Argument Strength in Probabilistic Argumentation using Confirmation Theory

Anthony Hunter[0000−0001−5602−7446]

Department of Computer Science, University College London, UK
anthony.hunter@ucl.ac.uk

Abstract. It is common for people to remark that a particular argument is a strong (or weak) argument. Having a handle on the relative strengths of arguments can help in deciding on which arguments to consider, and on which to present to others in a discussion. In computational models of argument, there is a need for a deeper understanding of argument strength. Our approach in this paper is to draw on confirmation theory for quantifying argument strength, and harness this in a framework based on probabilistic argumentation. We show how we can calculate strength based on the structure of the argument involving defeasible rules. The insights appear transferable to a variety of other structured argumentation systems.

Keywords: Argument strength · Probabilistic argumentation · Deductive argumentation · Defeasible logic.

1 Introduction

In real-world argumentation, it is common for arguments to be considered in terms of their strength. Yet in computational models of argument, we lack formalisms that adequately measure strength in arguments. Some variants of abstract argumentation touch on the notion of strength such as rankings (e.g. [6, 21, 1, 4, 2]), and probabilities (e.g. [11, 20, 34, 15, 30, 18]). However, these do not capture a notion of argument strength in terms of the quality of the contents of the premises and/or claim, rather they either assume that some kind of strength value is given for each argument and/or they calculate strength in terms of attacking and supporting arguments.

Using logical (i.e. structured) arguments allows the quality of the contents of the premises and claim to be directly considered. Some proposals assume strength is an input to the system (e.g. [8]). Others assess the strength of an argument in terms of the belief in it, often in terms of belief in the premises and claim (e.g. [14, 15, 29], and in terms of the conditional probability of the claim given the premises of the argument (e.g. [28, 36, 16]). So these draw on uncertainty in argumentation to quantify the strength. But as we shall see, these only give us an incomplete picture of the strength of an argument.

The use of defeasible logic is well-established in argumentation (see for example [27, 33, 13]), and key approaches to structured argumentation incorporate various kinds of defeasible rules [3]. There are also some proposals for probabilistic quantification of uncertainty in argumentation systems based on defeasible logic (e.g. [31, 11, 32]), but quantifying notions of strength have not been systematically considered in these formalisms.

To formalize argument strength for defeasible logic, we draw on measures from confirmation theory. These were originally proposed to determine the degree to which scientific evidence supports a hypothesis. They can capture how uncertainty associated with premises can impact a claim. To use these, we adapt the epistemic approach to probabilistic argumentation [34, 15, 18].

The proposal in this paper could be used in various ways. However, to illustrate and motivate, we focus on the following **audience scenario**: Someone presents us with an argument graph and the knowledgebase from which it has been constructed. The knowledge represents patterns that normally hold in the world (e.g., *if it is bird, then it is capable of flying*). As the audience, we are at liberty to identify a probability distribution over the possible worlds (in the following sections we make this precise) that represent our beliefs about the propositions in the language, and then we can use this probability distribution to analyze the strength of the arguments presented to us.

2 Defeasible logic

We assume a finite **set of atoms** $\mathcal{A}$. We form a **set of literals** $\mathcal{L}(\mathcal{A}) = \mathcal{A} \cup \{\neg\phi \mid \phi \in \mathcal{A}\}$. A **defeasible rule** is of the form $\psi_1 \wedge \ldots \wedge \psi_n \rightarrow \phi$ where $\psi_1, \ldots, \psi_n, \phi \in \mathcal{L}(\mathcal{A})$. For a rule ρ of the form $\psi_1 \wedge \ldots \wedge \psi_n \rightarrow \phi$, let $\mathsf{Tail}(\rho) = \{\psi_1, \ldots, \psi_n\}$ and $\mathsf{Head}(\rho) = \phi$. The **set of rules** is $\mathcal{R}(\mathcal{A})$ and the **set of formulae** is $\mathcal{F}(\mathcal{A}) = \mathcal{L}(\mathcal{A}) \cup \mathcal{R}(\mathcal{A})$. A **knowledgebase** is a subset of $\mathcal{F}(\mathcal{A})$.

Example 1. Consider $\mathtt{b}$ for *bird*, $\mathtt{p}$ for *penguin*, and $\mathtt{f}$ for *capable of flying*. Then $\Delta = \{\mathtt{b}, \mathtt{p}, \mathtt{b} \rightarrow \mathtt{f}, \mathtt{p} \rightarrow \neg\mathtt{f}\} \subseteq \mathcal{F}(\mathcal{A})$ is a knowledgebase.

Next, we present a variant of defeasible logic, incorporating *ex falso quodlibet*, to build arguments, and a semantics to analyze arguments.

Definition 1. *Let Δ be a knowledgebase and $\phi, \psi \in \mathcal{L}(\mathcal{A})$. The* **consequence relation***, denoted $\vdash$, is defined as follows: (1) $\Delta \vdash \phi$ if $\phi \in \Delta$; (2) $\Delta \vdash \phi$ if there is a $\psi_1 \wedge \ldots \wedge \psi_n \rightarrow \phi \in \Delta$ and $\Delta \vdash \psi_1$ and $\ldots$ and $\Delta \vdash \psi_n$; and (3) $\Delta \vdash \phi$ if $\Delta \vdash \psi$ and $\Delta \vdash \neg\psi$. Let* $\mathsf{Closure}(\Delta) = \{\phi \in \mathcal{A} \mid \Delta \vdash \phi\}$.

Example 2. For $\Delta = \{\mathtt{b}, \mathtt{b} \rightarrow \mathtt{f}\}$, $\mathsf{Closure}(\Delta) = \{\mathtt{b}, \mathtt{f}\}$.

The Tarskian properties (widely regarded as requirements for a logic) are satisfied (though for reflexivity, it is restricted to literals): (Reflexivity) $(\Delta \cap \mathcal{L}(\mathcal{A})) \subseteq \mathsf{Closure}(\Delta)$; (Monotonicity) $\mathsf{Closure}(\Delta) \subseteq \mathsf{Closure}(\Delta')$ if $\Delta \subseteq \Delta'$; and (Idempotency) $\mathsf{Closure}(\Delta) \subseteq \mathsf{Closure}(\mathsf{Closure}(\Delta))$.

A **model** is an assignment of true or false to the literals of the language. We represent each model by a subset of $\mathcal{A}$. The set of models of the language, denoted $\mathcal{M}(\mathcal{A})$, is the power set of $\mathcal{A}$. For $m \in \mathcal{M}(\mathcal{A})$, $\phi \in \mathcal{A}$, the **satisfaction**

relation, denoted $\models$, is: (1) $m \models \phi$ iff $\phi \in m$; And (2) $m \models \neg\phi$ iff $\phi \notin m$. We define the models for a set of formulae using the following fixpoint function.

Definition 2. *For $\Delta \subseteq \mathcal{F}(\mathcal{A})$, and $i \in \mathbb{N}$, the* **inference operators***, denoted* Inf^i*, are defined as:* $\mathsf{Inf}^1(\Delta) = \Delta \cap \mathcal{L}(\mathcal{A})$ *and* $\mathsf{Inf}^{i+1}(\Delta) = \mathsf{Inf}^i(\Delta) \cup \{\mathsf{Head}(\rho) \mid \rho \in \Delta \cap \mathcal{R}(\mathcal{A})$ *and for all* $\psi \in \mathsf{Tail}(\rho), \psi \in \mathsf{Inf}^i(\Delta)\}$*. Let* $\mathsf{Infer}(\Delta) = \mathsf{Inf}^k(\Delta)$ *where k is the smallest value s.t.* $\mathsf{Inf}^k(\Delta) = \mathsf{Inf}^{k+1}(\Delta)$*.*

Definition 3. *The* **satisfying models** *for $\Delta \subseteq \mathcal{F}(\mathcal{A})$, is* $\mathsf{Models}(\Delta) = \{m \in \mathcal{M}(\mathcal{A}) \mid m \models \phi$ *for all* $\phi \in \mathsf{Infer}(\Delta)\}$*.*

Example 3. For $\Delta = \{\mathtt{b}, \neg\mathtt{o}, \mathtt{b} \rightarrow \mathtt{f}, \mathtt{p} \rightarrow \neg\mathtt{f}\}$, where $\mathcal{A} = \{\mathtt{b}, \mathtt{p}, \mathtt{f}, \mathtt{o}\}$, $\mathsf{Infer}(\Delta) = \{\mathtt{b}, \neg\mathtt{o}, \mathtt{f}\}$, and $\mathsf{Models}(\Delta) = \{\{\mathtt{b}, \mathtt{f}\}, \{\mathtt{b}, \mathtt{p}, \mathtt{f}\}\}$.

Definition 4. *For $\Delta \subseteq \mathcal{F}(\mathcal{A})$, $\phi \in \mathcal{L}(\mathcal{A})$, the* **entailment relation** *holds, denoted $\Delta \models \phi$, iff* $\mathsf{Models}(\Delta) \subseteq \mathsf{Models}(\phi)$*.*

Example 4. For knowledgebase $\Delta = \{\mathtt{b}, \mathtt{p}, \mathtt{b} \rightarrow \mathtt{f}\}$, $\Delta \models \mathtt{f}$.

A knowledgebase is consistent iff it does not imply an atom and its negation. So Example 1 (respectively Example 4) is inconsistent (respectively consistent). Obviously, Δ is consistent iff $\mathsf{Models}(\Delta) \neq \emptyset$.

Proposition 1. *For $\Delta \subseteq \mathcal{F}(\mathcal{A})$, $\phi \in \mathcal{L}(\mathcal{A})$, $\Delta \vdash \phi$ iff $\Delta \models \phi$.*

When Δ is consistent, this correctness result can be shown via the notion of a proof tree where ϕ is at the root, each leaf is a literal in Δ, and each non-leaf node ϕ' is such that there is a rule $\psi_1 \wedge \ldots \wedge \psi_n \rightarrow \phi' \in \Delta$ and each $\psi_i \in \{\psi_1, \ldots, \psi_n\}$ is a child of ϕ'. So $\Delta \vdash \phi$ holds iff there is such a tree where each branch is finite. We can use the same tree to consider entailment. So $\Delta \models \phi$ holds iff the leaves and inferences are satisfied by all the models of Δ, and hence all these models satisfy ϕ. When Δ is inconsistent, the consequence relation entails any literal (Proof rule 3), and similarly the entailment relation for defeasible logic is trivializable in the sense that any literal is entailed by inconsistency. This is because when Δ is inconsistent, $\mathsf{Models}(\Delta) = \emptyset$, and therefore for any $\beta \in \mathcal{L}$, $\mathsf{Models}(\Delta) \subseteq \mathsf{Models}(\beta)$.

3 Probabilistic argumentation

In this section, we adapt the epistemic approach to probabilistic argumentation for use with defeasible logic. To present models, we use a **signature**, denoted $\mathcal{S}$, which is the atoms of the language $\mathcal{L}$ given in a sequence $\langle a_1, ..., a_n\rangle$, and then each model $m \in \mathcal{M}(\mathcal{A})$ is a binary number $b_1...b_n$ where for each digit b_i, if b_i is 1, then a_i is true in m, and if b_i is 0, then a_i is false in m. For example, for $\mathcal{S} = \langle \mathtt{a}, \mathtt{b}, \mathtt{c}\rangle$, $\mathcal{M}(\mathcal{A})$ is $\{111, 110, 101, 100, 011, 010, 001, 000\}$. So for $m = 101$, $\mathtt{a}$ is true, $\mathtt{b}$ is false, and $\mathtt{c}$ is true.

Definition 5. *A* **probability distribution** *P over $\mathcal{M}(\mathcal{A})$ is a function $P : \mathcal{M}(\mathcal{A}) \rightarrow [0, 1]$ s.t. $\sum_{m \in \mathcal{M}(\mathcal{A})} P(m) = 1$.*

Definition 6. *The* **probability of literal** $\phi \in \mathcal{L}(\mathcal{A})$ *w.r.t. probability distribution* P *is* $P(\phi) = \sum_{m \in \mathcal{M}(\mathcal{A}) \text{ s.t. } m \models \phi} P(m)$.

Next, for an argument: (1) the premises imply the claim; (2) the premises are consistent; and (3) the premises are minimal for entailing the claim.

Definition 7. *For* $\Phi \subseteq \mathcal{F}(\mathcal{A})$*, and* $\alpha \in \mathcal{L}(\mathcal{A})$*,* $\langle \Phi, \alpha \rangle$ *is an* **argument** *iff (1)* $\Phi \vdash \alpha$*; (2)* Φ *is consistent; and (3) there is no* $\Phi' \subset \Phi$ *such that* $\Phi' \vdash \alpha$.

Example 5. For $\Delta = \{\mathtt{b}, \mathtt{p}, \mathtt{b} \rightarrow \mathtt{f}, \mathtt{p} \rightarrow \neg\mathtt{f}\}$, the arguments are $\langle \{\mathtt{b}, \mathtt{b} \rightarrow \mathtt{f}\}, \mathtt{f} \rangle$, $\langle \{\mathtt{p}, \mathtt{p} \rightarrow \neg\mathtt{f}\}, \neg\mathtt{f} \rangle$, $\langle \{\mathtt{b}\}, \mathtt{b} \rangle$, and $\langle \{\mathtt{p}\}, \mathtt{p} \rangle$.

A **reflexive** argument is of the form $A = \langle \{\alpha\}, \alpha \rangle$. For argument $A = \langle \Phi, \alpha \rangle$, $\mathsf{Support}(A)$ returns Φ, $\mathsf{Claim}(A)$ returns α, $\mathsf{Facts}(A)$ returns the literals in Φ, $\mathsf{Rules}(A)$ returns the rules in Φ, and $\mathsf{Frame}(A)$ returns $\mathsf{Facts}(A) \cup \{\mathsf{Claim}(A)\}$.

The probability of an argument being acceptable is based on the facts in the premises and the claim of the argument. So returning to the audience scenario, if the audience has been presented with an argument graph, they can use their probability distribution to calculate the probability of each argument.

Definition 8. *The* **probability of argument** A *being acceptable is denoted* $P(A)$*, where* $P(A) = \sum_{m \in \mathsf{Models}(\mathsf{Frame}(A))} P(m)$.

Example 6. Continuing Ex. 1, with probability distribution P below, and signature $\mathcal{S} = \langle \mathtt{b}, \mathtt{p}, \mathtt{f} \rangle$, $P(\langle \{\mathtt{b}, \mathtt{b} \rightarrow \mathtt{f}\}, \mathtt{f} \rangle) = 0.95$ and $P(\langle \{\mathtt{p}, \mathtt{p} \rightarrow \neg\mathtt{f}\}, \neg\mathtt{f} \rangle) = 0.01$.

$\langle \mathtt{b}, \mathtt{p}, \mathtt{f} \rangle$	110	101	100
P	0.01	0.95	0.04

Definition 9. *For* $\phi, \phi' \in \mathcal{L}(\mathcal{A})$*, let* $\phi \equiv \phi'$ *denote* $\mathsf{Models}(\phi) = \mathsf{Models}(\phi')$ *and let* $\neg\neg\phi = \phi$*. For arguments* A *and* B*,* A *is a* **direct undercut** *of* B *if there is* $\phi \in \mathsf{Support}(B)$ *s.t.* $\mathsf{Claim}(A) \equiv \neg\phi$*.* A *is a* **rebuttal** *of* B *if* $\mathsf{Claim}(A) \equiv \neg\mathsf{Claim}(B)$*.* A **attacks** B *iff* A *is a direct undercut or* A *is a rebuttal of* B*.*

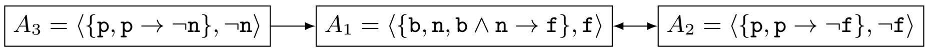
The following coherence property holds because the sum of belief in complementary literals is less than or equal to 1.

Proposition 2. *For probability distribution* P*, if* B *attacks* A*, then* $P(A) + P(B) \leq 1$.

We use the usual notion of an argument graph, where each node is an argument, and each arc denotes an attack by one argument on another [10]. For a knowledgebase, a **complete argument graph** contains all arguments and attacks (e.g. Figure 1). However, we are not using the argument graph to determine which arguments are acceptable. Rather, we use a probability distribution to determine the acceptable arguments as explained below. So the role of the argument graph is to provide a presentation of the arguments.

When $P(A) > 0.5$, then the argument is believed to be acceptable, whereas when $P(A) \leq 0.5$, then it is not believed to be acceptable. The epistemic extension for a graph G, denoted $\mathsf{Extension}(P, G)$, is the set of arguments that are

Fig. 1. The following is a complete argument graph (excluding reflexive arguments) from $\Delta = \{\mathtt{b}, \mathtt{p}, \mathtt{b} \land \mathtt{n} \to \mathtt{f}, \mathtt{p} \to \neg\mathtt{f}, \mathtt{p} \to \neg\mathtt{n}\}$ where $\mathtt{b}$ is *bird*, $\mathtt{p}$ is *penguin*, $\mathtt{f}$ is *capable of flying*, and $\mathtt{n}$ is a normality atom. Argument A_3 has the claim that negates $\mathtt{n}$, and so represents an attack on the use of the rule with $\mathtt{n}$ as condition in A_1.



believed to be acceptable (i.e. $A \in \mathsf{Extension}(P, G)$ iff A is in G and $P(A) > 0.5$). For example, for graph G in Figure 1, with $P(A_1) = 0.95$, and $P(A_2) = P(A_3) = 0.15$, $\mathsf{Extension}(P, G) = \{A_1\}$. As shown in [15], for any probability distribution P and graph G, $\mathsf{Extension}(P, G)$ is conflict-free.

The epistemic approach provides a finer-grained assessment of an argument graph than given by the definitions for Dung's extensions. By adopting constraints on the distribution, the epistemic approach subsumes Dung's definitions [34, 18, 26]. The epistemic approach also provides alternatives to Dung's approach. For instance, we may wish to represent disbelief in arguments even when they are unattacked [25].

4 Modelling normality

Defeasible rules are normally correct, but sometimes are incorrect, and so we need to attack them. To address this, we use normality atoms as illustrated in Figure 1. We assume the set of atoms $\mathcal{A}$ is partitioned into **normality atoms**, denoted $\mathcal{N}$, and **ordinary atoms**, denoted $\mathcal{Q}$. So $\mathcal{A} = \mathcal{Q} \cup \mathcal{N}$ and $\mathcal{Q} \cap \mathcal{N} = \emptyset$. We read the normality atom in the condition as saying that the context for applying the rule is normal. If there are reasons to believe that it is not a normal context, then a counterargument attacks this assumption of normality.

We will use the following **normality modelling convention**: Each rule has at most one normality atom as a condition, and if a normality atom appears as condition in a rule, it is unique to that rule. No other rule in the knowledgebase has the same normality atom as a condition. However, multiple rules in the knowledgebase can have the same negated normality atom as head. This convention helps us to specify an appropriate probability distribution over a set of atoms that includes normality atoms. We quantify the probability of a normality atom in terms of the unique rule that contains it as an antecedent: For a defeasible rule $\beta_1 \land \ldots \land \beta_n \land \gamma \to \alpha$ with normality atom γ, the probability of γ is $P(\beta_1 \land \ldots \land \beta_n \land \alpha)$.

Example 7. Consider probability distribution P, with the rules in Δ being $\mathtt{b} \land \mathtt{n}_1 \to \mathtt{f}$ and $\mathtt{p} \land \mathtt{n}_2 \to \neg\mathtt{f}$. So P satisfies our constraints that $P(\mathtt{n}_1) = P(\mathtt{b} \land \mathtt{f})$ and $P(\mathtt{n}_2) = P(\mathtt{p} \land \neg\mathtt{f})$.

$\langle \mathtt{b}, \mathtt{p}, \mathtt{f}, \mathtt{n}_1, \mathtt{n}_2 \rangle$	11001	10110	10000
P	0.01	0.95	0.04

The use of normality propositions to disable rules is analogous to the use of abnormality predicates in formalisms such as circumscription [22]. Furthermore, we can use normality atoms to capture the specificity principle where a more specific rule is preferred over a less specific rule.

5 Argument strength

Given an argument A and probability distribution P, we let $S_P(A)$ be an assignment in the $[-1, 1]$ interval to denote the strength of A. If $S_P(A) \leq 0$, then the support of the argument does not provide a good reason for the claim. As $S_P(A)$ rises above zero, then the support of the argument gives an increasingly good reason for the claim.

In the rest of this paper, we will focus on arguments that involve relationships between observations. We will assume that all the atoms, apart from the normality atoms, are observations. These are atoms that can ultimately be verified as true or false, though at any specific time, there may be uncertainty about which observations are true or false. Examples of observations include $\mathtt{b}$ = *bird*, $\mathtt{d}$ = *duck*, $\mathtt{p}$ = *penguin*, $\mathtt{e}$ = *eagle*, and $\mathtt{f}$ = *fly-thing*. Given an argument A, the **evidence** in A, denoted $\mathsf{Ev}(A)$, is the set of observations in the support of A that are not normality atoms. (i.e. $\mathsf{Ev}(A) = \mathsf{Facts}(A) \cap \mathcal{L}(\mathcal{Q})$). So when we investigate strength, we want to quantify some aspect of how believing the evidence in the premises supports the claim.

There are various ways of defining S_P. To clarify some of the issues consider the simple situation of an argument A of the form $\langle \{\mathtt{b}, \mathtt{b} \rightarrow \mathtt{a}\}, \mathtt{a} \rangle$. For this, consider the four models for signature $\langle \mathtt{a}, \mathtt{b} \rangle$. Mass on 11 indicates positive correlation between $\mathtt{a}$ and $\mathtt{b}$, and mass on 10 and 01 indicates negative correlation. On this basis, the conditional probabilities $P(\mathtt{a}|\mathtt{b})$ and $P(\mathtt{b}|\mathtt{a})$ indicate positive correlation, and the conditional probabilities $P(\mathtt{a}|\neg\mathtt{b})$, $P(\neg\mathtt{a}|\mathtt{b})$, $P(\mathtt{b}|\neg\mathtt{a})$, and $P(\neg\mathtt{b}|\mathtt{a})$, indicate negative correlation.

So there are multiple dimensions to connecting the evidence and claim. We will draw some of these out in the following seven properties that capture desirable, though not mandatory, features of argument strength. For a set of literals Γ, $\wedge\Gamma$ is the conjunction of the literals. We extend the definition for the probability of a literal so that if λ is a Boolean combination of literals in $\mathcal{L}(\mathcal{A})$, then $P(\lambda)$ is the sum of the probability of the models that classically satisfy λ. Also $P(\lambda|\lambda')$ is $P(\lambda \wedge \lambda')/P(\lambda')$, and if $P(\lambda) = 0$, then we let $P(\lambda|\lambda') = 0$.

$(X1)$ If $\mathsf{Claim}(A) \equiv \wedge\mathsf{Ev}(A)$ and $P(\mathsf{Claim}(A)) > 0,$ then $S(A) < 1$
$(X2)$ If $\wedge\mathsf{Ev}(A) \equiv \wedge\mathsf{Ev}(B)$ and $\mathsf{Claim}(A) = \mathsf{Claim}(B),$ then $S(A) = S(B)$
$(X3)$ If $S(A) > 0,$ then $P(\mathsf{Claim}(A)| \wedge\mathsf{Ev}(A)) > 0$
$(X4)$ If $P(\mathsf{Claim}(A)| \wedge\mathsf{Ev}(A)) = P(\mathsf{Claim}(A),$ then $S(A) = 0$
$(X5)$ If $P(\wedge\mathsf{Ev}(A)|\mathsf{Claim}(A)) = P(\wedge\mathsf{Ev}(A)|\neg\mathsf{Claim}(A)),$ then $S(A) = 0$
$(X6)$ If P is a uniform probability distribution, then $S(A) = 0$
$(X7)$ If $P(\mathsf{Claim}(A)| \wedge\mathsf{Ev}(A)) = P(\mathsf{Claim}(A)|\neg \wedge\mathsf{Ev}(A)),$ then $S(A) = 0$

We explain these properties as follows: (X1) If the evidence and claim are logically equivalent, and there is non-zero belief in the claim, then the argument is not providing a good reason for the claim because the reason is just the claim reiterated, and hence the strength is below 1; (X2) If two arguments are logically equivalent with respect to support and claim, then they provide equivalent reasons for equivalent claims, and so they are equally strong; (X3) Positive strength requires the probability of the claim conditional on the evidence to be non-zero; (X4) If the probability of the claim conditional on the evidence equals the probability of the claim, then the argument has zero strength as the premises are not giving a useful reason for the claim; (X5) If the probability of the evidence conditional on the claim equals the probability of the evidence conditional on the negation of the claim, then the argument has zero strength; (X6) If there is a uniform distribution, there is no material relationship between the evidence and claim and so the argument has zero strength; And (X7) If the probability of the claim conditional on the evidence equals the probability of the claim conditional on the negation of the evidence, then the argument has zero strength.

The first strength function we consider is the plausibility strength function which is the probability of the claim conditional on the evidence.

Definition 10. *The* **plausibility** *of an argument A w.r.t. probability distribution P, denoted S^p_P, is $P(\mathsf{Claim}(A)| \wedge \mathsf{Ev}(A))$.*

Example 8. Continuing Example 7, the plausibility of argument $A = \langle\{\mathtt{b}, \mathtt{b} \wedge \mathtt{n_1} \rightarrow \mathtt{f}\}, \mathtt{f}\rangle$ is $S^p_P(A) = 0.95$.

Proposition 3. *The plausibility measure S^p_P satisfies $X2$ and $X3$, but it does not satisfy $X1$, $X4$, $X5$, $X6$, or $X7$.*

Proof. For counterexamples for X4, X5 and X7, consider the signature $\langle \mathtt{a}, \mathtt{b}\rangle$ with argument $\langle\{\mathtt{b}, \mathtt{b} \rightarrow \mathtt{a}\}, \mathtt{a}\rangle$. (X1) Assume $\mathsf{Claim}(A) \equiv \wedge\mathsf{Ev}(A)$ and $P(\mathsf{Claim}(A)) > 0$. So $P(\mathsf{Claim}(A) \wedge \mathsf{Ev}(A))/P(\mathsf{Claim}(A)) = 1$. So $S^p_P(A) \not< 1$. (X2 and X3) Direct from definition. (X4) Let $P(11) = 0.2$, $P(10) = 0.3$, $P(01) = 0.2$, and $P(00) = 0.3$. (X5) Let $P(11) = 0.2$, $P(10) = 0.2$, $P(01) = 0.3$, and $P(00) = 0.3$. (X6) If P is uniform, then $P(\mathsf{Claim}(A) \wedge \mathsf{Ev}(A))|P(\wedge\mathsf{Ev}(A)) \neq 0$. So $S^p_P(A) \neq 0$. (X7) Let $P(11) = 0.2$, $P(10) = 0.3$, $P(01) = 0.2$, and $P(00) = 0.3$.

The failure of the $X1$, $X4$, $X5$, $X6$, and $X7$ properties suggests that there may be useful alternatives to the plausibility for measuring the strength of an argument. We are not suggesting that there is a single measure that tells us everything we need to know about the strength of a probabilistic argument based on defeasible rules, but we do expect that different measures can tell us different useful things about the strength of arguments.

6 Confirmation theory

For an alternative perspective on the strength of an argument, we turn to confirmation measures. Originally, confirmation measures were developed in the philosophy of science to investigate the development of scientific hypotheses [9]. The aim of a confirmation measure $C(E, H)$ is to capture the degree to which evidence E supports hypothesis H. Confirmation measures have been proposed as a measure of argument strength in [24] but only in the restricted context of an argument that is a conditional probability statement. For our purposes, we assume that the evidence is a set of literals Δ and the hypothesis is a literal ϕ. We review some well-known confirmation measures next.

Definition 11. *For $\Delta \subseteq \mathcal{L}(\mathcal{A})$, $\phi \in \mathcal{L}(\mathcal{A})$, and probability distribution P, the C^d_P [5], C^s_P [7], and C^k_P [19] confirmation measures are defined as follows.*

- $C^d_P(\Delta, \phi) = P(\phi| \wedge\Delta) - P(\phi)$
- $C^s_P(\Delta, \phi) = P(\phi| \wedge\Delta) - P(\phi|\neg(\wedge\Delta))$
- $C^k_P(\Delta, \phi) = \frac{P(\wedge\Delta|\phi) - P(\wedge\Delta|\neg\phi)}{P(\wedge\Delta|\phi) + P(\wedge\Delta|\neg\phi)}$ *when* $P(\wedge\Delta|\phi)+P(\wedge\Delta|\neg\phi) > 0$ *and 0 otherwise.*

We explain these measures as follows: (C^d_P) the increase in belief in the claim that can be attributed to believing the evidence to be true, i.e. for it to be positive, $P(\phi) < P(\phi \mid \wedge\Delta)$ holds; (C^s_P) the difference in belief in the claim conditioned on the evidence being true and belief in the claim conditioned on the evidence being untrue, i.e. for it to be positive, $P(\phi \mid \neg(\wedge\Delta)) < P(\phi \mid \wedge\Delta)$ holds; and (C^k_P) the difference in belief in the evidence conditioned on the claim being true and belief in the evidence conditioned on the claim being untrue, normalized by the maximum range for the value, i.e. for it to be positive, $P(\wedge\Delta \mid \neg\phi) < P(\wedge\Delta \mid \phi)$ holds.

Example 9. Consider $S = \langle \mathtt{a}, \mathtt{b} \rangle$ with the following the distribution (left) and strength measures (right) for $\Delta = \{\mathtt{b}\}$ and $\phi = \mathtt{a}$. Here, the conditional probability gives a quite high score, whereas the confirmation measures give lower scores, reflecting the mass assigned to the models 10 and 01.

$\langle \mathtt{a}, \mathtt{b} \rangle$	11	10	01	00
P	0.5	0.1	0.2	0.2

	$P(.)$	S^p_P	S^d_P	S^s_P	S^k_P
A	0.5	0.71	0.11	0.38	0.25

We can harness confirmation theory for argumentation (where arguments concern relationships between observations) as below, where the greater the value, the stronger the argument. In order to focus on the evidence, we consider the ordinary facts (i.e. observations) in the support of the argument.

Definition 12. *The* **confirmation strength** *of argument A w.r.t. probability distribution P and confirmation measure C^x_P, for $x \in \{d, s, k\}$, denoted $S^x_P(A)$, is $S^x_P(A) = C^x_P(\mathsf{Ev}(A), \mathsf{Claim}(A))$.*

Example 10. Consider the following probability distribution (left), with signature $\langle \mathtt{d}, \mathtt{p}, \mathtt{f} \rangle$, for an insectarium, where $\mathtt{d}$ is *dragonfly*, $\mathtt{p}$ is *pollinator*, and $\mathtt{f}$ is *flying insect*.

$\langle \mathtt{d}, \mathtt{p}, \mathtt{f} \rangle$	101	011
P	0.2	0.8

	$P(.)$	S^p_P	S^d_P	S^s_P	S^k_P
A_1	0.00	0.00	-0.80	-1.00	-1.00
A_2	0.20	1.00	0.80	1.00	1.00

For arguments $A_1 = \langle\{\mathtt{d}, \mathtt{d} \rightarrow \mathtt{p}\}, \mathtt{p}\rangle$ and $A_2 = \langle\{\mathtt{d}, \mathtt{d} \rightarrow \neg\mathtt{p}\}, \neg\mathtt{p}\rangle$, the values for argument strength are given in the table on the right above where $P(.)$ is the probability of the argument being acceptable. Since dragonflies are not pollinators (as shown by the probability distribution), A_1 has low scores, and A_2 has high scores, for confirmation strength.

Example 11. Consider the probability distribution P, with signature $\langle\mathtt{b}, \mathtt{p}, \mathtt{f}, \mathtt{n}\rangle$, where $\mathtt{b}$ is *bird*, $\mathtt{p}$ is *penguin*, $\mathtt{f}$ is *capable of flying*, and $\mathcal{N} = \{\mathtt{n}\}$, for a zoo with a large aviary.

$\langle\mathtt{b}, \mathtt{p}, \mathtt{f}, \mathtt{n}\rangle$	1100	1011	1000	0000
P	0.01	0.75	0.04	0.2

	$P(.)$	S^p_P	S^d_P	S^s_P	S^k_P
A_1	0.75	0.94	0.19	0.94	0.67
A_2	0.01	1.00	0.75	0.76	1.00
A_3	0.01	1.00	0.75	0.76	1.00

For the arguments $A_1 = \langle\mathtt{b}, \mathtt{b} \wedge \mathtt{n} \rightarrow \mathtt{f}\}, \mathtt{f}\rangle$, $A_2 = \langle\mathtt{p}, \mathtt{p} \rightarrow \neg\mathtt{f}\}, \neg\mathtt{f}\rangle$, and $A_3 = \langle\{\mathtt{p}, \mathtt{p} \rightarrow \neg\mathtt{n}\}, \neg\mathtt{n}\rangle$, the strengths are given in the table on the above right. So A_1 has a high probability of acceptability, and some good scores for confirmation strength, and A_2 and A_3 have a low probability of acceptability but are quite strong arguments.

Proposition 4. *The table captures satisfaction ✓, or non-satisfaction ×, of the X1 to X7 properties.*

	X1	X2	X3	X4	X5	X6	X7
S^d_P	✓	✓	✓	✓	✓	✓	✓
S^s_P	×	✓	✓	×	✓	✓	✓
S^k_P	×	✓	✓	×	✓	✓	✓

Proof. For counterexamples, consider the signature $\langle\mathtt{a}, \mathtt{b}\rangle$ with $\langle\{\mathtt{a}, \mathtt{a} \rightarrow \mathtt{a}\}, \mathtt{a}\rangle$ for X1, and $\langle\{\mathtt{b}, \mathtt{b} \rightarrow \mathtt{a}\}, \mathtt{a}\rangle$ for X4 and X7. (S^d_P) (X1) Assume $\wedge\mathsf{Cn}(\mathsf{Ev}(A)) = \wedge\mathsf{Cn}(\mathsf{Claim}(A)) = \phi$. So $S(A) = P(\phi|\phi) - P(\phi) = 1 - P(\phi)$. So if $P(\phi) > 0$, then $S(A) < 1$. (X2) Direct from definition. (X3) If $S^d_P(A) > 0$, then $P(\mathsf{Claim}(A)| \wedge \mathsf{Ev}(A)) - P(\mathsf{Claim}(A)) > 0$. So $P(\mathsf{Claim}(A)|\wedge\mathsf{Ev}(A)) > 0$. (X4) Direct from definition. (X5) From the assumption $P(\wedge\mathsf{Ev}(A)|\mathsf{Claim}(A)) = P(\wedge\mathsf{Ev}(A)|\neg\mathsf{Claim}(A))$, we can show $P(\mathsf{Claim}(A)| \wedge \mathsf{Ev}(A)) = P(\mathsf{Claim}(A))$. (X6) If P is uniform, then $P(\mathsf{Claim}(A)| \wedge \mathsf{Ev}(A)) = 0.5$ and $P(\mathsf{Claim}(A)) = 0.5$. So $S^d_P(A) = 0$. (X7) From the assumption $P(\mathsf{Claim}(A)| \wedge \mathsf{Ev}(A)) = P(\mathsf{Claim}(A)|\neg \wedge \mathsf{Ev}(A))$ we can show $P(\mathsf{Claim}(A)| \wedge \mathsf{Ev}(A)) = P(\mathsf{Claim}(A))$. ($S^s_P$) (X1) Let $P(11) = 0.5$ and $P(00) = 0.5$. (X2) Direct from definition. (X3) If $S^s_P(A) > 0$, then $P(\mathsf{Claim}(A)|\wedge\mathsf{Ev}(A)) - P(\mathsf{Claim}(A)|\neg \wedge \mathsf{Ev}(A)) > 0$. So $P(\mathsf{Claim}(A)| \wedge \mathsf{Ev}(A)) > 0$. (X4) Let $P(11) = 1$. (X5) Follows from the fact that $P(\mathsf{Claim}(A)| \wedge \mathsf{Ev}(A)) = P(\mathsf{Claim}(A)|\neg \wedge \mathsf{Ev}(A))$ holds iff $P(\wedge\mathsf{Ev}(A)|\mathsf{Claim}(A)) = P(\wedge\mathsf{Ev}(A)|\neg\mathsf{Claim}(A))$ holds. (X6) If P is uniform, then $P(\mathsf{Claim}(A)| \wedge \mathsf{Ev}(A)) = 0.5$ and $P(\mathsf{Claim}(A)|\neg \wedge \mathsf{Ev}(A)) = 0.5$. So $S^s_P(A) = 0$. (X7) Direct from definition. (S^k_P) (X1) Let $P(11) = 0.5$ and $P(00) = 0.5$. (X2) Direct from defn. (X3) If $S^p_P(A) > 0$, then $P(\mathsf{Claim}(A)|\wedge\mathsf{Ev}(A)) > 0$. So $P(\mathsf{Ev}(A)|\mathsf{Claim}(A)) > 0$. (X4) Let $P(11) = 1$. (X5) Direct from definition. (X6) If P is uniform, then then $P(\wedge\mathsf{Ev}(A) \wedge \mathsf{Claim}(A)) = P(\wedge\mathsf{Ev}(A))/2$ and $P(\wedge\mathsf{Ev}(A) \wedge$

$\neg\mathsf{Claim}(A)) = P(\wedge\mathsf{Ev}(A))/2$. Hence $P(\wedge\mathsf{Ev}(A)|\mathsf{Claim}(A)) = ((P(\wedge\mathsf{Ev}(A))/2)/0.5 = P(\wedge\mathsf{Ev}(A))$ and $P(\wedge\mathsf{Ev}(A)|\neg\mathsf{Claim}(A)) = ((P(\wedge\mathsf{Ev}(A))/2)/0.5 = P(\wedge\mathsf{Ev}(A))$. So $\frac{P(\wedge\mathsf{Ev}(A)|\mathsf{Claim}(A))-P(\wedge\mathsf{Ev}(A)|\neg\mathsf{Claim}(A)}{P(\wedge\mathsf{Ev}(A)|\mathsf{Claim}(A))+P(\wedge\mathsf{Ev}(A)|\neg\mathsf{Claim}(A)} = \frac{P(\wedge\mathsf{Ev}(A))-P(\wedge\mathsf{Ev}(A))}{P(\wedge\mathsf{Ev}(A))+P(\wedge\mathsf{Ev}(A))} = 0$ (X7) Same as for S^s_P $X5$.

Recall that the plausibility measure S^p_P only satisfies $X2$ and $X3$. So S^d_P, S^s_P and S^k_P offer potentially valuable alternatives to S^p_P as they satisfy most/all of the properties. Note, four of these properties concern what zero strength means, and so in future work, we will consider further properties to explore what positive and negative strength mean, and to differentiate the measures.

7 Multiple defeasible rules

We now consider the question of the strength of an argument with multiple defeasible rules in the premises. To illustrate some of our concerns, consider the following arguments where `b` denotes *bird*, `w` denotes *has wings*, `y` denotes *yellow*, `f` denotes *capable of flying*.

- $A_1 = \langle\{\mathtt{b}, \mathtt{b} \rightarrow \mathtt{f}\}, \mathtt{f}\rangle$
- $A_2 = \langle\{\mathtt{b}, \mathtt{b} \rightarrow \mathtt{w}, \mathtt{w} \rightarrow \mathtt{f}\}, \mathtt{f}\rangle$
- $A_3 = \langle\{\mathtt{b}, \mathtt{b} \rightarrow \mathtt{y}, \mathtt{y} \rightarrow \mathtt{f}\}, \mathtt{f}\rangle$

Intuitively, A_1 is a reasonably strong argument since most birds have the capability to fly. But does A_2 have the same strength as A_1 since it starts from the same fact (i.e. bird) or is it stronger because it makes the intermediate point concerning having wings? And does A_3 have the same strength as A_1 or is it weaker because it makes the intermediate point (i.e. being yellow) that is irrelevant (and unlikely to be correct)? Assuming the probability distribution over the atoms `b` and `f` is the same for each argument, then the strength of each argument is the same since it is based on `b` and `f`. However, taking the rules into account, we might expect the following: (A_2) A strong confirmation by *birds* for *has wings*, and by *has wings* for *capable of flying*; and (A_3) A weak confirmation by *birds* for *yellow*, and a weak confirmation by *yellow* for *capable of flying*. To capture this, we consider how the assessment of the strength of an argument can depend on its intermediate steps.

Definition 13. *Argument B is an* **intermediate** *of argument A iff* $\mathsf{Rules}(B) \subseteq \mathsf{Rules}(A)$. *Let* $\mathsf{Intermediates}(A) = \{B \mid B$ *is an intermediate of* $A\}$.

Example 12. For $A_1 = \langle\{\mathtt{b}, \mathtt{b} \rightarrow \mathtt{w}, \mathtt{w} \rightarrow \mathtt{f}\}, \mathtt{f}\rangle$, the intermediates are A_1, $B_1 = \langle\{\mathtt{b}, \mathtt{b} \rightarrow \mathtt{w}\}, \mathtt{w}\rangle$, and $B_2 = \langle\{\mathtt{w}, \mathtt{w} \rightarrow \mathtt{f}\}, \mathtt{f}\rangle$,

If B is a strict intermediate of A (i.e. $\mathsf{Rules}(B) \subset \mathsf{Rules}(A)$), and $\mathsf{Claim}(B) \neq \mathsf{Claim}(A)$, then there is defeasible rule $\beta_1 \wedge \ldots \wedge \beta_n \rightarrow \phi \in \mathsf{Support}(A)$ where $\mathsf{Claim}(B) \in \{\beta_1, \ldots, \beta_n\}$ (e.g. B_1 in Example 12). This is because arguments are minimal, and so if the claim of the intermediate differs from that of the argument, then it is also a condition in a defeasible rule. Also, if B is a strict intermediate of A, it is not necessarily the case that $\mathsf{Support}(B) \subset \mathsf{Support}(A)$ (e.g. B_2 in Example 12). In order to consider the intermediates in the derivation of a claim from its premises, we use the following definition that judges not just the argument but also the intermediates.

Definition 14. *An argument A is* **compositionally strong** *with respect to strength measure S_P and threshold $\tau \in [-1, 1]$ iff for all $B \in \mathsf{Intermediates}(A)$, $S_P(B) \geq \tau$.*

We now return to the arguments A_2 and A_3 from the introduction of this section, and analyze them in the following two examples.

Example 13. For $A_2 = \langle\{\mathtt{b}, \mathtt{b} \to \mathtt{w}, \mathtt{w} \to \mathtt{f}\}, \mathtt{f}\rangle$, with the following probability distribution for a zoo, where $\mathcal{S} = \langle \mathtt{b}, \mathtt{w}, \mathtt{f}\rangle$, then A_2 and its strict intermediates have high strength.

$\langle \mathtt{b}, \mathtt{w}, \mathtt{f}\rangle$	111	110	001	000
P	0.09	0.01	0.02	0.88

	S^p_P	S^d_P	S^s_P	S^k_P
$\langle\{\mathtt{b}, \mathtt{b} \to \mathtt{w}, \mathtt{w} \to \mathtt{f}\}, \mathtt{f}\rangle$	0.90	0.79	0.88	0.97
$\langle\{\mathtt{b}, \mathtt{b} \to \mathtt{w}\}, \mathtt{w}\rangle$	1.00	0.90	1.00	1.00
$\langle\{\mathtt{w}, \mathtt{w} \to \mathtt{f}\}, \mathtt{f}\rangle$	0.90	0.79	0.88	0.97

Example 14. For $A_3 = \langle\{\mathtt{b}, \mathtt{b} \to \mathtt{y}, \mathtt{y} \to \mathtt{f}\}, \mathtt{f}\rangle$, with the following probability distribution for a zoo, where $\mathcal{S} = \langle \mathtt{b}, \mathtt{y}, \mathtt{f}\rangle$, then A_3 has the same high strength as A_2 in the previous example (because the marginals involving $\mathtt{b}$ and $\mathtt{f}$ are the same) but low strength for the strict intermediates.

$\langle \mathtt{b}, \mathtt{y}, \mathtt{f}\rangle$	111	101	100	010	001	000
P	0.01	0.08	0.01	0.02	0.02	0.86

	S^p_P	S^d_P	S^s_P	S^k_P
$\langle\{\mathtt{b}, \mathtt{b} \to \mathtt{y}, \mathtt{y} \to \mathtt{f}\}, \mathtt{f}\rangle$	0.90	0.79	0.88	0.97
$\langle\{\mathtt{b}, \mathtt{b} \to \mathtt{y}\}, \mathtt{y}\rangle$	0.10	0.07	0.08	0.56
$\langle\{\mathtt{y}, \mathtt{y} \to \mathtt{f}\}, \mathtt{f}\rangle$	0.33	0.22	0.23	0.60

In the same way that we consider the strength of an argument, we can consider the strength of a defeasible rule.

Definition 15. *The* **strength** *of rule $\psi_1 \wedge \ldots \wedge \psi_n \to \phi$ is $S^x_P(A)$ where $A = \langle\{\psi_1, \ldots, \psi_n, \psi_1 \wedge \ldots \wedge \psi_n \to \phi\}, \phi\rangle$ and P is a probability distribution and $x \in \{p, d, s, k\}$.*

The following result says that if we want compositionally strong arguments, then we only need to consider strong defeasible rules.

Proposition 5. *If argument A is compositionally strong w.r.t. strength measure S^x_P and threshold τ, then the strength of any rule $\rho \in \mathsf{Support}(A)$ is greater than or equal to τ.*

Proof. An argument A is compositionally strong with respect to strength measure S^x_P and threshold $\tau \in [-1, 1]$ iff for all $B \in \mathsf{Intermediates}(A)$, $S^x_P(B) \geq \tau$. So for all $B \in \mathsf{Intermediates}(A)$ of the form $B = \langle\{\psi_1, \ldots, \psi_n, \psi_1 \wedge \ldots \wedge \psi_n \to \phi\}, \phi\rangle$, $S^x_P(B) \geq \tau$. So for all rules $\rho = \psi_1 \wedge \ldots \wedge \psi_n \to \phi \in \mathsf{Support}(A)$, the strength of ρ is greater than or equal to τ.

The above result means that if we do select strong defeasible rules for our knowledgebase, then we do not risk missing strong arguments. In other words, by rejecting weak defeasible rules, rules that are not going to lead to strong arguments are eliminated from the knowledgebase.

8 Discussion

Returning to the **audience scenario** given in the introduction, the confirmation measures give us different ways to judge the arguments, and decide which we regard as strong arguments. For those arguments that we identify as weak, we may argue with the person who provided the arguments in order to question those arguments. Possibly, they may provide supporting arguments to back-up the arguments under question, and if we are convinced by those supporting arguments, we have the option to then update our probability distribution. The protocol for this, the criteria for being convinced by the supporting arguments, and the method for updating the probability distribution, are beyond the scope of this paper, but this expanded scenario indicates how being able to analyze the strength of the arguments presented by others is potentially useful as part of an argumentation process.

Another application is in a **persuasion scenario** (for a review of persuasion see [17]). Assume we have a knowledgebase from which we can construct arguments. In a dialogue with another agent, we may want to select the best arguments to present in order to maximize the likelihood that they will be persuaded. For this, we can construct a probability distribution that reflects what we think the other agent believes about the world. Then using that probability distribution, we could select the stronger arguments to present.

A third example of an application is an **analytical scenario**. If we have acquired knowledge (perhaps from multiple sources), we may want to analyze the quality of the arguments generated from that knowledge. We can construct multiple probability distributions in order to investigate the arguments. Each probability distribution could reflect a possible modelling of the world, and so the change in strength for specific arguments could be investigated. Robustness could be investigated by identifying how extreme the modelling would be for arguments to be substantially weakened or strengthened. We leave the framework for undertaking robustness analysis to future work.

So there are potential applications for measuring argument strength, but it is an insufficiently understood notion in the literature on computational models of arguments. To address this, we consider a very simple defeasible logic with a clear semantics. This is so that we can get a clear understanding of the key concepts. We could have used an existing proposal for argumentation, but then the underlying issues we wanted to explore would be less clear in a more complex framework (e.g. defeasible logic programming [12]). Nonetheless, we believe this paper provides insights relevant for other argumentation systems, and so in future work, we will adapt existing proposals for structured argumentation systems (e.g. [12, 23, 35]) to quantify strength in a probabilistic context.

References

1. Amgoud, L., Ben-Naim, J.: Axiomatic foundations of acceptability semantics. In: Proceedings of KR'16. pp. 2–11. AAAI Press (2016)
2. Baroni, P., Rago, A., Toni, F.: From fine-grained properties to broad principles for gradual argumentation: A principled spectrum. International Journal of Approximate Reasoning **105**, 252–286 (2019)
3. Besnard, P., García, A., Hunter, A., Modgil, S., Prakken, H., Simari, G., Toni, F.: Introduction to structured argumentation. Argument and Computation **5**(1), 1–4 (2014)
4. Bonzon, E., Delobelle, J., Konieczny, S., Maudet, N.: A comparative study of ranking-based semantics for abstract argumentation. In: Proceedings of AAAI'16 (2016)
5. Carnap, R.: Logical Foundations of Probability. University of Chicago Press, 2nd edn. (1962)
6. Cayrol, C., Lagasquie-Schiex, M.: Graduality in argumentation. Journal of Artificial Intelligence Research **23**, 245–297 (2005)
7. Christensen, D.: Measuring confirmation. Journal of Philosophy **96**, 437–461 (1999)
8. Cohen, A., Gottifredi, S., Tamargo, L., Garca, A., Simari, G.: An informant-based approach to argument strength in defeasible logic programming. Argument and Computation **12**(1), 115–147 (2021)
9. Crupi, V.: Confirmation. In: The Stanford Encyclopedia of Philosophy. Metaphysics Research Lab, Stanford University, Spring 2020 edn. (2020)
10. Dung, P.: 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 (1995)
11. Dung, P., Thang, P.: Towards (probabilistic) argumentation for jury-based dispute resolution. In: Proceedings of COMMA'10. pp. 171–182. IOS Press (2010)
12. Garcia, A., Simari, G.R.: Defeasible logic programming: an argumentative approach. Theory and Practice of Logic Programming **4**(1–2), 95–138 (2004)
13. Governatori, G., Maher, M., Antoniou, G., Billington, D.: Argumentation semantics for defeasible logic. J. Logic & Comp. **14**(5), 675–702 (2004)
14. Haenni, R.: Modelling uncertainty with propositional assumptions-based systems. In: Applications of Uncertainty Formalisms, LNCS, vol. 1455, pp. 446–470. Springer (1998)
15. Hunter, A.: A probabilistic approach to modelling uncertain logical arguments. International Journal of Approximate Reasoning **54**(1), 47–81 (2013)
16. Hunter, A.: Generating instantiated argument graphs from probabilistic information. In: Proceedings of ECAI'20. IOS Press (2020)
17. Hunter, A., Chalaguine, L., Czernuszenko, T., Hadoux, E., Polberg, S.: Towards computational persuasion via natural language argumentation dialogues. In: Proceedings of KI'19. LNCS, vol. 11793, pp. 18–33. Springer (2019)
18. Hunter, A., Thimm, M.: Probabilistic reasoning with abstract argumentation frameworks. Journal of Artificial Intelligence Research **59**, 565–611 (2017)
19. Kemeny, J., Oppenheim, P.: Degrees of factual support. Philosophy of Science **19**, 307–324 (1952)
20. Li, H., Oren, N., Norman, T.J.: Probabilistic argumentation frameworks. In: Proceedings of TAFA'11. LNCS, vol. 7132, pp. 1–16. Springer (2011)
21. Matt, P.A., Toni, F.: A game-theoretic perspective on the notion of argument strength in abstract argumentation. In: Proceedings of JELIA'08. vol. LNCS, pp. 285–297. Springer (2008)

22. McCarthy, J.: Circumscription: A form of non-monotonic reasoning. Artificial Intelligence **28**(1), 89–116 (1980)
23. Modgil, S., Prakken, H.: The ASPIC+ framework for structured argumentation: a tutorial. Argument and Computation **5**, 31–62 (2014)
24. Pfeifer, N.: On argument strength. In: Bayesian Argumentation, Synthese Library, vol. 362, pp. 185–193. Springer (2012)
25. Polberg, S., Hunter, A.: Empirical evaluation of abstract argumentation: Supporting the need for bipolar and probabilistic approaches. Int. J. Approx. Reasoning **93**, 487–543 (2018)
26. Polberg, S., Hunter, A., Thimm, M.: Belief in attacks in epistemic probabilistic argumentation. In: Proceedings of the International Conference on Scalable Uncertainty Management (SUM'17). LNCS, vol. 10564, pp. 223–236. Springer (2017)
27. Pollock, J.: Defeasible reasoning. Cognitive Science **11**(4), 481–518 (1987)
28. Pollock, J.: Cognitve Carpentry. MIT Press (1995)
29. Prakken, H.: Probabilistic strength of arguments with structure. In: Proceedings of KR'18. pp. 158–167. AAAI Press (2018)
30. Riveret, R., Governatori, G.: On learning attacks in probabilistic abstract argumentation. In: Proc. of AAMAS'16. pp. 653–661 (2016)
31. Riveret, R., Rotolo, A., Sartor, G., Prakken, H., Roth, B.: Success chances in argument games: a probabilistic approach to legal disputes. In: Proceedings of JURIX'07. pp. 99–108. IOS Press (2007)
32. Shakarian, P., Simari, G., Moores, G., Paulo, D., Parsons, S., Falappa, M., Aleali, A.: Belief revision in structured probabilistic argumentation - model and application to cyber security. Ann. Math. Artif. Intell. **78**(3-4), 259–301 (2016)
33. Simari, G., Loui, R.: A mathematical treatment of defeasible reasoning and its implementation. Artificial Intelligence **53**(2-3), 125–157 (1992)
34. Thimm, M.: A probabilistic semantics for abstract argumentation. In: Proceedings of ECAI'12 (2012)
35. Toni, F.: A tutorial on assumption-based argumentation. Argument & Computation **5**(1), 89–117 (2014)
36. Verheij, B.: Arguments and their strength: revisiting Pollock's anti-probabilistic starting points. In: Proceedings of COMMA'14. IOS Press (2014)