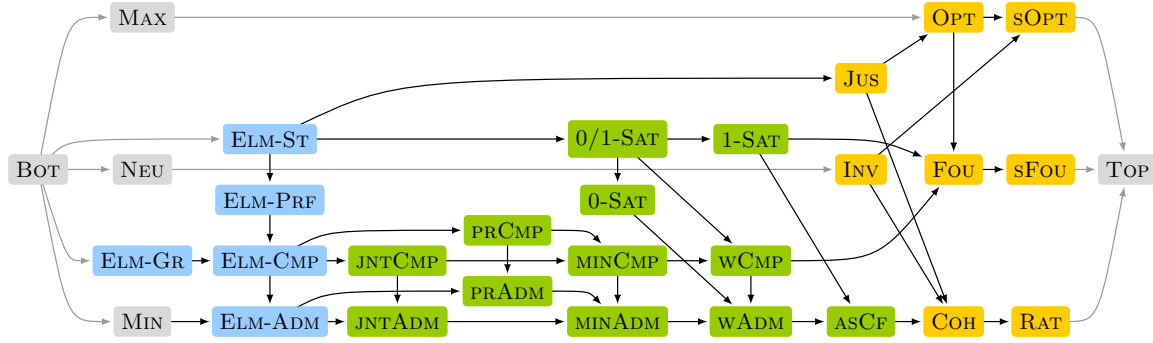


Figure 6: Hierarchy of probabilistic argumentation semantics. Light gray boxes indicate trivial semantics, the boxes on the left-hand side stand for element-wise lifted classical semantics, the boxes on the right refer to the probabilistic semantics introduced by Hunter and Thimm (2017), and the boxes in the middle to the new notions of admissibility (Section 3.2), complete semantics (Definition 29), and saturation (Definition 38).

Lemma 61 ($\text{wCMP} \Rightarrow \text{FOU}$). *If μ is weakly complete, then μ is founded.*

Proof. Let $A \in \text{Arg}$ be initial. Due to $\neg A = \emptyset$, A is trivially almost-surely defended by μ , i.e., $A \in \text{asDefend}(\mu)$. By wCMP , $\text{Arg}_\mu = \text{asDefend}(\mu)$, so $A \in \text{Arg}_\mu$ holds and thus $\mu(A) = 1$. \square

Since any 0/1-saturated distribution is weakly complete, Lemma 61 also implies that 0/1-saturated distributions are founded. The latter holds already for 1-saturated distributions:

Lemma 62 ($1\text{-SAT} \Rightarrow \text{FOU}$). *If μ is 1-saturated, then μ is founded.*

Proof. Assume that μ is 1-saturated but not founded. This means there is some initial A for which $\mu(A) < 1$. But then since μ is 1-saturated, A must be attacked with positive probability, which cannot be the case as A is initial. \square

On the other hand, 0-SAT does not imply FOU as the next example shows.

Example 63 ($0\text{-SAT} \not\Rightarrow \text{sFOU}$). *Consider the AF with only one argument A and no attacks and the distribution μ with $\mu(A) = 2/5$ and $\mu(\perp) = 3/5$. Then μ is 0-saturated trivially but μ is not semi-founded (and, hence, also not founded) as $\mu(A) < 1/2$.*

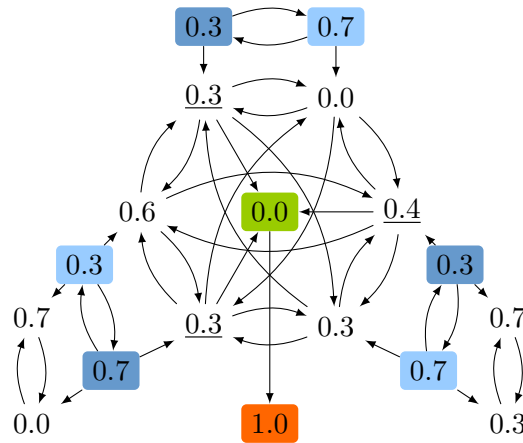
In the setting of Hunter and Thimm (2017), COH and JUS are generalizations of conflict-free argument sets and the complete semantics, respectively. By our definition (cf. Definition 5), coherence is a conservative extension of CF, though this is not the case for JUS and CMP: The assignment β from Example 59 is complete but Dirac_β is not justifiable.

Example 64. *To illustrate some of the differences between our semantics and, in particular, justifiability semantics, we return to our motivating example from the introduction where we*

considered an AF for a semi-autonomous car. Consider the following distribution μ given below as the probabilities of all events in its support, where $\mu(\mathcal{S})$ abbreviates $\mu(\text{id}_{\mathcal{S}})$.

$$\begin{aligned} \mu(\{\text{st}, \bar{\text{cl}}, \bar{\text{cl}} \Rightarrow \bar{\text{l}}, \bar{\text{cl}} \Rightarrow \bar{\text{m}}, \bar{\text{cr}}, \bar{\text{cr}} \Rightarrow \bar{\text{m}}, \bar{\text{cr}} \Rightarrow \bar{\text{r}}, \bar{\text{ld}}\}) &= 0.2 \\ \mu(\{\text{st}, \text{ld}, \bar{\text{cl}}, \bar{\text{cl}} \Rightarrow \bar{\text{l}}, \bar{\text{cl}} \Rightarrow \bar{\text{m}}, \bar{\text{cr}}, \bar{\text{cr}} \Rightarrow \bar{\text{m}}, \bar{\text{cr}} \Rightarrow \bar{\text{r}}\}) &= 0.1 \\ \mu(\{\text{st}, \text{cr}, \text{cr} \Rightarrow \text{r}, \text{ld}, \bar{\text{cl}}, \bar{\text{cl}} \Rightarrow \bar{\text{l}}, \bar{\text{cl}} \Rightarrow \bar{\text{m}}\}) &= 0.3 \\ \mu(\{\text{st}, \text{cr}, \text{cr} \Rightarrow \text{m}, \text{cr} \Rightarrow \text{r}, \text{ld}, \text{ld} \Rightarrow \text{m}, \bar{\text{cl}}, \bar{\text{cl}} \Rightarrow \bar{\text{l}}\}) &= 0.1 \\ \mu(\{\text{st}, \text{cl}, \text{cl} \Rightarrow \text{m}, \text{cr}, \text{cr} \Rightarrow \text{m}, \text{cr} \Rightarrow \text{r}, \bar{\text{ld}}\}) &= 0.1 \\ \mu(\{\text{st}, \text{cl}, \text{cl} \Rightarrow \text{m}, \text{cr}, \text{cr} \Rightarrow \text{m}, \text{cr} \Rightarrow \text{r}, \text{ld}, \text{ld} \Rightarrow \text{m}\}) &= 0.2 \end{aligned}$$

Based on Figure 2, we depict the resulting μ -likelihood of each argument in the figure below:



Though not visible when looking only at the likelihoods, argument st (bottom center, red) with $\mu(\text{st}) = 1$ is only defended by the three (underlined) attackers of ct (center, green) with a total probability of

$$\mu(\text{cl} \Rightarrow \text{m} \vee \text{ld} \Rightarrow \text{m} \vee \text{cr} \Rightarrow \text{m}) = 0.4.$$

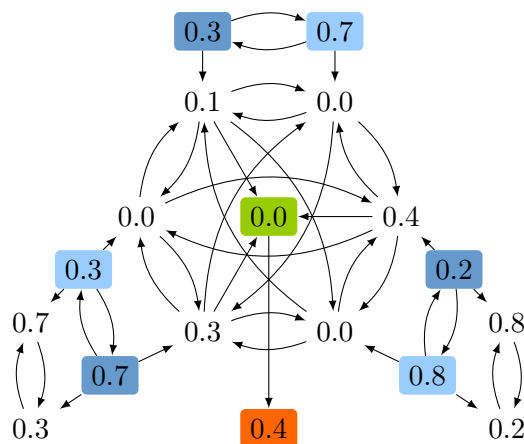
That is, st is not almost-surely defended. However, this is required by weak admissibility (and thus by all of our notions of admissibility and complete semantics) for arguments like st that are almost-surely accepted.

Justifiability, instead, considers solely the likelihoods of the immediate attackers. For the arguments in question, the optimism constraints for argument st and argument ct hold:

$$\begin{aligned} \mu(\neg \text{st}) = 0 &\leq 0 = \mu(\text{ct}) \\ \mu(\neg \text{ct}) = 1 &\leq 0.3 + 0.3 + 0.4 \\ &= \mu(\text{cl} \Rightarrow \text{m}) + \mu(\text{ld} \Rightarrow \text{m}) + \mu(\text{cr} \Rightarrow \text{m}) \end{aligned}$$

In fact, OPT holds for all arguments, so μ is justifiable as coherency is given as well.

Example 65. In the figure below, we show the argument likelihoods induced by a distribution μ that is joint-attack complete but not justifiable.



We see μ is not justifiable as, e.g., for argument st , the OPT constraint is violated: $\mu(\neg st) = 0.6 \not\leq 0 = \mu(ct)$. Note that JNTCMP cannot be verified from the likelihoods alone and requires reasoning on the full joint distribution on arguments. We refer to Example 78 for details on how μ was constructed.

The Bigger Picture. Apart from the technical relations between the probabilistic argumentation semantics shown in Figure 6, there is a foundational difference between the semantical notions by Hunter and Thimm (2017) and the notions we propose. The former only impose constraints on the likelihood of single arguments and therefore tend to be more coarse than our semantics from Section 3. The latter crucially exploit the possibility to impose constraints across the *joint probability distributions*, and this makes it possible to express dependencies as needed already to spell out admissibility in an adequate manner.

Specifically, when restricting oneself to constraints over the likelihood of single arguments, dependencies between the truth values of arguments cannot be expressed: Whenever a system of constraints for $\mu(A)$ can be satisfied by at least one distribution, then the solution space contains at least (i) one distribution μ where the arguments are pairwise independent and (ii) one distribution μ such that for all arguments A, B where $\mu(A) \geq \mu(B)$, it is implied that $B \rightarrow A$ holds for every assignment in the support of μ . If B attacks A and the likelihood of B is positive then both (i) and (ii) are in contrast to standard argumentation semantics. In particular, linear constraints for $\mu(A)$ cannot express that arguments are complements of each other and thus mutually exclusive, which is however required to faithfully model our vehicle example. This observation should not be read as a critique at the earlier work by Hunter and Thimm as their focus is on modeling the belief of arguments and the induced three-values labelings, rather than a conservative extension of standard concepts (such as conflict-freeness or admissibility) on the level of distributions.

4.1 Further Related Work

There is a large body of work on probabilistic extensions of AFs. In general, one distinguishes between the *constellations* approach (Dung & Thang, 2010; Li et al., 2011; Hunter, 2012; Fazzinga, Flesca, & Parisi, 2015) where uncertainty pertains the topology of the framework, and the *epistemic* approach (Hunter & Thimm, 2017; Potyka, 2019) where the

framework is fixed and uncertainty revolves around the acceptance of arguments. This paper falls into the latter category.

Baroni, Giacomin, and Vicig (2014) approach epistemic probabilities from the angle of de Finetti’s theory of subjective probabilities (de Finetti, 1974). They consider rationality conditions based on the notions of defense and reinstatement, which are closely related to admissibility and complete semantics.

An investigation of variants of semantics giving uniform distributions over the complete, preferred, and semi-stable *labelings* of an AF is given by Rienstra, Thimm, Liao, and van der Torre (2018). They show that the schemes investigated produce semantics which are founded, rational, and coherent (see Definition 58). The authors introduce new principles for probabilistic semantics based on SCC-decomposability and SCC-factorability.

Hunter, Polberg, and Thimm (2020) extended the approach by Hunter and Thimm (2017) and proposed *epistemic graphs*. Besides the notion of *support*³ of arguments, a notion that complements attacks (Boella, Gabbay, van der Torre, & Villata, 2010), they augment argument graphs by constraints that restrict the degree of beliefs in arguments as well as how these beliefs influence each other. These constraints can be formulated as Boolean combinations of polynomial inequalities over terms that stand for the probability of acceptance of argument sets (represented with propositional atoms). The semantics considered by Hunter et al. (2020) associates sets of distributions over the powerset of arguments with each epistemic graph. The authors consider three forms of semantics or types of constraints on the distributions, the simplest being the satisfaction semantics, which simply returns the distributions consistent with the constraints of an argument graph. A central impulse for the development of epistemic graphs appears to be work on using argumentation for persuasion (Hadoux, Hunter, & Polberg, 2021), which was already in part also addressed in Hunter et al. (2020). Epistemic graphs are a very general framework. As a matter of fact, all constraints appearing in our work can be cast into the setting of epistemic graphs and its satisfaction semantics.

5. Reasoning Problems and Complexity

Classically, a number of different reasoning problems are of interest given an AF and a specific argumentation semantics. Foremost, one might ask whether there exists an argument set satisfying the semantics’ constraints. Further, one is surely interested in whether a particular argument set satisfies the constraints of a semantics. Finally, checking *credulous* and *skeptical acceptance* of single arguments is of interest, that is, deciding whether a given argument is contained in at least one argument set or, respectively, all argument sets that satisfy the semantics’ constraints. A comprehensive overview is given by Dvořák and Dunne (2018). Based on these tasks, we define five reasoning problems for probabilistic argumentation semantics.

Definition 66 (Reasoning problems). *Let ρ be any probabilistic argumentation semantics. We define the following reasoning problems:*

(NE- ρ) Non-emptiness: *Given an AF \mathcal{F} , decide whether $[[\mathcal{F}]]_\rho$ is non-empty.*

3. Not to be confused with the support of a distribution as defined in Section 2.3.

- (NTNE- ρ) Non-trivial non-emptiness: *Given an AF \mathcal{F} , decide whether $\llbracket \mathcal{F} \rrbracket_\rho \setminus \{\text{Dirac}_\emptyset\}$ is non-empty.*
- (CA- ρ) Credulous acceptance: *Given an AF $\mathcal{F} = \langle \text{Arg}, \text{Att} \rangle$, an argument $A \in \text{Arg}$, and a threshold $\theta \in \mathbb{Q} \cap (0, 1]$, decide if there is at least one $\mu \in \llbracket \mathcal{F} \rrbracket_\rho$ with $\mu(A) \geq \theta$.*
- (SA- ρ) Skeptical acceptance: *Given an AF $\mathcal{F} = \langle \text{Arg}, \text{Att} \rangle$, an argument $A \in \text{Arg}$, and a threshold $\theta \in \mathbb{Q} \cap (0, 1]$, decide if $\mu(A) \geq \theta$ holds for all $\mu \in \llbracket \mathcal{F} \rrbracket_\rho$.*
- (M- ρ) Membership: *Given an AF \mathcal{F} and a distribution $\nu \in \text{Distr}(\mathcal{F})$ as input, decide whether $\nu \in \llbracket \mathcal{F} \rrbracket_\rho$.*

When considering the complexity of these reasoning problems, we assume a simple graph encoding for the input AF and a binary encoding of the rational threshold θ . The input distribution for the membership problem is encoded explicitly as vector of rational values in $\mathbb{Q} \cap [0, 1]$, one for each assignment.

For the first three decision problems, i.e., both variants of non-emptiness and credulous acceptance, a positive answer would usually be accompanied by a witness distribution. For skeptical acceptance, in case of a negative answer one might be interested in a distribution acting as a counterexample.

In the following sections we investigate how the reasoning problems provided in Definition 66 can be algorithmically solved and investigate their complexities when instantiated with the probabilistic semantics introduced in this paper. An overview of all results is given in Table 1.

5.1 Generic Upper Bounds

We show how each of the reasoning problems from Definition 66 can be computed in a generic way for all semantics introduced in this paper. This establishes decidability for each of the decision problems and yields a general upper bound on their time complexity.

With the exception of their weak variants, the introduced admissible and complete semantics share the characteristics that each of them imposes a certain set of linear constraints on the joint probability distributions over assignments. The non-emptiness problem NE- ρ is thus solvable in exponential time since the feasibility of linear constraint systems can be checked in polynomial time and the number of variables grows exponentially in the number of arguments. For weak admissibility, weak complete semantics, and the saturation semantics, the imposed constraints are non-linear as they contain implications that condition the likelihood of arguments. However, an exponential time bound for the existence problem under these semantics can still be obtained by enumerating all subsets of *Arg* as candidates for the arguments which are accepted with the respective likelihood and checking the feasibility of the resulting—then linear—constraint system for each of them.

NE- ρ is trivially solved if ρ is any semantics implied by MIN in Figure 6, as MIN always contains the distribution $\mu_0 = \text{Dirac}_\emptyset$ with $\mu_0(A) = 0$ for all arguments. This motivated the formulation of the non-trivial non-emptiness problem NTNE- ρ where μ_0 is excluded. The switch from NE- ρ to NTNE- ρ can be achieved by adding a single linear constraint, such that the exponential time bound remains the same for all semantics. The same holds for the credulous acceptance problem CA- ρ which can be phrased as existence problem

semantics ρ	NE- ρ	NTNE- ρ	CA- ρ	SA- ρ	M- ρ
ELM-CF	trivial	in P	in P	trivial	in P
ELM-GR	trivial	in P	P -c	P -c	P -c
WADM	trivial	in P	NP -h	trivial	in P
MINADM	trivial		NP -h	trivial	in P
JNTADM	trivial		NP -h	trivial	in P
PRADM	trivial		NP -h	trivial	in P
ELM-ADM	trivial	NP -c	NP -c	trivial	in P
WCMP	trivial	in P	NP -h	P -c	in P
MINCMP	trivial		NP -h	P -c	in P
JNTCMP	trivial		NP -h	P -c	in P
PRCMP	trivial		NP -h	P -c	in P
ELM-CMP	trivial	NP -c	NP -c	P -c	in P
0-SAT			NP -h		in P
1-SAT	in P	in P	in P	in P	in P
0/1-SAT			NP -h		in P
ELM-ST	NP -c	NP -c	NP -c	$coNP$ -c	in P

Table 1: Summary of complexity results for the reasoning problems from Definition 66. *EXPTIME* is an implicit upper bound for all empty fields and entries where only a hardness result is given (see Section 5.1), entries for the element-wise lifted semantics are discussed in Section 5.2, results for NE- ρ and NTNE-WADM and NTNE-WCMP in Section 5.3.1, and skeptical acceptance SA- ρ in Section 5.3.2. 1-SAT is covered in Section 5.3.3 and the hardness results for CA- ρ are given in Section 5.4. Note that for the membership problem M- ρ , the polynomial time bounds are subject to a trivial full encoding of the input distribution.

with an additional linear constraint encoding $\mu(A) \geq \theta$ for the given argument A and threshold θ . For the skeptical acceptance problem SA- ρ , we can check for the existence of a counterexample by looking for a distribution which satisfies the constraints of ρ but also the linear constraint encoding $\mu(A) < \theta$.

For the membership problem M- ρ , the constraints enforced by a semantics ρ for a given AF have to be evaluated under the given distribution. The number of constraints is linear in the number of arguments in the AF, though the size of each constraint can be exponentially large. However, given the trivial encoding of the input distribution as a full list of entries for each possible assignment, the distribution is itself exponentially large compared to the AF. In that case, taking into account the whole input size, evaluating the constraints is possible in polynomial time. The complexity of M- ρ is left open in case the distribution is given by some more efficient encoding, e.g., when the zero entries for the assignments that are not in the support of the distribution are omitted.

In summary, *EXPTIME* is a general upper bound on the time complexity for all reasoning problems. Upper and lower bounds can however be sharpened for many problem instances, as we show in the following sections.

5.2 Classical Semantics Instances

Decision problems for element-wise lifting semantics inherit their complexity from the corresponding decision problems in the classical setting.

Lemma 67 (Reasoning problems under element-wise lifting). *Let $\mathcal{F} = \langle \text{Arg}, \text{Att} \rangle$ be an AF and σ a classical argumentation semantics. Then the following equivalences hold for the respective reasoning problem:*

- (a) *Non-emptiness: $[\mathcal{F}]_\sigma = \emptyset$ iff $\llbracket \mathcal{F} \rrbracket_{\text{ELM-}\sigma} = \emptyset$.*
- (b) *Non-trivial non-emptiness: $[\mathcal{F}]_\sigma \setminus \{\emptyset\} = \emptyset$ iff $\llbracket \mathcal{F} \rrbracket_{\text{ELM-}\sigma} \setminus \{\text{Dirac}_\emptyset\} = \emptyset$.*
- (c) *Credulous acceptance: For any argument $A \in \text{Arg}$, there is a set $\mathcal{S} \in [\mathcal{F}]_\sigma$ with $A \in \mathcal{S}$ iff there is a distribution $\mu \in \llbracket \mathcal{F} \rrbracket_{\text{ELM-}\sigma}$ and a threshold $\theta \in (0, 1]$ s.t. $\mu(A) \geq \theta$.*
- (d) *Skeptical acceptance: For any $A \in \text{Arg}$, A is contained in each $\mathcal{S} \in [\mathcal{F}]_\sigma$ iff there is threshold $\theta \in (0, 1]$ s.t. $\mu(A) \geq \theta$ holds for all $\mu \in \llbracket \mathcal{F} \rrbracket_{\text{ELM-}\sigma}$.*

Proof. (a) and (b): The existence of an argument set $\mathcal{S} \in [\mathcal{F}]_\sigma$ implies $\text{Dirac}_\mathcal{S} \in \llbracket \mathcal{F} \rrbracket_{\text{ELM-}\sigma}$, and the existence of a distribution $\mu \in \llbracket \mathcal{F} \rrbracket_{\text{ELM-}\sigma}$ implies there is at least one assignment $\beta \in \text{Supp}(\mu)$ such that $\text{Arg}_\beta \in [\mathcal{F}]_\sigma$.

(c): If argument A is contained in at least one argument set $\mathcal{S} \in [\mathcal{F}]_\sigma$, then $\text{Dirac}_\mathcal{S} \in \llbracket \mathcal{F} \rrbracket_{\text{ELM-}\sigma}$ and $\text{Dirac}_\mathcal{S}(A) = 1 \geq \theta$ for ever $\theta \in (0, 1]$. Vice versa, if $\mu(A) \geq \theta$ for any $\theta > 0$ and a distribution $\mu \in \llbracket \mathcal{F} \rrbracket_{\text{ELM-}\sigma}$, then there is at least one $\beta \in \text{Supp}(\mu)$ with $\beta(A) = \top$. Thus, $A \in \text{Arg}_\beta$ and $\text{Arg}_\beta \in [\mathcal{F}]_\sigma$.

(d): If A is in each set $\mathcal{S} \in [\mathcal{F}]_\sigma$, then $\beta(A) = \top$ for each assignment $\beta \in \text{Supp}(\mu)$ where $\mu \in \llbracket \mathcal{F} \rrbracket_{\text{ELM-}\sigma}$ and thus $\mu(A) = 1$. For the reverse direction, note that for every element-wise lifted semantics σ there is a distribution $\hat{\mu} \in \llbracket \mathcal{F} \rrbracket_{\text{ELM-}\sigma}$ where $\text{Supp}(\hat{\mu}) = \{\text{id}_\mathcal{S} : \mathcal{S} \in [\mathcal{F}]_\sigma\}$. Now if $\mu(A) \geq \theta$ holds for all $\mu \in \llbracket \mathcal{F} \rrbracket_{\text{ELM-}\sigma}$, we also have $\hat{\mu}(A) \geq \theta$ and thus $A \in \mathcal{S}$ for all $\mathcal{S} \in [\mathcal{F}]_\sigma$. \square

Using complexity results from classical abstract argumentation (see Dvořák and Dunne (2018) for a summary), the non-trivial non-emptiness problem and the credulous acceptance problem are *NP*-complete (notated as *NP-c*) for the element-wise lifted versions of admissible, complete, and stable semantics. Further, for ELM-ST also the regular non-emptiness problem is *NP*-complete and the skeptical acceptance problem is *coNP*-complete.

5.3 Tractable Instances

Apart from the membership problem and the non-emptiness problem for the various admissibility semantics, several other decision problems can be solved in polynomial time.

5.3.1 NON-TRIVIAL NON-EMPTINESS

Every argumentation framework has a uniquely defined grounded argument set \mathcal{S}_{GR} (see Definition 2) which is also a complete set and can be computed in polynomial time. As each of our probabilistic variants of complete semantics conservatively extends classical complete semantics, they always contain the distribution $\text{Dirac}_{\mathcal{S}_{\text{GR}}}$, thus trivializing the non-emptiness problem. However, as the grounded argument set can be the empty set, this does not effect the non-trivial non-emptiness problem NTNE- ρ .

For the weakly admissible semantics, the constraint $\text{Arg}_{\mu} \subseteq \text{asDefend}(\mu)$ is trivially satisfied for all distributions μ where $\mu(A) < 1$ holds for each argument A . At first glance this fact might suggest that hence also a non-trivial distribution always exists where $\mu(A) > 0$ for at least one argument. However, special care is needed in case there are self-attacking arguments. As every weakly admissible distribution needs to be almost-surely conflict-free, we get $\mu(A) = 0$ for each self-attacking argument A . Thus, if all arguments of an AF are self-attacking (which can be checked in polynomial time) then no non-trivial weakly distribution exists. Otherwise, let $\text{Arg}_{\not\subseteq} \subseteq \text{Arg}$ denote the set of non-self-attacking arguments and consider the distribution $\tilde{\mu}$ with

$$\tilde{\mu}(\text{id}_{\emptyset}) = \frac{1}{|\text{Arg}_{\not\subseteq}| + 1} \quad \text{and} \quad \tilde{\mu}(\text{id}_{\{A\}}) = \frac{1}{|\text{Arg}_{\not\subseteq}| + 1} \quad \text{for each } A \in \text{Arg}_{\not\subseteq}.$$

Then $\tilde{\mu}$ is almost-surely conflict-free, weakly admissible, and different to the trivial distribution Dirac_{\emptyset} . Thus, the problem NTNE-WADM lies in P .

Similarly, a polynomially time-bounded algorithm to compute a non-trivial weakly complete distribution can be obtained. Again, we first check whether all arguments are self-attacking. If so, no non-trivial weakly complete distribution exists. Otherwise, we proceed to compute the unique grounded argument set \mathcal{S}_{GR} . If it is non-empty, we are finished and return the distribution $\text{Dirac}_{\mathcal{S}_{\text{GR}}}$. Otherwise, when $\mathcal{S}_{\text{GR}} = \emptyset$, we know that all arguments in the AF must be attacked (as unattacked arguments necessarily appear in the grounded set). In those cases, the distribution $\tilde{\mu}$ is weakly complete: $\tilde{\mu}(\Delta(C)) < 1$ holds for all non-initial arguments C by the construction of $\tilde{\mu}$ as $\tilde{\mu}(\text{id}_{\emptyset}) > 0$ and the assignment id_{\emptyset} only appears in $\Delta(C)$ for initial arguments. Thus, NTNE-WCMP is in P as well.

5.3.2 SKEPTICAL ACCEPTANCE

The skeptical acceptance problem SA- ρ asks whether the likelihood of a given argument A exceeds a given positive threshold θ under all ρ -distributions. Like the non-emptiness problem, this problem is trivially solvable for all semantics that necessarily contain the distribution $\mu_0 = \text{Dirac}_{\emptyset}$, as $\mu_0(A) = 0 \not\geq \theta$ for any $\theta > 0$. Further, for all our notions of complete semantics, the problem is equivalent to deciding if the argument A is part of the grounded set \mathcal{S}_{GR} , as shown by the following corollary of Lemma 33. By (Dvořák, 2012), checking if an argument is in the grounded argument set is P -complete.

Corollary 68. *Let ρ be any of the complete semantics from Definition 29. Given an AF $\mathcal{F} = \langle \text{Arg}, \text{Att} \rangle$, an argument $A \in \text{Arg}$ is skeptically accepted w.r.t. to any threshold $\theta \in (0, 1]$ under ρ if and only if A is in the grounded argument set \mathcal{S}_{GR} of \mathcal{F} .*

Proof. By Lemma 33, we have $\mathcal{S}_{\text{GR}} = \bigcap_{\mu \in \llbracket \mathcal{F} \rrbracket_\rho} \text{Arg}_\mu$. For $A \in \mathcal{S}_{\text{GR}}$, this entails $\mu(A) = 1 \geq \theta$ for any threshold in $(0, 1]$, so A is skeptically accepted under ρ . Vice-versa, for $\theta = 1$ the set $\bigcap_{\mu \in \llbracket \mathcal{F} \rrbracket_\rho} \text{Arg}_\mu$ contains exactly the skeptically accepted arguments. \square

5.3.3 1-SATURATION SEMANTICS

As shown in Section 3.4.1, a non-trivial 0/1-saturated distribution always exists for AFs that do not contain any self-attacking arguments, though the presence of self-attacks alone is not sufficient for no 0/1-saturated distribution to exist. We now focus on 1-saturation and present a polynomial-time bounded algorithm to decide whether a (non-trivial) 1-saturated distribution exists. The algorithm suffices to answer almost all reasoning problems for 1-SAT (see Lemma 72). The exception is credulous acceptance checking under a threshold of $\theta = 1$, which is subsequently covered with the help of a second algorithm in Lemma 74.

The general idea is as follows. Due to almost-sure conflict-freeness, all self-attacking arguments must have a likelihood of zero. If an argument is attacked solely by such arguments (or, alternatively, has no attackers), 1-saturation requires its likelihood to be one. Now again by conflict-freeness, arguments attacked by at least a single argument with likelihood one necessarily have a likelihood of zero. Thus, originating from self-attacking arguments, we can iteratively construct two sets of arguments that need to have a likelihood of zero or one under every 1-SAT distribution. The existence of a (non-trivial) 1-saturated distribution then hinges on the coverage and disjointness of these two sets. For example, a self-attacking argument with no other attackers would need to belong to both sets. As this cannot be, no 1-saturated distribution exists for AFs containing arguments that are solely attacked by themselves.

Formally, let $\mathcal{F} = \langle \text{Arg}, \text{Att} \rangle$ be an AF and let $\mathcal{X}_0 := \{A \in \text{Arg} : (A, A) \in \text{Att}\}$ denote the set of self-attacking arguments. We define two set-valued functions $\mathcal{G}, \mathcal{H}: 2^{\text{Arg}} \rightarrow 2^{\text{Arg}}$ as follows:

$$\begin{aligned} \mathcal{G}(\mathcal{X}) &:= \{B \in \text{Arg} : \neg B \subseteq \mathcal{X}\} \\ \mathcal{H}(\mathcal{X}) &:= \mathcal{X}_0 \cup \mathcal{G}(\mathcal{X})^\rightarrow \end{aligned}$$

That is, $\mathcal{G}(\mathcal{X})$ denotes the set of all arguments where all attackers belong to \mathcal{X} , and an argument is in $\mathcal{H}(\mathcal{X})$ iff it is self-attacking or has an attacker in $\mathcal{G}(\mathcal{X})$. Note that both \mathcal{G} and \mathcal{H} are *monotonic*, i.e., for all $\mathcal{A}, \mathcal{B} \in 2^{\text{Arg}}$ with $\mathcal{A} \subseteq \mathcal{B}$, we have $\mathcal{G}(\mathcal{A}) \subseteq \mathcal{G}(\mathcal{B})$ and $\mathcal{H}(\mathcal{A}) \subseteq \mathcal{H}(\mathcal{B})$. By the Knaster-Tarski theorem, the monotonicity of \mathcal{H} implies the existence of a least fixed-point which we denote by \mathcal{X}^* . Further, let $\mathcal{Y}^* := \mathcal{G}(\mathcal{X}^*)$. Now consider the two increasing sequences (\mathcal{X}_n) and (\mathcal{Y}_n) given by $\mathcal{X}_{n+1} := \mathcal{H}(\mathcal{X}_n)$ and $\mathcal{Y}_n := \mathcal{G}(\mathcal{X}_n)$ for $n \in \mathbb{N}$ and \mathcal{X}_0 as above. Then \mathcal{X}^* and \mathcal{Y}^* are the limits of the sequences (\mathcal{X}_n) and (\mathcal{Y}_n) , i.e., there exists $k \leq |\text{Arg}|$ such that $\mathcal{X}^* = \mathcal{X}_k = \bigcup_n \mathcal{X}_n$ and $\mathcal{Y}^* = \mathcal{Y}_k = \bigcup_n \mathcal{Y}_n$. This entails that the sets \mathcal{X}^* and \mathcal{Y}^* can be computed in polynomial time by determining \mathcal{X}_0 and iterating \mathcal{H} on \mathcal{X}_0 until \mathcal{X}^* is found, which also yields $\mathcal{Y}^* = \mathcal{G}(\mathcal{X}^*)$.

Example 69. *We demonstrate the algorithm on the AF from Figure 7. Argument A is the only self-attacking argument, so $\mathcal{X}_0 = \{A\}$. Then $\mathcal{Y}_0 = \mathcal{G}(\mathcal{X}_0) = \{B, C\}$, as $\neg B = \{A\} \subseteq \mathcal{X}_0$ and $\neg C = \{A\} \subseteq \mathcal{X}_0$ holds. While D and A itself are also attacked by A , note that they are*

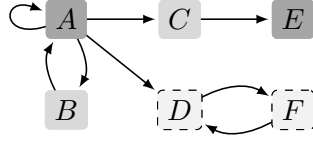


Figure 7: Under 1-saturated distributions the arguments A and E must have likelihood zero, B and C must have likelihood one, and no constraints arise for D and F

not part of \mathcal{Y}_0 as they each have another attacker not in \mathcal{X}_0 . For the next iteration, we get:

$$\begin{aligned}\mathcal{X}_1 &= \mathcal{H}(\mathcal{X}_0) = \{A\} \cup \{B, C\}^\rightarrow = \{A, E\} \\ \mathcal{Y}_1 &= \mathcal{G}(\{A, E\}) = \{B, C\}.\end{aligned}$$

That is, \mathcal{X}_1 contains all self-attacking arguments as well as those attacked by \mathcal{Y}_0 , namely A and E . As E does not have any outgoing attacks, \mathcal{Y}_1 remains unchanged and equals \mathcal{Y}_0 . Subsequently, we also get $\mathcal{X}_2 = \mathcal{X}_1$, thus $\mathcal{X}^* = \{A, E\}$ and $\mathcal{Y}^* = \{B, C\}$.

Now assume we remove the attack from B to A . Then A is also part of \mathcal{Y}_0 and we get:

$$\begin{aligned}\mathcal{X}_0 &= \{A\} \\ \mathcal{Y}_0 &= \mathcal{G}(\mathcal{X}_0) = \{A, B, C\} \\ \mathcal{X}_1 &= \mathcal{H}(\mathcal{X}_0) = \{A\} \cup \{A, B, C\}^\rightarrow = \{A, B, C, D, E\} \\ \mathcal{Y}_1 &= \mathcal{G}(\mathcal{X}_1) = \{A, B, C, D, E, F\}.\end{aligned}$$

In that case, both \mathcal{X}^* and \mathcal{Y}^* would contain all arguments of the AF.

Lemma 70. For each 1-saturated distribution μ , we have

- (a) $\mu(A) = 0$ for each $A \in \mathcal{X}^*$, and
- (b) $\mu(B) = 1$ for each $B \in \mathcal{Y}^*$.

Proof. We show $\mu(A) = 0$ for each $A \in \mathcal{X}_n$ and $\mu(B) = 1$ for $B \in \mathcal{Y}_n$ by induction on n .
 Base case. (a): Each $A \in \mathcal{X}_0$ is self-attacking, so $\mu(A) = 0$ follows from the fact that μ is almost-surely conflict-free. (b): For $B \in \mathcal{Y}_0$, we know each attacker of B is in \mathcal{X}_0 and thus $\mu(\bigvee_{A \in \rightarrow B} A) = 0$. By the 1-SAT constraint, this implies $\mu(B) = 1$.
 Step of induction. (a): Let $A \in \mathcal{X}_{i+1} = \mathcal{H}(\mathcal{X}_i) = \mathcal{X}_0 \cup \mathcal{G}(\mathcal{X}_i)^\rightarrow$. If $A \in \mathcal{X}_0$ the same reasoning as in the base case applies. For $A \in \mathcal{G}(\mathcal{X}_i)^\rightarrow = \mathcal{Y}_i^\rightarrow$, there is a $B \in \mathcal{Y}_i$ which attacks A and, by the induction hypothesis, $\mu(B) = 1$. Then $\mu(A)$ must be zero by the conflict-freeness of μ . (b): Let $B \in \mathcal{Y}_{i+1} = \mathcal{G}(\mathcal{X}_{i+1})$. Then each attacker A of B is in \mathcal{X}_{i+1} , so as before we have $\mu(A) = 0$ and thus $\mu(\bigvee_{A \in \rightarrow B} A) = 0$ and $\mu(B) = 1$ as μ is 1-saturated. \square

An immediate consequence of Lemma 70 is that no 1-saturated distribution can exist for AFs where \mathcal{X}^* and \mathcal{Y}^* are not disjoint, i.e., where $\mathcal{X}^* \cap \mathcal{Y}^* \neq \emptyset$. Let $\mathcal{Z}^* := \text{Arg} \setminus (\mathcal{X}^* \cup \mathcal{Y}^*)$ denote the set of remaining arguments that are neither in \mathcal{X}^* or \mathcal{Y}^* . We will need the following observations about the arguments in \mathcal{X}^* , \mathcal{Y}^* , and \mathcal{Z}^* .

Lemma 71. *For an AF \mathcal{F} where \mathcal{X}^* and \mathcal{Y}^* are disjoint, it holds:*

- (a) *For each $C \in \mathcal{Z}^*$, the set $\mathcal{Y}^* \cup \{C\}$ is conflict-free.*
- (b) *Each $C \in \mathcal{Z}^*$ has at least one attacker in \mathcal{Z}^* , i.e., $\neg C \cap \mathcal{Z}^* \neq \emptyset$.*
- (c) *Each $A \in \mathcal{X}^*$ has at least one attacker in \mathcal{Y}^* or \mathcal{Z}^* , i.e., $\neg A \cap (\mathcal{Y}^* \cup \mathcal{Z}^*) \neq \emptyset$.*

Proof. (a): First note that \mathcal{Y}^* is conflict-free as $\neg B \subseteq \mathcal{X}^*$ for all $B \in \mathcal{Y}^*$. As $C \notin \mathcal{X}^*$, this entails that C does not attack any argument in \mathcal{Y}^* . Further, as $\mathcal{X}^* = \mathcal{H}(\mathcal{X}^*) = \mathcal{X}_0 \cup \mathcal{G}(\mathcal{X}^*)^\rightarrow = \mathcal{X}_0 \cup \mathcal{Y}^{*\rightarrow}$, we have $C \notin \mathcal{Y}^{*\rightarrow}$, so C has no attackers in \mathcal{Y}^* .

(b): As $C \notin \mathcal{Y}^*$, we have $\neg C \not\subseteq \mathcal{X}^*$, so $\neg C$ is non-empty. By (a), we also have $\neg C \cap \mathcal{Y}^* = \emptyset$, thus $\neg C \cap \mathcal{Z}^* \neq \emptyset$ must hold.

(c): Suppose by contradiction that $\neg A \cap (\mathcal{Y}^* \cup \mathcal{Z}^*) = \emptyset$ for some $A \in \mathcal{X}^*$. Then $\neg A \subseteq \mathcal{X}^*$. But then $A \in \mathcal{G}(\mathcal{X}^*) = \mathcal{Y}^*$, which contradicts the assumption that \mathcal{X}^* and \mathcal{Y}^* are disjoint. \square

With Lemmas 70 and 71 at hand, we are now ready to show how the reasoning problems NE-1-SAT, NTNE-1-SAT, CA-1-SAT (for $\theta < 1$), and SA-1-SAT can be answered given the sets \mathcal{X}^* and \mathcal{Y}^* . Thus, they are tractable in polynomial time.

Lemma 72 (Reasoning problems for 1-SAT). *Let $\mathcal{F} = \langle \text{Arg}, \text{Att} \rangle$ be an AF and $\mathcal{X}^*, \mathcal{Y}^*, \mathcal{Z}^* \subseteq \text{Arg}$ as before.*

- (a) *There is a 1-saturated distribution for \mathcal{F} iff $\mathcal{X}^* \cap \mathcal{Y}^* = \emptyset$.*
- (b) *There is a non-trivial 1-saturated distribution for \mathcal{F} iff $\mathcal{X}^* \cap \mathcal{Y}^* = \emptyset$ and \mathcal{Z}^* or \mathcal{Y}^* are non-empty.*
- (c) *For $A \in \text{Arg}$ and a threshold $\theta \in (0, 1)$, there is a 1-saturated distribution for \mathcal{F} with $\mu(A) \geq \theta$ iff $\mathcal{X}^* \cap \mathcal{Y}^* = \emptyset$ and $A \in \mathcal{Y}^*$ or $A \in \mathcal{Z}^*$.*
- (d) *For $A \in \text{Arg}$ and $\theta \in (0, 1]$, then $\mu(A) \geq \theta$ holds for all 1-saturated distributions for \mathcal{F} iff either $\mathcal{X}^* \cap \mathcal{Y}^*$ is non-empty (in which case there is no 1-saturated distribution) or $\mathcal{X}^* \cap \mathcal{Y}^* = \emptyset$ and $A \in \mathcal{Y}^*$.*

Proof. For (a), the direction 1-SAT $\Rightarrow \mathcal{X}^* \cap \mathcal{Y}^* = \emptyset$ follows directly from Lemma 70. The other direction follows from (b) in case \mathcal{Z}^* or \mathcal{Y}^* are non-empty, otherwise it is easy to see that the distribution $\text{Dirac}_{\mathcal{Y}^*}$ is 1-saturated.

(b) immediately follows from (c). From Lemma 70, we also get $\mu(A) \geq \theta$ for all $A \in \mathcal{Y}^*$, $\theta \in (0, 1]$, and 1-saturated distributions μ . It remains to show the case $A \in \mathcal{Z}^*$ for (c) and (d), i.e., that there is a 1-saturated distribution μ with $\mu(A) \geq \theta$ for $A \in \mathcal{Z}^*$, but this does not hold for all 1-saturated distributions. We show this by constructing a 1-saturated distribution μ_A^θ with $\mu_A^\theta(A) = \theta$ for $A \in \mathcal{Z}^*$ and any $\theta \in (0, 1)$.

For any $C \in \mathcal{Z}^*$, let β_C denote the assignment $\text{id}_{\mathcal{Y}^* \cup \{C\}}$. Then we define the distribution μ_A^θ as follows:

$$\mu_A^\theta(\beta_A) = \theta \quad \text{and} \quad \mu_A^\theta(\beta_C) = \frac{1 - \theta}{|\mathcal{Z}^*| - 1} \quad \text{for } C \in \mathcal{Z}^* \setminus \{A\}.$$

For all $C \in \mathcal{Z}^* \setminus \{A\}$, we then have a positive likelihood $\mu_A^\theta(C) > 0$ as $|\mathcal{Z}^*| \geq 2$ (from Lemma 71b) and $\theta < 1$. The distribution μ_A^θ is 1-saturated iff it is almost-surely conflict-free and $\mu_A^\theta(B) < 1$ implies $\mu_A^\theta(\bigvee_{C \in \rightarrow B} C) > 0$ for all $B \in \text{Arg}$. The former follows as each assignment β_C in the support of μ_A^θ is conflict-free by Lemma 71a. For $B \in \mathcal{Y}^*$, we have $\mu_A^\theta(B) = 1$, so the condition holds trivially. For $B \in \mathcal{Z}^*$, we have $\mu_A^\theta(B) < 1$, but also $\mu_A^\theta(\bigvee_{C \in \rightarrow B} C) > 0$ by Lemma 71b. For $B \in \mathcal{X}^*$, we have $\mu_A^\theta(B) = 0$, but again $\mu_A^\theta(\bigvee_{C \in \rightarrow B} C) > 0$ by Lemma 71c. \square

Example 73. *Continuing Example 69, we have $\mathcal{X}^* = \{A, E\}$, $\mathcal{Y}^* = \{B, C\}$, and $\mathcal{Z}^* = \{D, F\}$ for the AF from Figure 7. By Lemma 72, as \mathcal{X}^* and \mathcal{Y}^* are disjoint and \mathcal{Z}^* is non-empty, we know that a non-trivial 1-saturated distributions exists. On the other hand, when dropping the edge from B to A, \mathcal{X}^* and \mathcal{Y}^* both contain all arguments, thus no 1-SAT distribution can exist.*

Note that Lemma 72 (c) covers the credulous acceptance problem only for thresholds θ that are strictly less than one. This is because requiring $\mu(C) = 1$ for a 1-saturated distribution comes with a number of implications:

1. $\mu(A) = 1$ implies $\mu(B) = 0$ for $B \in \rightarrow A \cup A \rightarrow$ by ASCF.
2. $\mu(A) < 1$ implies $\mu(\bigvee_{B \in \rightarrow A} B) > 0$ by 1-SAT.
3. $\mu(\bigvee_{B \in \rightarrow A} B) = 0$ implies $\mu(A) = 1$ by 1-SAT.

Therefore, when argument C belongs to \mathcal{Z}^* , requiring $\mu(C) = 1$ can lead to more arguments that need to be added to \mathcal{X}^* and \mathcal{Y}^* . However, the credulous acceptance problem is still tractable for threshold $\theta = 1$ as shown by Algorithm 1 and Lemma 74. Intuitively, the algorithm tries to add C to \mathcal{X}^* while keeping track of the three implications from above. If a violation occurs, C is not credulously accepted, otherwise a 1-SAT distribution with $\mu(C) = 1$ necessarily exists.

Lemma 74. *For an AF $\mathcal{F} = \langle \text{Arg}, \text{Att} \rangle$ and an argument $C \in \text{Arg}$, Algorithm 1 decides in polynomial time whether C is credulously accepted under 1-saturated semantics and threshold $\theta = 1$.*

Proof. C is credulously accepted iff there is a distribution $\mu \in \llbracket \mathcal{F} \rrbracket_{1\text{-SAT}}$ with $\mu(C) = 1$. Lines 2–6 of Algorithm 1 are covered by Lemma 72, for the remainder we can assume $C \in \mathcal{Z}^*$. Note that the algorithm observes the following invariants when a 1-saturated distribution μ with $\mu(C) = 1$ exists:

- $\mu(A) = 0$ for all $A \in \mathcal{S}^0$,
- $\mu(A) = 1$ for all $A \in \mathcal{S}^1$,
- $\mu(\bigvee_{A \in \mathcal{S}} A) > 0$ for all $\mathcal{S} \in \mathfrak{D}$.

We first consider the case that the algorithm returns **no**, i.e., we end up in line 14, 21, or 27. In line 14, the current argument B is contained in \mathcal{S}^0 (and thus $\mu(B) = 0$), but should be added to \mathcal{S}^1 , which would imply $\mu(B) = 1$. Hence, no such distribution μ can exist. The

Algorithm 1: Decide credulous acceptance for argument C under 1-SAT and $\theta = 1$

Input: $C \in Arg$
Output: yes or no

- 1 **Function** IsCredAccept(C)
- 2 Compute \mathcal{X}^* , \mathcal{Y}^* , \mathcal{Z}^* as above
- 3 **if** $\mathcal{X}^* \cap \mathcal{Y}^* \neq \emptyset$ **or** $C \in \mathcal{X}^*$ **then**
- 4 **return no**
- 5 **if** $C \in \mathcal{Y}^*$ **then**
- 6 **return yes**
- 7 $\mathcal{S}^0 \leftarrow \mathcal{X}^*$ // \mathcal{S}^0 collects all arguments that are set to zero
- 8 $\mathcal{S}^1 \leftarrow \mathcal{Y}^*$ // \mathcal{S}^1 collects all arguments that are set to one
- 9 $\mathcal{D} \leftarrow \{\neg A \cap \mathcal{Z}^* : A \in \mathcal{S}^0 \text{ where } \neg A \cap \mathcal{S}^1 = \emptyset\}$
 // \mathcal{D} collects sets of arguments that cannot all be set to zero
- 10 SetToOne(C)
- 11 **return yes**
- 12 **Function** SetToOne(B)
- 13 **if** $B \in \mathcal{S}^0$ **then**
- 14 **return no**
- 15 $\mathcal{S}^1 \leftarrow \mathcal{S}^1 \cup \{B\}$
- 16 $\mathcal{D} \leftarrow \{S \in \mathcal{D} : B \notin S\}$ // Remove all sets in \mathcal{D} that contain B
- 17 **for** $B' \in (\neg B \cup B^\rightarrow) \setminus \mathcal{S}^0$ **do**
- 18 SetToZero(B') // All arguments attacked by or attacking B are set to zero
- 19 **Function** SetToZero(B)
- 20 **if** $B \in \mathcal{S}^1$ **then**
- 21 **return no**
- 22 $\mathcal{S}^0 \leftarrow \mathcal{S}^0 \cup \{B\}$
- 23 $\mathcal{D} \leftarrow \{S \setminus \{B\} : S \in \mathcal{D}\}$ // Remove B from all sets in \mathcal{D}
- 24 **if** $\neg B \cap \mathcal{S}^1 = \emptyset$ **then**
- 25 $\mathcal{D} \leftarrow \mathcal{D} \cup \{\neg B \setminus \mathcal{S}^0\}$
- 26 **if** $\emptyset \in \mathcal{D}$ **then**
- 27 **return no** // all arguments in a set in \mathcal{D} were set to zero
- 28 **for** $B' \in B^\rightarrow \setminus \mathcal{S}^1$ **do**
- 29 **if** $\neg B' \setminus \mathcal{S}^0 = \emptyset$ **then**
- 30 SetToOne(B')
 // All its attackers are set to zero, so B' is set to one

same holds for line 21 with \mathcal{S}^0 and \mathcal{S}^1 switched. In line 27, we have $\emptyset \in \mathfrak{D}$. But then, by the invariant, $\mu(\bigvee_{A \in \emptyset} A) = 0 \not\equiv 0$.

Now assume the algorithm returns **yes**. Let $\mathcal{S}^? = \text{Arg} \setminus (\mathcal{S}^0 \cup \mathcal{S}^1)$. Then any distribution $\tilde{\mu}$ with support $\text{Supp}(\tilde{\mu}) = \{\text{id}_{\mathcal{S}^1}\} \cup \{\text{id}_{\{A\} \cup \mathcal{S}^1} : A \in \mathcal{S}^?\}$ satisfies $\tilde{\mu}(C) = 1$ and $0 < \tilde{\mu}(A) < 1$ for all $A \in \mathcal{S}^?$. We show that any such distribution $\tilde{\mu}$ is 1-saturated. Let $A \in \mathcal{S}^?$. Since $\tilde{\mu}(A) < 1$, we have to show $\tilde{\mu}(\bigvee_{B \in \rightarrow A} B) > 0$. Note that no attacker of A can belong to \mathcal{S}^1 as each argument in \mathcal{S}^1 either is in \mathcal{Y}^* (which does not attack \mathcal{Z}^* by Lemma 71a) or is added by **SetToOne**, in which case A would have been added to \mathcal{S}^0 by line 18. Further, at least one attacker of A belongs to $\mathcal{S}^?$: By Lemma 71c there is at least one attacker in \mathcal{Z}^* , and not all attackers can belong to \mathcal{S}^0 as otherwise A would be added to \mathcal{S}^1 by line 30. Thus, $\tilde{\mu}(\bigvee_{B \in \rightarrow A} B) > 0$ holds.

Finally, note that the algorithm always terminates as the set \mathcal{Z}^* is finite and **SetToOne** and **SetToZero** are called at most once for each argument. All set operations are possible in polynomial time and the size of \mathfrak{D} is bounded by the number of arguments, so the overall runtime is polynomial. \square

5.4 Hardness Results

We now show *NP*-hardness of the credulous acceptance problem for most of our semantics via a polynomial reduction from 3SAT. We make use of the “standard translation” of a CNF formula as given in (Dvořák & Dunne, 2018). Let ψ be a 3CNF formula over variables in $X = \{x_1, \dots, x_n\}$ with k clauses $\kappa_1, \dots, \kappa_k$. That is, $\psi = \kappa_1 \wedge \kappa_2 \wedge \dots \wedge \kappa_k$ where $\kappa_i = \xi_{1,i} \vee \xi_{2,i} \vee \xi_{3,i}$ and the $\xi_{i,j}$ ’s are literals in $\text{Lit} = \{x_1, \dots, x_n, \neg x_1, \dots, \neg x_n\}$ (for $j \in \{1, \dots, k\}$ and $i \in \{1, 2, 3\}$). W.l.o.g. ψ has no tautologic clauses, i.e., there is no $j \in \{1, \dots, k\}$ and no $x \in X$ with $\{x, \neg x\} \subseteq \{\xi_{1,j}, \xi_{2,j}, \xi_{3,j}\}$. Now we define an AF $\mathcal{F}_\psi = \langle \text{Arg}_\psi, \text{Att}_\psi \rangle$ induced by ψ as follows. The set Arg_ψ contains arguments for all literals, all clauses, and the formula ψ itself. That is:

$$\text{Arg}_\psi := \{A_z : z \in \text{Lit} \cup \text{Clauses} \cup \{\psi\}\} \quad \text{where} \quad \text{Clauses} = \{\kappa_1, \dots, \kappa_k\}.$$

The attack relation Att_ψ consists of the following pairs (where we identify each clause with the set of its literals, i.e., κ_j is identified with $\{\xi_{1,j}, \xi_{2,j}, \xi_{3,j}\}$):

- $(A_x, A_{\neg x})$ and $(A_{\neg x}, A_x)$ for each variable $x \in X$,
- (A_ξ, A_κ) for each $\kappa \in \text{Clauses}$ and each literal $\xi \in \kappa$,
- (A_κ, A_ψ) for each $\kappa \in \text{Clauses}$.

Thus, $\rightarrow A_{\kappa_j} = \{A_{\xi_{1,j}}, A_{\xi_{2,j}}, A_{\xi_{3,j}}\}$, $A_{\kappa_j} \rightarrow = \{A_\psi\}$, and $\rightarrow A_\psi = \{A_{\kappa_1}, \dots, A_{\kappa_k}\}$. An example is given in Figure 8.

Lemma 75. *Let ψ be a 3CNF formula and $\mathcal{F}_\psi = \langle \text{Arg}_\psi, \text{Att}_\psi \rangle$ the corresponding AF obtained by the standard construction. Then the following statements are equivalent:*

- (a) ψ is satisfiable.
- (b) \mathcal{F}_ψ has a stable set containing A_ψ .

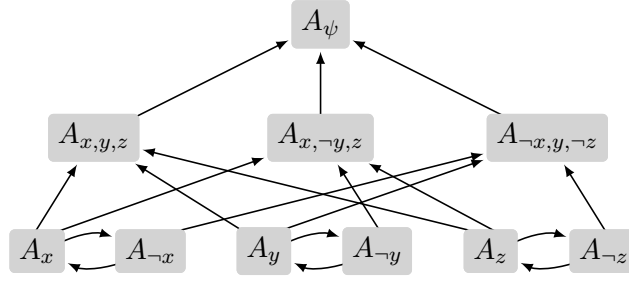


Figure 8: The AF \mathcal{F}_ψ corresponding to the formula $\psi = (x \vee y \vee z) \wedge (x \vee \neg y \vee z) \wedge (\neg x \vee y \vee \neg z)$.

- (c) *There is an element-wise stable distribution μ for \mathcal{F}_ψ with $\mu(A_\psi) = 1$.*
 (d) *There is a weakly admissible distribution μ for \mathcal{F}_ψ with $\mu(A_\psi) = 1$.*

Proof. The equivalence of (a) and (b) follows from the construction of \mathcal{F}_ψ and is detailed in (Dvořák & Dunne, 2018). Intuitively, the formula ψ holds iff all its clauses are satisfied, and likewise, argument A_ψ is in a stable set iff all its attackers, the arguments corresponding to the clauses, are not in the set. Each clause κ is satisfied iff at least one of its literals is true. Likewise, A_κ is not in a stable set if at least one of its attackers, the arguments corresponding to the literals, is part of the set. Finally, a model for ψ is boolean assignment to all variables in X , whereas in a stable set either A_x or $A_{\neg x}$ need to be contained for every variable $x \in X$.

(b) \iff (c) follows from Lemma 7, and (c) \implies (d) from the established relationships as summarized in Figure 6. To complete the proof it suffices to show (d) \implies (a). Assume $\mu \in \text{Distr}(\mathcal{F}_\psi)$ is weakly admissible and $\mu(A_\psi) = 1$. Then argument A_ψ is almost-surely defended, i.e., $\mu(\Delta(A_\psi)) = 1$, and there is at least one assignment $\beta \in \text{Supp}(\mu)$ which satisfies

$$\Delta(A_\psi) = \bigwedge_{A_\kappa \in \neg A_\psi} \bigvee_{A_\xi \in \neg A_\kappa} A_\xi = \bigwedge_{j=1}^k (A_{\xi_{1,j}} \vee A_{\xi_{2,j}} \vee A_{\xi_{3,j}}).$$

As $\beta \models \Delta(A_\psi)$, we have $\beta \models A_{\xi_{1,j}} \vee A_{\xi_{2,j}} \vee A_{\xi_{3,j}}$ for every $j \in \{1, \dots, k\}$. That is, for each clause κ of ψ there is a literal $\xi \in \kappa$ with $\beta(A_\xi) = \top$.

Let $\alpha \in \text{Asg}(X)$ be given by $\alpha(x) = \beta(A_x)$ for each variable $x \in X$. We now show that α is a satisfying assignment for ψ . Let X' denote the set of variables $x \in X$ such that $\beta(A_x) = \top$ or $\beta(A_{\neg x}) = \top$, and $\text{Lit}' = \{x, \neg x : x \in X'\}$. (The variables in $X \setminus X'$ are those where $\beta(A_x) = \beta(A_{\neg x}) = \text{f}$. These variables are, however, irrelevant for the proof that α is satisfying for ψ .) As β is conflict-free and the literals x and $\neg x$ attack each other:

$$\text{For each literal } \xi \in \text{Lit}' : \quad \alpha \models \xi \text{ iff } \beta(A_\xi) = \top.$$

But then, the above yields that for each clause κ of ψ there is a literal $\xi \in \kappa \cap \text{Lit}'$ with $\alpha \models \xi$. Thus, α is a satisfying assignment for the 3CNF formula ψ . \square

Corollary 76. *Deciding the credulous acceptance problem CA- ρ is NP-hard for the admissibility and complete semantics from Definitions 14–19 and 29, as well as the 0-SAT and 0/1-SAT semantics.*

6. Implementation and Evaluation

We have developed a prototypical implementation to support the understanding and evaluation of abstract argumentation frameworks and the proposed probabilistic semantics. Our tool is capable of solving the reasoning problems specified in Definition 66 for all semantics given in the taxonomy overview in Figure 6, including those by Hunter and Thimm (2017). The non-emptiness problem is covered by the basic functionality to synthesize one (or more) distributions that satisfy the given semantics’ constraints, or to assert that no such distribution exists. The non-trivial non-emptiness problem can easily be checked using the features to enforce multiple semantics simultaneously and to consider complement semantics: A synthesized distribution then satisfies all the constraints of each regular semantics, but violates at least one constraint of each specified complement semantics. Thus, to check if a non-trivial distribution exists under some semantics, one simply requires the complement of the MIN semantics to hold as well.

Further features and use cases supported by the tool include the following:

Likelihood optimization. For a given AF \mathcal{F} , an argument A , and a semantics ρ , find a distribution $\mu \in \llbracket \mathcal{F} \rrbracket_\rho$ such that $\mu(A)$ is maximal (or minimal) w.r.t. all other distributions in $\llbracket \mathcal{F} \rrbracket_\rho$. This task requires the chosen semantics ρ to be linear, i.e., to induce only linear constraints. As mentioned before all semantics except the weakly admissible, weakly complete, and the saturation semantics are linear.

Solution space enumeration. For each linear semantics ρ , the solution space (i.e., the set $\llbracket \mathcal{F} \rrbracket_\rho$ as subspace of \mathbb{R}^n when each distribution is viewed as n -dimensional vector) is a convex polytope. This feature allows to export the distributions at the corner of this polytope, yielding a finite representation of all solutions even in case there are infinitely many. Then each distribution in $\llbracket \mathcal{F} \rrbracket_\rho$ arises as convex combination from these corner distributions.

SMT constraints. In addition to the semantics’ constraints, context specific constraints can be added to an AF in SMT-LIB format (Barrett, Fontaine, & Tinelli, 2016). Among others, this allows to specify the following constraints on a probability distribution μ over arguments A and B :

- *Interval containment* constraints $\mu(A) \in [u, \ell]$ specify that argument A is accepted with some probability $u \leq \mu(A) \leq \ell$.
- *Complement* constraints $\mu(A) = 1 - \mu(B)$ specify that the interpretation of the arguments A and B are complementary, i.e., their likelihoods sum up to one as counter events do.
- *Scalar dependency* constraints $\mu(A) = k \cdot \mu(B)$ specify that the likelihoods of arguments A and B are proportional w.r.t. some factor $k \in \mathbb{R}$.
- *Conditional probability* constraints $\mu(A \mid B) = p$ specify that p is the probability of A being accepted given that argument B is accepted. This constraint can be represented by a variant of the scalar dependency: $\mu(A \wedge B) = p \cdot \mu(B)$.

Back-end selection. A number of different state-of-the-art SMT solvers can be selected as back-end via the pySMT interface (Gario & Micheli, 2015) and number of linear solvers is available via CVXOPT (Andersen, Dahl, & Vandenberghe, 2014).

Labelings. All labeling schemes discussed in Section 3.5 can be used to generate one or all labelings for the chosen semantics.

Let us now provide some examples for the application of our tool on instances of our running example.

Example 77. *Example 64 was in fact produced by our tool, synthesizing a distribution that satisfies all constraints of the semantics JUS and ELM-CF, as well as the complement semantics of WCMP. Additionally, we imposed the following constraints on the likelihoods of some of the arguments: $\mu(\overline{\text{cl}}) = 0.7$, $\mu(\text{ld}) = 0.7$, $\mu(\text{cr} \Rightarrow \text{r}) \in [0.7, 1]$, and $\mu(\text{cr} \Rightarrow \text{m}) \in [0.4, 1]$. Intuitively, e.g., the latter means that the probability of camera right truly detecting an object in the middle in front of the car is between 0.4 and 1.*

Example 78. *For Example 65, the semantics JNTCMP and the complement of JUS were enforced, along with the same constraints on argument likelihoods as above. Furthermore, some context-specific constraints were imposed in an additional file in SMT-LIB format:*

- *In the example, the arguments for the three sensors come with a complement, e.g., ld and $\overline{\text{ld}}$. We added constraints such that these arguments are in fact their respective inverse by complement constraints such as $\mu(\text{ld}) = 1 - \mu(\overline{\text{ld}})$.*
- *A maximum 2% risk of a false positive detection was enforced via the conditional probability constraint $\mu(\text{cr} \Rightarrow \text{r} \vee \text{cr} \Rightarrow \text{m} \mid \text{cr}) \geq 0.98$.*
- *In line with the sensor arrangement visualized in Figure 1, we enforced a scalar dependency of arguments $\text{cr} \Rightarrow \text{r}$ and $\text{cr} \Rightarrow \text{m}$ by the constraint $\mu(\text{cr} \Rightarrow \text{r}) = 2 \cdot \mu(\text{cr} \Rightarrow \text{m})$. Intuitively, this means that one third of the right camera view angle is monitoring the (overlapping) middle.*

6.1 Tool Architecture

Each semantics is implemented as a function that takes a representation of an AF as input and returns a set of constraints on the induced joint distribution. This approach enables the combination of semantics as needed, including the option to consider complement semantics, and supports easy addition of new probabilistic semantics to the tool. Tasks like credulous or skeptical acceptance checking are tackled by adding additional constraints as described in Section 5.1.

Two kinds of solver back-ends are available to tackle different tasks. First, SMT solvers like Z3 (de Moura & Bjørner, 2008) are able to handle arbitrary polynomial constraints in the existential theory of the reals. This covers all semantics considered in this paper as well as their complements. Second, linear-optimization solvers can be used for likelihood optimization and solution space enumeration, provided the selected semantics' constraints are linear.

instance	#nodes	distribution synthesis		max probability of ct		solution polytope	
cl-ld-cr	6	0.017	(0.004)	0.006	(0.003)	0.022	(0.003)
cl-ld-cr	8	0.055	(0.018)	0.030	(0.018)	1.076	(0.018)
cl-ld-cr	12	1.286	(0.651)	2.736	(0.762)	> 600	timeout
cl-ld-cr	14	7.423	(3.913)	43.093	(15.887)	> 600	timeout
cl-ld-cr	18	361.033	(119.800)	> 600	timeout	> 600	timeout

Table 2: Running times in seconds on subgraphs of the vehicle example with increasing size for different tasks. cl-ld-cr is the full example, otherwise the arguments in grey stand for omitted sensors that are dropped including related edges. The time to generate the constraints is given in parenthesis.

Example 79 (Constraint encoding). *For a simplistic AF $\mathcal{F} = \langle \{A, B\}, \{(A, B)\} \rangle$ with two arguments A and B and a single attack $A \rightarrow B$, a constraint like $\mu(A) = 1/2$ is encoded as SMT constraint as follows: First of all, for each assignment $\beta \in \text{Asg}(\{A, B\})$, a floating-point variable v_β is created, along with the constraints $v_\beta \geq 0$ and $v_\beta \leq 1$. Exemplary, for $\beta = \{A=\top, B=F\}$, we obtain the variable $v_{A\bar{B}}$. Additionally, we have to enforce the constraint $\sum_\beta \mu(\beta) = 1$, i.e., $v_{\bar{A}\bar{B}} + v_{\bar{A}B} + v_{A\bar{B}} + v_{AB} = 1$. Then $\mu(A) = 1/2$ is encoded as $v_{A\bar{B}} + v_{AB} = 0.5$.*

For linear constraints, μ is seen as a vector of length 4. We always get the constraints $\langle 1 \ 1 \ 1 \ 1 \rangle \cdot \mu = 1$ as well as $\langle 1 \ 0 \ 0 \ 0 \rangle \cdot \mu \geq 0$, $\langle 1 \ 0 \ 0 \ 0 \rangle \cdot \mu \leq 1$, etc. Then $\mu(A) = 1/2$ is encoded as $\langle 0 \ 0 \ 1 \ 1 \rangle \cdot \mu = 0.5$.

The tool is implemented in Python and can either be used via a rich command-line interface or directly by including the provided Python package. Further implementation details are given in (Käfer, 2020).

6.2 Running Times

All experiments were conducted on an Intel i9-10900K machine with 64GB of RAM, running Ubuntu 20.10 and Python 3.8.6. Computing solutions according to Example 77 took seventy minutes, while less than five minutes were required to compute the solution for Example 78. Table 2 provides further statistics for typical tasks applied to the vehicle example and some of its subgraphs under PRCMP semantics. The majority of the time spend in the tool itself is taken up by the constraint generation (given in parentheses), the remaining time is mostly used by the back-end solvers to solve the imposed constraint system.

With distributions ranging over all possible assignments of an AF’s arguments, it is clear that increasing the number of arguments in an AF leads to an exponential growth of the associated distributions. This also affects the size of the generated constraints and is reflected both in the time spend in the tool as well as the time needed by the solver, as evident by the timings for the distribution synthesis task. The optimization task to find a distribution under which $\mu(\text{ct})$ is maximal is more involved but still succeeds in time for all but the largest instance. The final task of finding the distributions at the corner points of the solution polytope comes with the additional challenge that especially for larger instances the

number of such corner distributions can become very large itself, so memory consumption may arise as additional limiting factor.

The tool, further documentation, and all experimental data are publicly available at <https://www.perspicuous-computing.science/cpraa/>.

7. Conclusion

In this paper we contributed to the quest for quantitative abstract argumentation frameworks from a probability-theoretic perspective. At the core of our approach, we are viewing each semantics as inducing sets of constraints on the joint distribution over argument sets. We have provided a profound study of admissibility and complete semantics and have discussed a hierarchy of resulting semantics, also in relation to earlier work. We have formulated decision problems for probabilistic semantics and provided a range of complexity results.

We experimented with these semantic notions on a probabilistic abstract argumentation framework inspired by an autonomous driving scenario. For this, we implemented a tool based on SMT solvers to harvest present and future advances. In particular, by providing generic support for including additional constraints, our tool is capable of addressing a variety of adapted semantic notions well beyond the notions of admissibility and complete semantics spelled out in this paper, as well as adapting them towards context-specific needs. Our tool is a research prototype, built for the ease of experimentation with semantic notions across the wider spectrum of probabilistic abstract argumentation frameworks. Indeed, it turned out very helpful for the authors of this paper to sharpen their intuition.

7.1 Future Work

The tool already offers some elementary optimization tasks like maximizing the likelihood of selected arguments. Expanding on this functionality, we plan to investigate quantifications of how close distributions are to satisfy a certain semantics. This could allow to find, e.g., the “most” min-admissible distribution even if no distribution exists that completely satisfies the MINADM constraints, or pave the way for a probabilistic adaption of the semi-stable semantics.

Finally, we identified two further directions that might be worthwhile to explore under the lens of our probabilistic approach. In *dynamic abstract argumentation*, nodes and edges can be added or removed from an initial AF, giving rise to the question of how these dynamic interventions affect, e.g., the accepted argument sets (see, e.g., Diller, Haret, Linsbichler, Rümmele, and Woltran (2018) or Doutre and Maily (2018) for a survey). In the probabilistic setting, changes to the resulting distributions could be quantified. Secondly, several logical languages for abstract argumentation have been developed (see, e.g., YALLA (Dupin de Saint-Cyr, Bisquert, Cayrol, & Lagasque-Schiex, 2016)). Such languages allow to encode and subsequently reason about the basic notions of abstract argumentation, and extensions in the probabilistic domain could prove useful.

Acknowledgments

This research was partially funded by the DFG through the Collaborative Research Center TRR 248 (see <https://perspicuous-computing.science>, project ID 389792660), Cluster of Excellence EXC 2050/1 (CeTI, project ID 390696704, as part of Germany’s Excellence Strategy), the Research Training Groups QuantLA (GRK 1763) and RoSI (GRK 1907), by the Bundesministerium für Bildung und Forschung (BMBF) Förderkennzeichen 01IS20056_NAVAS, by the European Research Council through the Advanced Investigators Grant 695614 (POWVER), and by the Key-Area Research and Development Program Grant 2018B010107004 of Guangdong Province.

References

- Amgoud, L., & Cayrol, C. (2002). A reasoning model based on the production of acceptable arguments. *Annals of Mathematics and Artificial Intelligence*, 34(1-3), 197–215.
- Amgoud, L., Cayrol, C., Lagasque-Schieux, M., & Livet, P. (2008). On bipolarity in argumentation frameworks. *International Journal of Intelligent Systems*, 23(10), 1–32.
- Andersen, M., Dahl, J., & Vandenberghe, L. (2014). Cvxopt – python software for convex optimization. <http://cvxopt.org/>.
- Baier, C., Diller, M., Dubslaff, C., Gaggl, S. A., Hermanns, H., & Käfer, N. (2021). Admissibility in Probabilistic Argumentation. In *Proceedings of the 18th International Conference on Principles of Knowledge Representation and Reasoning*, pp. 87–98.
- Baroni, P., Caminada, M., & Giacomin, M. (2011). An introduction to argumentation semantics. *Knowledge Engineering Review*, 26(4), 365–410.
- Baroni, P., Giacomin, M., & Vicig, P. (2014). On rationality conditions for epistemic probabilities in abstract argumentation. In Parsons, S., Oren, N., Reed, C., & Cerutti, F. (Eds.), *Proceedings of the 5th International Conference on Computational Models of Argument (COMMA 2014)*, Vol. 266 of *Frontiers in Artificial Intelligence and Applications*, pp. 121–132. IOS Press.
- Barrett, C., Fontaine, P., & Tinelli, C. (2016). The Satisfiability Modulo Theories Library (SMT-LIB). www.SMT-LIB.org.
- Baumann, R., Brewka, G., & Ulbricht, M. (2020). Revisiting the foundations of abstract argumentation - semantics based on weak admissibility and weak defense. In *The Thirty-Fourth AAAI Conference on Artificial Intelligence, AAAI 2020, The Thirty-Second Innovative Applications of Artificial Intelligence Conference, IAAI 2020, The Tenth AAAI Symposium on Educational Advances in Artificial Intelligence, EAAI 2020, New York, NY, USA, February 7-12, 2020*, pp. 2742–2749. AAAI Press.
- Bench-Capon, T. J. M. (2003). Persuasion in practical argument using value-based argumentation frameworks. *Journal of Logic and Computation*, 13(3), 429–448.
- Boella, G., Gabbay, D. M., van der Torre, L. W. N., & Villata, S. (2010). Support in abstract argumentation. In Baroni, P., Cerutti, F., Giacomin, M., & Simari, G. R. (Eds.), *Proceedings of the 3rd International Conference on Computational Models of Argument*

- (*COMMA 2010*), Vol. 216 of *Frontiers in Artificial Intelligence and Applications*, pp. 111–122. IOS Press.
- Brewka, G., Straß, H., Ellmauthaler, S., Wallner, J. P., & Woltran, S. (2013). Abstract Dialectical Frameworks Revisited. In Rossi, F. (Ed.), *Proceedings of the 23rd International Joint Conference on Artificial Intelligence (IJCAI 2013)*, pp. 803–809. IJCAI/AAAI.
- Brewka, G., & Woltran, S. (2010). Abstract Dialectical Frameworks. In Lin, F., Sattler, U., & Truszczyński, M. (Eds.), *Proceedings of the 12th International Conference on Principles of Knowledge Representation and Reasoning (KR 2010)*, pp. 102–111. AAAI Press.
- Caminada, M., Carnielli, W., & Dunne, P. (2006). Semi-stable semantics.. *22*(5), 1207–1254.
- Caminada, M., & Gabbay, D. (2009). A Logical Account of Formal Argumentation. *Studia Logica*, *93*(2-3), 109–145.
- de Finetti, B. (1974). *Theory of Probability*. Wiley.
- de Moura, L., & Bjørner, N. (2008). Z3: An Efficient SMT Solver. In *Tools and Algorithms for the Construction and Analysis of Systems*, Vol. 4963 of *Lecture Notes in Computer Science*, pp. 337–340. Springer Berlin Heidelberg.
- Diller, M., Haret, A., Linsbichler, T., Rümmele, S., & Woltran, S. (2018). An extension-based approach to belief revision in abstract argumentation. *Int. J. Approx. Reason.*, *93*, 395–423.
- Doutre, S., & Mailly, J. (2018). Constraints and changes: A survey of abstract argumentation dynamics. *Argument Comput.*, *9*(3), 223–248.
- Dung, P. M. (1995). On the acceptability of arguments and its fundamental role in non-monotonic reasoning, logic programming and n-person games. *Artificial Intelligence*, *77*(2), 321–358.
- Dung, P. M., & Thang, P. M. (2010). Towards (probabilistic) argumentation for jury-based dispute resolution. In Baroni, P., Cerutti, F., Giacomin, M., & Simari, G. R. (Eds.), *Proceedings of the 3rd International Conference on Computational Models of Argument (COMMA 2010)*, Vol. 216 of *Frontiers in Artificial Intelligence and Applications*, pp. 171–182. IOS Press.
- Dunne, P. E., Hunter, A., McBurney, P., Parsons, S., & Wooldridge, M. (2011). Weighted argument systems: Basic definitions, algorithms, and complexity results. *Artificial Intelligence*, *175*(2), 457–486.
- Dupin de Saint-Cyr, F., Bisquert, P., Cayrol, C., & Lagasquie-Schiex, M.-C. (2016). Argumentation update in yalla (yet another logic language for argumentation). *International Journal of Approximate Reasoning*, *75*, 57–92.
- Dvořák, W., & Dunne, P. E. (2018). *Computational problems in formal argumentation and their complexity*, pp. 631–687. College Publications.
- Dvořák, W. (2012). *Computational Aspects of Abstract Argumentation*. Ph.D. thesis, Technischen Universität Wien.

- Faqeh, R., Fetzner, C., Hermanns, H., Hoffmann, J., Klauck, M., Köhl, M. A., Steinmetz, M., & Weidenbach, C. (2020). Towards dynamic dependable systems through evidence-based continuous certification. In Margaria, T., & Steffen, B. (Eds.), *Proceedings of the 9th International Symposium on Leveraging Applications of Formal Methods, ISOFA 2020, Verification and Validation: Engineering Principles, Part II*, Vol. 12477 of *Lecture Notes in Computer Science*, pp. 416–439. Springer.
- Fazzinga, B., Flesca, S., & Parisi, F. (2015). On the complexity of probabilistic abstract argumentation frameworks. *ACM Transactions on Computational Logic*, 16(3), 22:1–22:39.
- Gaggl, S. A., Rudolph, S., & Straß, H. (2021). On the decomposition of abstract dialectical frameworks and the complexity of naive-based semantics. *Journal of Artificial Intelligence Research*, 70, 1–64.
- Gario, M., & Micheli, A. (2015). Pysmt: a solver-agnostic library for fast prototyping of smt-based algorithms. In *SMT Workshop 2015*.
- Hadoux, E., Hunter, A., & Polberg, S. (2021). Strategic argumentation dialogues for persuasion: Framework and experiments based on modelling the beliefs and concerns of the persuadee. *CoRR*, abs/2101.11870.
- Hunter, A. (2012). Some foundations for probabilistic abstract argumentation. In Verheij, B., Szeider, S., & Woltran, S. (Eds.), *Proceedings of the International Conference on Computational Models of Argument (COMMA 2012)*, Vol. 245 of *Frontiers in Artificial Intelligence and Applications*, pp. 117–128. IOS Press.
- Hunter, A. (2013). A probabilistic approach to modelling uncertain logical arguments. *International Journal of Approximate Reasoning*, 54(1), 47–81.
- Hunter, A., Polberg, S., & Thimm, M. (2020). Epistemic graphs for representing and reasoning with positive and negative influences of arguments. *Artificial Intelligence*, 281, 103236.
- Hunter, A., & Thimm, M. (2017). Probabilistic reasoning with abstract argumentation frameworks. *Journal of Artificial Intelligence Research*, 59, 565–611.
- Käfer, N. (2020). A general framework for probabilistic abstract argumentation. Master’s thesis, Saarland University.
- Li, H., Oren, N., & Norman, T. J. (2011). Probabilistic argumentation frameworks. In Modgil, S., Oren, N., & Toni, F. (Eds.), *Revised Selected Papers of the 1st International Workshop on Theory and Applications of Formal Argumentation (TAFIA 2011)*, Vol. 7132 of *Lecture Notes in Computer Science*, pp. 1–16. Springer.
- Modgil, S., & Prakken, H. (2018). Abstract Rule-Based Argumentation. In Baroni, P., Gabbay, D., Giacomin, M., & van der Torre, L. (Eds.), *Handbook of Formal Argumentation*, chap. 6. College Publications.
- Nouioua, F., & Risch, V. (2011). Argumentation frameworks with necessities. In Benferhat, S., & Grant, J. (Eds.), *Proceedings of the 5th International Conference on Scalable Uncertainty Management (SUM 2011)*, Vol. 6929 of *Lecture Notes in Computer Science*, pp. 163–176. Springer.

- Polberg, S., & Doder, D. (2014). Probabilistic abstract dialectical frameworks. In Fermé, E., & Leite, J. (Eds.), *Proceedings of the 14th European Conference on Logics in Artificial Intelligence (JELIA 2014)*, Vol. 8761 of *Lecture Notes in Computer Science*, pp. 591–599. Springer.
- Potyka, N. (2019). A polynomial-time fragment of epistemic probabilistic argumentation. *International Journal of Approximate Reasoning*, 115, 265–289.
- Rienstra, T., Thimm, M., Liao, B., & van der Torre, L. W. N. (2018). Probabilistic abstract argumentation based on SCC decomposability. In Thielscher, M., Toni, F., & Wolter, F. (Eds.), *Proceedings of the 16th International Conference on Principles of Knowledge Representation and Reasoning (KR 2018)*, pp. 168–177. AAAI Press.
- Straß, H., & Wallner, J. P. (2015). Analyzing the Computational Complexity of Abstract Dialectical Frameworks via Approximation Fixpoint Theory. *Artificial Intelligence*, 226, 34–74.
- Thimm, M. (2012). A probabilistic semantics for abstract argumentation. In Raedt, L. D., Bessiere, C., Dubois, D., Doherty, P., Frasconi, P., Heintz, F., & Lucas, P. J. F. (Eds.), *Proceedings of the 20th European Conference on Artificial Intelligence (ECAI 2012). Including Prestigious Applications of Artificial Intelligence (PAIS-2012) System Demonstrations Track*, Vol. 242 of *Frontiers in Artificial Intelligence and Applications*, pp. 750–755. IOS Press.
- Thimm, M., Baroni, P., Giacomin, M., & Vicig, P. (2017). Probabilities on extensions in abstract argumentation. In Black, E., Modgil, S., & Oren, N. (Eds.), *Proceedings of the 4th International Workshop on Theory and Applications of Formal Argumentation (TAFA 2017)*, Vol. 10757 of *Lecture Notes in Computer Science*, pp. 102–119. Springer.