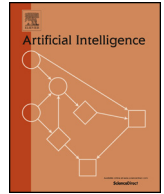


Contents lists available at [ScienceDirect](#)

# Artificial Intelligence

journal homepage: [www.elsevier.com/locate/artint](http://www.elsevier.com/locate/artint)

## Syntactic reasoning with conditional probabilities in deductive argumentation

Anthony Hunter<sup>a,\*</sup>, Nico Potyka<sup>b</sup><sup>a</sup> Department of Computer Science, University College London, London, UK<sup>b</sup> Department of Computing, Imperial College London, London, UK

### ARTICLE INFO

#### Article history:

Received 30 March 2022

Received in revised form 10 March 2023

Accepted 18 April 2023

Available online 2 May 2023

#### Keywords:

Probabilistic argumentation

Deductive argumentation

Reasoning with evidence

Probabilistic logic

### ABSTRACT

Evidence from studies, such as in science or medicine, often corresponds to conditional probability statements. Furthermore, evidence can conflict, in particular when coming from multiple studies. Whilst it is natural to make sense of such evidence using arguments, there is a lack of a systematic formalism for representing and reasoning with conditional probability statements in computational argumentation. We address this shortcoming by providing a formalization of conditional probabilistic argumentation based on probabilistic conditional logic. We provide a semantics and a collection of comprehensible inference rules that give different insights into evidence. We show how arguments constructed from proofs and attacks between them can be analyzed as arguments graphs using dialectical semantics and via the epistemic approach to probabilistic argumentation. Our approach allows for a transparent and systematic way of handling uncertainty that often arises in evidence.

© 2023 The Author(s). Published by Elsevier B.V. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

In many fields, in particular in science, technology, medicine, and social sciences, evidence comes in the form of conditional probabilities (either explicitly or implicitly) that have been obtained by analyzing data that may have come from surveys, databases, or scientific experiments. Systematic use of evidence is important in diverse areas of professional and administrative life including healthcare, ecology management, environmental management, education, policing and security administration, and health and safety management.

However, the rapidly increasing amount of evidence available on a subject means that it is difficult for a decision maker to locate, or even be aware of, new evidence that is relevant to their needs. Even if the decision maker locates the necessary evidence, it is difficult for them to effectively and efficiently assimilate and fully exploit the current state of all the evidence. In addition to the difficulty presented by the sheer volume of information, the evidence is often conceptually complex, heterogeneous, incomplete and inconsistent. So for the decision maker, it is imperative to abstract away from the details of individual items of evidence, and to aggregate the evidence in a way that reduces the volume, complexity, inconsistency and incompleteness.

To illustrate these issues, consider how evidence is acquired and used in healthcare. Many studies such as randomized clinical trials, and observational studies, are published every year in the medical literature. Healthcare professionals need

\* Corresponding author.

E-mail address: [anthony.hunter@ucl.ac.uk](mailto:anthony.hunter@ucl.ac.uk) (A. Hunter).

to use this evidence when making decisions for specific patients. To help them, there are systematic ways to collate and analyze evidence in the form of meta-analyses, systematic reviews, and evidence-based guidelines. So individual healthcare professionals might use a mix of primary evidence (such as from randomized clinical trials, and observational studies) as well as meta-analyses, systematic reviews, and guidelines [34,35]. In order to do this, the quality of the evidence is taken into account as we illustrate in the following example.

**Example 1.** Suppose there are two treatments  $T_1$  and  $T_2$  for a disease  $D$ , and one study  $S_1$  suggests that  $T_1$  cures  $D$  in 75% of patients, but with a 1% chance of an unpleasant side-effect, and another study  $S_2$  suggests that  $T_2$  cures  $D$  in 99% of patients, also with a 1% chance of an unpleasant side-effect. So  $T_2$  appears to be better than  $T_1$ . However, also suppose that  $S_1$  was a very large study conducted to very high standards and supported by an independent and highly reputable healthcare charity whereas  $S_2$  is a small study that was not entirely randomized and was sponsored by the manufacturer of the treatment  $T_2$ . So now there is uncertainty about whether the evidence in  $S_2$  is indeed reliable.

For domains such as in healthcare, Example 1 illustrates that we require computational techniques that represent the key evidence from each study, and computational techniques to analyze this evidence. These techniques should support the querying and aggregation of the knowledge. Furthermore, since the evidence, will be incomplete, uncertain, and inconsistent in various ways, these techniques should reflect how professionals deal with such information, which is often by constructing and evaluating arguments and counterarguments (for example, see an analysis of argumentation in clinical research publications [31], development of techniques for argument mining from clinical research publications [57], development of computational argumentation for aggregating evidence from randomized clinical trials [49], and development of computational argumentation for making recommendations from potentially conflicting clinical guidelines [18,87]).

Central to the development of computational techniques for reasoning with evidence is the need to have an appropriate representation. Since we assume evidence is obtained from scientific studies (e.g. randomized trials, cohort studies, meta-analyses, etc), there is normally an underlying formal structure to the evidence. For instance, in medical evidence the information we need to extract consists of specific categories of data from published papers about clinical trials, as specified by the PICO format, where  $P$  is for the patient class of the trial,  $I$  is for the intervention class (i.e. treatment) to which some patients were assigned,  $C$  is for the control class (e.g. alternative treatment) to which the remaining patients were assigned, and  $O$  is the outcome being measure the difference between the intervention and control classes (see [42] for more information on the PICO format). To illustrate, consider an example based on an actual paper as follows [28].

...patients with axillary lymph node-negative, estrogen receptor-positive breast cancer. ... chemotherapy plus tamoxifen resulted in significantly better disease-free survival than tamoxifen alone (90% for MFT versus 85% for tamoxifen [ $P = .01$ ]; ...

From this information, we can extract details of the patient class for the trial, including the facts that they are breast cancer patients who are axillary lymph node-negative and estrogen receptor-positive, that the treatments are either chemotherapy plus tamoxifen or tamoxifen alone, and that with the outcome disease-free survival, chemotherapy plus tamoxifen was significantly better than tamoxifen. We can for instance model this evidence as follows.

**Example 2.** Let  $bc$  denote breast cancer patient,  $ln$  denote axillary lymph node-negative,  $ep$  denote estrogen receptor-positive,  $ct$  denote chemotherapy plus tamoxifen,  $ta$  denote tamoxifen alone, and  $ds$  denote disease-free survival. The evidence in [28] quoted above can be represented by the following conditional probability statements.

$$P(ds|ln \wedge ep \wedge ct) = 0.9 \quad P(ds|ln \wedge ep \wedge ta) = 0.85$$

So Example 2 illustrates how representing evidence from clinical trials in the form of conditional probabilities provides an ideal format for capturing the object-level uncertainty. Efficacy of any treatment can then be considered in terms of conditional probability (i.e. the probability of that the treatment has a specified desired outcome given the patient class and circumstance and the treatment). There may be multiple desired outcomes such as disease-free survival, and survival over 5 years, and so there could be multiple conditional probability statements for the same patient class and treatment. This means that reasoning with the evidence becomes a multi-dimensional problem. This is further compounded by considering side-effects. As illustrated by Example 3 below, we can capture the different perspectives (different benefits and drawbacks) in the form of arguments and counterarguments.

**Example 3.** Consider treatments  $T_1$  and  $T_2$ , where  $T_1$  is better than  $T_2$  for treating the disorder according to a particular positive outcome but with respect to a particular side-effect,  $T_1$  is worse than  $T_2$ . Deciding whether  $T_1$  is a better choice than  $T_2$  depends on the two outcomes (the positive outcome and the side-effect) and their magnitudes (for example if the positive outcome is only marginally in favor of  $T_1$  but the side-effect is very severe, it may be preferable to choose  $T_2$ ). Preferences over outcome indicators and their magnitude can be represented by counterarguments. For instance, an argument that assumes  $T_1$  as a treatment can have a counterargument that expresses the sentiment against the treatment. In this way, the relative utility of one option over another can be captured in terms of argumentation.

Medical evidence can also come from analyzing healthcare databases such as electronic health records. Unfortunately, general patterns obtained by data analytics can be lower grade evidence by comparison with randomized clinical trials and meta-analyses. Furthermore, biases and confounding factors can commonly and profoundly affect relationships between interventions and outcomes when compared with randomized clinical trials [34,35]. These are issues of the quality of evidence. They raise different perspectives on the evidence, and so can also be analyzed using arguments and counterarguments as illustrated by Example 4 below.

**Example 4.** When looking at a patient database, it might be seen that patients with a particular disorder might be treated with either  $T_1$  or  $T_2$ , and that patients with  $T_1$  tend to have low mortality but patients with  $T_2$  tend to have higher mortality. From this, it might be inferred that  $T_1$  is better than  $T_2$ . However, it might be the case that the data is biased because  $T_1$  is only given to patients who have a mild form of the disorder whereas  $T_2$  is given to patients with a serious form of the disorder. In other words, there is a counterargument based on the observation that the two patients groups are not equivalent.

The effectiveness of any medical test can also be considered in terms of conditional probability (i.e. the probability that a patient has a disorder given the patient class and circumstance and the test being positive). In order to reason about medical tests, we may also need to use Bayes theorem, where the known conditional probability is the test being positive given the patient class and circumstances and the patient has the disease.

Whilst conditional probabilities are a natural and widely used format for representing evidence in many domains including in healthcare, there are challenges to using this in automated reasoning systems for argumentation.

**Transparent reasoning with conditional probabilities** We need a knowledge representation and reasoning formalism that allows for representation of conditional probabilities but also allows for transparent reasoning. Probabilistic graphical models like Bayesian networks have been applied to model medical diagnosis problems [11,40,52]. However, Bayesian networks are less flexible in that they only allow conditioning a variable on a fixed set of parents and require a full parametrization for the conditional probability of the variable given its parents. Probabilistic conditional logic has been applied in medical expert systems before [92]. However, with the exception of [29], reasoning algorithms are usually based on solving numerical optimization problem [6,41]. While this approach has computational advantages, it often remains unclear why an inference follows from a premise. Relatively little consideration has been given to a structured way of breaking inference into easier to understand steps.

**Argumentation with conditional probabilities** We need to incorporate reasoning with conditional probabilities into a framework for computational argumentation. There are numerous proposals for structured (i.e. logic-based) argumentation (for a review see [9]), but relatively few proposals for using probabilistic logics. Moreover, there has been a lack of consideration of how conditional probabilities can be directly handled in structured argumentation, though as we discuss in the literature review, some proposals handle conditional probabilities as defeasible rules (e.g. [21]). As part of argumentation with conditional probabilities, we need to represent meta-level uncertainty about the quality of evidence and harness it in the argumentation so as to reflect how both object-level and meta-level uncertainty are harnessed in the computational argumentation. Existing approaches to handling meta-level uncertainty in argumentation include the use of preferences (e.g. [3]), and the use of Dempster-Shafer theory (e.g. [63,91]), there is relatively little consideration of meta-level uncertainty of probabilistic evidence in argumentation.

The solution to the above challenges that we present in this paper is the following: (1) A collection of comprehensible inference rules for probabilistic logic with conditional probability statements that allows for individual inference steps to be clearly represented in the argumentation; and (2) A version of deductive argumentation that is based on probabilistic logic with conditional probability statements.

The advantage of our proposal is that we can develop applications where evidence is reasoned with automatically, and argument graphs can be constructed automatically, or existing arguments and counterarguments can be represented and checked for correctness. Furthermore, meta-level uncertainty can be represented and used to ensure that acceptable arguments are believed. By constructing arguments and counterarguments, where the premises are conditional probability statements, we can have processes for analyzing evidence that are transparent, systematic, and auditable.

We proceed as follows in the rest of the paper: In Section 2, we review the aspects of probabilistic logic that we require for our proposal, and we review abstract argumentation; In Section 3, we present a set of inference rules for probabilistic logic; In Section 4, we present our framework for deductive argumentation based on probabilistic logic with conditional probability statements, including the construction of arguments and attacks, and analysis of arguments graphs using dialectical semantics and via the epistemic approach to probabilistic argumentation; In Section 5, we compare with the related literature; And in Section 6, we revisit the goals set out in the introduction, and discuss some ideas for future work.

## 2. Preliminaries

In this section, we review a well-known approach to probabilistic logic, and we review the key definitions for abstract argumentation.

## 2.1. Probabilistic logic

We assume classical propositional logic for describing aspects of the world, and we model uncertainty about the formulae using a probability distribution over models of the propositional formulae (as reviewed in [60,64]).

A classical propositional language  $\mathcal{L}(\mathcal{A})$  of formulae is composed from a finite set of propositional atoms  $\mathcal{A}$  and the logical connectives  $\wedge$ ,  $\vee$ , and  $\neg$  in the usual way. We use  $\alpha$  and  $\beta$  to denote arbitrary propositional formulae and we use  $\perp$  to denote a contradictory formula and  $\top$  to denote a tautological formula. We assume the usual machinery of classical propositional logic.

Given a language  $\mathcal{L}(\mathcal{A})$ , the set of models (i.e. interpretations) of the language is denoted  $\mathcal{M}(\mathcal{A})$ . Each **model** is an assignment of true or false to the propositional formulae of the language defined in the usual way for classical logic. We represent a model  $m$  by a set of atoms. So  $\mathcal{M}(\mathcal{A})$  is the power set of atoms. For  $\alpha \in \mathcal{L}(\mathcal{A})$ , and  $m \in \mathcal{M}(\mathcal{A})$ , we say that  $m$  satisfies  $\alpha$  and write  $m \models \alpha$  iff  $\alpha \in m$ . For each  $\alpha \in \mathcal{L}(\mathcal{A})$ , we let  $\text{Models}(\alpha) = \{m \in \mathcal{M}(\mathcal{A}) \mid m \models \alpha\}$  denote the **set of models** of  $\alpha$ .

In order to assign probabilities to formulae, we consider probability distributions over the models of our language.

**Definition 1.** Let  $\mathcal{M}(\mathcal{A})$  be the models of the language  $\mathcal{L}(\mathcal{A})$ . A **probability distribution**  $P$  on  $\mathcal{M}(\mathcal{A})$  is a function  $P : \mathcal{M}(\mathcal{A}) \rightarrow [0, 1]$  such that  $\sum_{m \in \mathcal{M}(\mathcal{A})} P(m) = 1$ .

As usual, we extend the domain of  $P$  to formulae by defining the probability of formulae as the probability of their models.

**Definition 2.** Let  $\mathcal{M}(\mathcal{A})$  be the models of the language  $\mathcal{L}(\mathcal{A})$ , and let  $P$  be a probability distribution on  $\mathcal{M}(\mathcal{A})$ . The **belief in a formula**  $\alpha \in \mathcal{L}(\mathcal{A})$  w.r.t.  $P$  is  $P(\alpha) = \sum_{m \in \text{Models}(\alpha)} P(m)$ .

**Example 5.** Let  $\mathcal{L}(\mathcal{A})$  be the propositional language that can be formed from  $\{a, b\}$ . Now suppose  $P(\{a, b\}) = 0.8$  and  $P(\{a\}) = 0.2$ . Then,  $P(a) = 1$ ,  $P(a \wedge b) = 0.8$ ,  $P(b \vee \neg b) = 1$ ,  $P(a \wedge \neg b) = 0.2$ , etc.

A **probabilistic conditional formula**, or **conditional** for short, is an expression of the form  $(\alpha|\beta)[v, w]$  where  $\alpha, \beta \in \mathcal{L}(\mathcal{A})$  and  $v, w \in [0, 1]$  such that  $v \leq w$ . Intuitively,  $(\alpha|\beta)[v, w]$  expresses the belief that  $\alpha$  holds given  $\beta$  is between  $v$  and  $w$ . Point-value formulae are obtained for the special case  $(\alpha|\beta)[v, v]$ , and these are abbreviated by  $(\alpha|\beta)[v]$ . Unconditioned formulae are obtained for the special case  $(\alpha|\top)[v, w]$ , where the condition is tautological, and these are abbreviated by  $(\alpha)[v, w]$ . We use  $\phi$  and  $\psi$  for arbitrary conditionals.

**Example 6.** Consider the conditional  $(\text{flu}|\text{fever} \wedge \text{ache})[0.75, 0.85]$  which represents the probability of flu, denoted  $\text{flu}$ , is between 0.75 and 0.85 given the symptoms of fever, denoted  $\text{fever}$ , and muscle ache, denoted  $\text{ache}$ .

For a conditional  $\phi = (\alpha|\beta)[v, w]$ , let  $\text{Head}(\phi) = \alpha$ ,  $\text{Body}(\phi) = \beta$ ,  $\text{Lower}(\phi) = v$ , and  $\text{Upper}(\phi) = w$ . We let  $\mathcal{F}(\mathcal{A})$  be the set of conditionals formed from  $\mathcal{L}(\mathcal{A})$ . A **knowledgebase**  $\Delta$  is a finite set of conditionals.

Intuitively, a probability distribution  $P$  satisfies  $(\beta|\alpha)[v, w]$  iff the conditional probability of  $\beta$  given  $\alpha$  under  $P$  is between  $v$  and  $w$ . The conditional probability of  $\beta$  given  $\alpha$  is usually defined as  $P(\beta|\alpha) = \frac{P(\alpha \wedge \beta)}{P(\alpha)}$ . However, since this expression is undefined when  $P(\alpha) = 0$ , satisfaction of conditionals is often defined in a slightly different way [54].

**Definition 3.** A probability distribution  $P$  **satisfies** a conditional  $(\beta|\alpha)[v, w]$ , denoted  $P \models (\beta|\alpha)[v, w]$ , iff

$$P(\alpha) \cdot v \leq P(\alpha \wedge \beta) \leq P(\alpha) \cdot w \quad (1)$$

For a conditional  $\phi$ , let  $\text{Sat}(\phi)$  be the set of probability distributions satisfying  $\phi$  (i.e.  $\text{Sat}(\phi) = \{P \mid P \models \phi\}$ ). For a knowledgebase  $\Delta \subseteq \mathcal{F}(\mathcal{A})$ ,  $\text{Sat}(\Delta) = \bigcap_{\phi \in \Delta} \text{Sat}(\phi)$ .

**Example 7.** For  $\Delta = \{(a|b)[0.4, 0.6], (b|c)[0.6, 0.7], (c)[0.7, 0.8]\}$ , a probability distribution where  $P(a \wedge b \wedge c) = 0.25$ ,  $P(a \wedge \neg b \wedge c) = 0.25$ ,  $P(\neg a \wedge b \wedge c) = 0.25$ , and  $P(\neg a \wedge \neg b \wedge \neg c) = 0.25$ , satisfies  $\Delta$ .

Note that when dividing every term in (1) by  $P(\alpha)$ , the equation just demands that the conditional probability is between  $v$  and  $w$ . However, the constraint is also defined if  $P(\alpha) = 0$ . In this case, the constraint is satisfied even though the conditional probability is undefined. This can be seen as a generalization of the idea that the logical implication  $\alpha \rightarrow \beta$  is satisfied whenever  $\alpha$  is false (no matter what  $\beta$  is).

Following [54], we define two notions of entailment from a knowledgebase.

**Definition 4.** Let  $\Delta \subseteq \mathcal{F}(\mathcal{A})$  be a knowledgebase and let  $\phi \in \mathcal{F}(\mathcal{A})$  be a conditional. We say that  $\Delta$  **entails**  $\phi$  and write  $\Delta \models \phi$  iff  $\text{Sat}(\Delta) \subseteq \text{Sat}(\phi)$ . We say that  $\Delta$  **tightly entails**  $(\beta|\alpha)[v, w]$  and write  $\Delta \models_t (\beta|\alpha)[v, w]$  iff  $\text{Sat}(\Delta) \neq \emptyset$  and  $\Delta \models (\beta|\alpha)[v', w']$  implies that  $[v, w] \subseteq [v', w']$ .

Intuitively, a knowledgebase entails a conditional if every probability distribution that satisfies the knowledgebase also satisfies the conditional. However, this notion is not very strong because if a knowledgebase satisfies a conditional  $(\beta|\alpha)[v, w]$ , it also satisfies the conditional  $(\beta|\alpha)[v', w']$  whenever  $[v, w] \subseteq [v', w']$ . In particular, every knowledgebase entails  $(\beta|\alpha)[0, 1]$  for arbitrary formulae  $\alpha, \beta$  because probabilities are necessarily between 0 and 1. Tight entailment captures the tightest (and most informative) interval for which the entailment relation holds. Since this notion only makes sense when the knowledgebase is consistent, we also assume  $\text{Sat}(\Delta) \neq \emptyset$ . If  $\text{Sat}(\Delta) = \emptyset$ , tight entailment is not defined and we write  $\Delta \models (\beta|\alpha)\emptyset$ . The tight entailment result is uniquely defined and can be computed by solving two linear optimization problems (one for the lower and one for the upper bound) [41]. More formally, tight entailment gives rise to the *probabilistic entailment problem*: given a knowledgebase  $\Delta \subset \mathcal{F}(\mathcal{A})$  and formulae  $\alpha, \beta \in \mathcal{F}(\mathcal{A})$ , solve the two optimization problems  $\min_{P \models \Delta} / \max_{P \models \Delta} P(\beta | \alpha)$ , s.t.  $P(\alpha) > 0$ . Strictly speaking, the optimization problems are fractional. However, they can be transformed into equivalent linear optimization problems [41]. The problem is feasible whenever there exists a probability function  $P$  such that  $P \models \Delta$ . Checking this property is called the *probabilistic satisfiability problem* and can be done by solving a linear optimization problem again. It is worth pointing out that we can also have “weak inconsistencies” in this context. These are the probabilistic analogue of “classical weak inconsistencies”. For example, the formulae  $\alpha \rightarrow \beta$  and  $\alpha \rightarrow \neg\beta$  are consistent even though their conclusion is clearly inconsistent. Intuitively, consistency is established by letting  $\alpha$  be false. Similarly, the conditionals  $(b|a)[0.9]$  and  $(b|a)[0.1]$  are consistent. However, they entail  $(a)[0]$ . Just like in the classical setting, this can be undesirable in applications. The *strict probabilistic satisfiability problem* is a stronger form of the probabilistic satisfiability problem that also checks if there is a probability function that assigns non-zero probability to the condition of every conditional. This problem can again be solved by a linear optimization problem [76]. For future reference, we summarize the main results from the literature in the following proposition.

**Proposition 1** (*Probabilistic Reasoning with Linear Programming [32,41,76]*). *Let  $\Delta \subset \mathcal{F}(\mathcal{A})$  be a finite knowledgebase over a finite set of propositional atoms  $\mathcal{A}$ . Then the probabilistic satisfiability problem, the strict probabilistic satisfiability problem and the probabilistic entailment problem can be solved in worst-case exponential time with respect to the number of atoms.*

For the explicit form of the corresponding linear programs and the correctness proofs, we refer to Proposition 4.24 (probabilistic satisfiability problem), Proposition 4.30 and Proposition 4.37 in [76]. Let us note that while the worst-case runtime is exponential, the problems can often be solved significantly faster by applying column generation techniques [14, 27,32]. This often allows deciding the satisfiability of knowledge bases with hundreds of atoms in seconds. While classical SAT solvers can deal with significantly larger knowledge bases, this size is completely sufficient for our purposes.

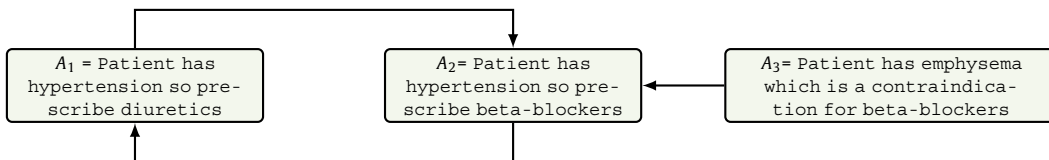
2.2. Abstract argumentation

We provide a brief review of abstract argumentation as proposed by Dung [26]. In this approach, each argument is treated as an atom, and so no internal structure of the argument needs to be identified.

**Definition 5.** An **argument graph** is a pair  $G = (\mathcal{N}, \mathcal{R})$  where  $\mathcal{N}$  is a set and  $\mathcal{R}$  is a binary relation over  $\mathcal{N}$  (in symbols,  $\mathcal{R} \subseteq \mathcal{N} \times \mathcal{N}$ ). Let  $\text{Nodes}(G)$  be the set of nodes in  $G$  and let  $\text{Arcs}(G) \subseteq \text{Nodes}(G) \times \text{Nodes}(G)$  be the set of arcs in  $G$ .

So an argument graph is a directed graph. Each element  $A \in \mathcal{N}$  is called an **argument** and  $(A_i, A_j) \in \mathcal{R}$  means that  $A_i$  **attacks**  $A_j$  (accordingly,  $A_i$  is said to be an **attacker** of  $A_j$ ). So  $A_i$  is a **counterargument** for  $A_j$  when  $(A_i, A_j) \in \mathcal{R}$  holds.

**Example 8.** Consider arguments  $A_1 =$  “Patient has hypertension so prescribe diuretics”,  $A_2 =$  “Patient has hypertension so prescribe beta-blockers”, and  $A_3 =$  “Patient has emphysema which is a contraindication for beta-blockers”. Here, we assume that  $A_1$  and  $A_2$  attack each other because we should only give one treatment and so giving one precludes the other, and we assume that  $A_3$  attacks  $A_2$  because it provides a counterargument to  $A_2$ . Hence, we get the following abstract argument graph.



Arguments can work together as a coalition by attacking other arguments and by defending their members from attack as follows.

**Definition 6.** Let  $S \subseteq \mathcal{N}$  be a set of arguments.

- $S$  **attacks**  $A_j \in \mathcal{N}$  iff there is an argument  $A_i \in S$  such that  $A_i$  attacks  $A_j$ .

- **S defends**  $A_i \in \mathcal{N}$  iff for each argument  $A_j \in \mathcal{N}$ , if  $A_j$  attacks  $A_i$  then  $S$  attacks  $A_j$ .

The following gives a requirement that should hold for a coalition of arguments to make sense. If it holds, it means that the arguments in the set offer a consistent view on the topic of the argument graph.

**Definition 7.** A set  $S \subseteq \mathcal{N}$  of arguments is **conflict-free** iff there are no arguments  $A_i$  and  $A_j$  in  $S$  such that  $A_i$  attacks  $A_j$ .

Now, we consider how we can find an acceptable set of arguments from an abstract argument graph. The simplest case of arguments that can be accepted is as follows.

**Definition 8.** A set  $S \subseteq \mathcal{N}$  of arguments is **admissible** iff  $S$  is conflict-free and defends all its arguments.

The intuition here is that for a set of arguments to be accepted, we require that, if any one of them is challenged by a counterargument, then they offer grounds to challenge, in turn, the counterargument. There always exists at least one admissible set: The empty set is always admissible.

Clearly, the notion of admissible sets of arguments is the minimum requirement for a set of arguments to be accepted. We will focus on the following classes of acceptable arguments.

**Definition 9.** Let  $\Gamma$  be a conflict-free set of arguments, and let  $\text{Defended} : \wp(\mathcal{N}) \mapsto \wp(\mathcal{N})$  be a function such that  $\text{Defended}(\Gamma) = \{A \mid \Gamma \text{ defends } A\}$ .

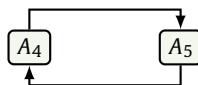
1.  $\Gamma$  is a **complete extension** iff  $\Gamma = \text{Defended}(\Gamma)$
2.  $\Gamma$  is a **grounded extension** iff it is the minimal (w.r.t. set inclusion) complete extension.
3.  $\Gamma$  is a **preferred extension** iff it is a maximal (w.r.t. set inclusion) complete extension.
4.  $\Gamma$  is a **stable extension** iff it is a preferred extension that attacks every argument that is not in the extension.

The grounded extension is always unique, whereas there may be multiple preferred extensions. We illustrate these definitions with the following examples. As can be seen from the examples, the grounded extension provides a skeptical view on which arguments can be accepted, whereas each preferred extension takes a credulous view on which arguments can be accepted.

**Example 9.** Continuing Example 8, there is only one complete set, and so this is both grounded and preferred. Note,  $\{A_1, A_2\}$ ,  $\{A_2, A_3\}$ , and  $\{A_1, A_2, A_3\}$  are not conflict-free subsets. Only the conflict-free subsets are given in the table.

	Conflict-free	Admissible	Complete	Grounded	Preferred	Stable
$\{\}$	✓	✓	×	×	×	×
$\{A_1\}$	✓	✓	×	×	×	×
$\{A_2\}$	✓	×	×	×	×	×
$\{A_3\}$	✓	✓	×	×	×	×
$\{A_1, A_3\}$	✓	✓	✓	✓	✓	✓

**Example 10.** Consider the following argument graph. For this, there are two preferred sets, neither of which is grounded. Note  $\{A_4, A_5\}$  is not conflict-free. Only the conflict-free subsets are given in the table.



	Conflict-free	Admissible	Complete	Grounded	Preferred	Stable
$\{\}$	✓	✓	✓	✓	×	×
$\{A_4\}$	✓	✓	✓	×	✓	✓
$\{A_5\}$	✓	✓	✓	×	✓	✓

The formalization we have reviewed in this section is abstract because both the nature of the arguments and the nature of the attack relation are ignored. In particular, the internal (logical) structure of each of the arguments is not made explicit. Nevertheless, Dung's proposal for abstract argumentation is valuable for clearly representing arguments and counterarguments, and for intuitively determining which arguments should be accepted (depending on whether we want to take a credulous or skeptical perspective).

Given an argument graph, let  $\text{Extensions}_\sigma(G)$  denote the set of extensions according to  $\sigma$  where  $\sigma = \text{co}$  denotes the complete extensions,  $\sigma = \text{pr}$  denotes the preferred extensions,  $\sigma = \text{gr}$  denotes the grounded extensions, and  $\sigma = \text{st}$  denotes the stable extensions. So continuing Example 9,  $\text{Extensions}_{\text{pr}}(G) = \{\{A_1, A_3\}\}$ , whereas continuing Example 10,  $\text{Extensions}_{\text{pr}}(G) = \{\{A_4\}, \{A_5\}\}$ .

### 3. Inference rules for probabilistic argumentation

In this section, we present a set of sound inference rules for probabilistic logic. We can then use these rules, or a subset of them, as a proof system. This will allow us to construct arguments in a transparent way from a set of basic assumptions and proofs that derive interesting conclusions from them.

#### 3.1. Probabilistic logical proof systems

State-of-the-art reasoning algorithms for the probabilistic entailment problem are based on solving linear optimization problems [14,27,43]. While this approach is computationally convenient, it is not well suited for probabilistic argumentation because the reasoning that leads to a conclusion is not transparent. A probabilistic proof system would be better suited for our needs. Unfortunately, to the best of our knowledge, there are currently no complete proof systems for probabilistic logic. However, 32 sound inference rules for probabilistic logic have been presented in [29]. The authors also showed that 5 of the rules yield a complete calculus for the fragment that contains only probabilistic facts  $(\alpha)[v, w]$  and deterministic conditionals  $(\beta | \alpha)[1]$ . While this fragment is not sufficiently expressive for our purposes, we can also work with a sound, but incomplete calculus. Unfortunately, no correctness proofs have been given in [29] and the inference rules are rather abstract since they have been designed for being most general and not most comprehensible. Therefore, we will develop a small set of sound and intuitive inference rules that seem useful for probabilistic argumentation in this section, and then we will investigate how they can be harnessed in a form of deductive argumentation in Section 4. We will discuss relationships to the rules in [29] in Section 5.

#### 3.2. A collection of sound inference rules

To begin with, we will develop some sound inference rules that seem useful for probabilistic argumentation. We represent *inference rules* in the following form, where  $\psi_1, \dots, \psi_l$  are the premises of the rule and  $\phi$  is its conclusion.

$$\frac{\psi_1, \dots, \psi_l}{\phi}$$

For example, the classical inference rule *Modus Ponens* can be written as the following inference rule where  $A$  and  $A \rightarrow B$  are the premises and  $B$  is the conclusion.

$$\frac{A, A \rightarrow B}{B}$$

Formally, an inference rule is *sound* if, for every set of assumptions  $\Delta$  (including the empty set),  $\Delta \models \psi_i$  for  $i = 1, \dots, l$  implies that  $\Delta \models \phi$ . That is, the conclusion of the rule follows logically from its premises. Combinations of these rules can be used to construct different proof systems with different expressiveness and proof complexity. Since we focus on sound inference rules, all proof systems are sound in the sense that  $\Delta \vdash_{\text{PS}} \phi$  implies  $\Delta \models \phi$ , where  $\vdash_{\text{PS}}$  denotes the derivability relation with respect to our proof system. However, we do not attempt to get completeness. That is, we do not expect that  $\Delta \models \phi$  implies  $\Delta \vdash_{\text{PS}} \phi$ . To the best of our knowledge, there are no known complete proof systems for the probabilistic logic that we consider here.

Before designing some interesting rules, let us note some simple but useful calculation rules.

**Lemma 1.** Let  $\alpha, \beta \in \mathcal{L}(\mathcal{A})$  and let  $P$  be a probability distribution on  $\mathcal{M}(\mathcal{A})$ . Then

1.  $P(\alpha) = 1 - P(\neg\alpha)$ .
2.  $P(\alpha) = 1$  if  $\alpha \equiv \top$  is a tautology and  $P(\alpha) = 0$  if  $\alpha \equiv \perp$  is a contradiction.
3.  $P(\alpha) = P(\alpha \wedge \beta) + P(\alpha \wedge \neg\beta)$ .
4.  $P(\alpha \wedge \beta) \leq P(\alpha)$ .
5.  $P(\alpha \vee \beta) \geq P(\alpha)$ .

**Proof.** 1. We have  $P(\alpha) + P(\neg\alpha) = \sum_{m \in \text{Models}(\alpha)} P(m) + \sum_{m \in \text{Models}(\neg\alpha)} P(m) = \sum_{m \in \mathcal{M}(\mathcal{A})} P(m) = 1$ . Subtracting  $P(\neg\alpha)$  from both sides yields the claim.

2. Follows immediately by observing that  $\text{Models}(\alpha) = \mathcal{M}(\mathcal{A})$  if  $\alpha$  is tautological and  $\text{Models}(\neg\alpha) = \emptyset$  if  $\alpha$  is a contradiction.

$$3. P(\alpha) = \sum_{m \in \text{Models}(\alpha)} P(m) = \sum_{m \in \text{Models}(\alpha \wedge \beta)} P(m) + \sum_{m \in \text{Models}(\alpha \wedge \neg\beta)} P(m) = P(\alpha \wedge \beta) + P(\alpha \wedge \neg\beta).$$

$$4. \text{Rewriting item 3, we get } P(\alpha \wedge \beta) = P(\alpha) - P(\alpha \wedge \neg\beta) \leq P(\alpha).$$

$$5. P(\alpha \vee \beta) = \sum_{m \in \text{Models}(\alpha \vee \beta)} P(m) \geq \sum_{m \in \text{Models}(\alpha)} P(m) = P(\alpha). \quad \square$$

We will now list a number of soundness results. That is, results of the form “If  $\Delta \models \psi_1, \dots, \Delta \models \psi_l$ , then  $\Delta \models \phi$ ”. We will summarize the corresponding inference rules in a table at the end of this section.

We first note that item 1 from Lemma 1 can be generalized to conditionals. This gives us a first inference rule that we call *Conditional Negation* and abbreviate by *CN*.

**Proposition 2 (CN).** *If  $\Delta \models (\beta \mid \alpha)[l, u]$ , then  $\Delta \models (\neg\beta \mid \alpha)[1 - u, 1 - l]$ .*

**Proof.** Consider an arbitrary probability function  $P$  such that  $P \models \Delta$ . Then, we have  $P(\alpha) \cdot l \leq P(\alpha \wedge \beta) \leq P(\alpha) \cdot u$  by definition. Since the models of  $\alpha \wedge \beta$ ,  $\alpha \wedge \neg\beta$ ,  $\neg\alpha \wedge \beta$ ,  $\neg\alpha \wedge \neg\beta$  form a partition of all models in our language, we have

$$P(\alpha \wedge \neg\beta) = 1 - P(\alpha \wedge \beta) - P(\neg\alpha \wedge \beta) - P(\neg\alpha \wedge \neg\beta) = 1 - P(\alpha \wedge \beta) - P(\neg\alpha).$$

From this, we first get the upper bound  $P(\alpha \wedge \neg\beta) \leq 1 - P(\alpha) \cdot l - (1 - P(\alpha)) = P(\alpha) \cdot (1 - l)$ . Symmetrically, we can derive the lower bound  $P(\alpha \wedge \neg\beta) \geq 1 - P(\alpha) \cdot u - (1 - P(\alpha)) = P(\alpha) \cdot (1 - u)$ . Hence, by definition,  $\Delta \models (\neg\beta \mid \alpha)[1 - u, 1 - l]$ .  $\square$

Note that item 1 from Lemma 1 corresponds to the special case  $\Delta \models (\beta)[p]$  of CN.

We can also generalize items 4 and 5 from Lemma 1. We call the rules *Conjunction Elimination* and *Disjunction Elimination* and abbreviate them by *CE* and *DE*, respectively.

**Proposition 3 (CE).** *If  $\Delta \models (\beta \wedge \gamma \mid \alpha)[l, u]$ , then  $\Delta \models (\beta \mid \alpha)[l, 1]$ .*

**Proof.** Consider an arbitrary probability function  $P$  such that  $P \models \Delta$ . Then, we have  $P(\alpha \wedge \beta) \geq P(\alpha \wedge \beta \wedge \gamma) \geq l \cdot P(\alpha)$ .  $\square$

**Proposition 4 (DE).** *If  $\Delta \models (\beta \vee \gamma \mid \alpha)[l, u]$ , then  $\Delta \models (\beta \mid \alpha)[0, u]$ .*

**Proof.** Since  $\beta \vee \gamma \equiv \neg(\neg\beta \wedge \neg\gamma)$ , we have  $\Delta \models (\neg(\neg\beta \wedge \neg\gamma) \mid \alpha)[l, u]$ . CN implies that  $\Delta \models (\neg\beta \wedge \neg\gamma \mid \alpha)[1 - u, 1 - l]$ . Applying CE implies  $\Delta \models (\neg\beta \mid \alpha)[1 - u, 1]$ . Finally, applying CN again implies  $\Delta \models (\beta \mid \alpha)[0, u]$ .  $\square$

It is often useful to go from two individual probabilities to their joint probability. For this purpose, we introduce the rules *Conjunction Insertion* and *Disjunction Insertion* that we abbreviate by *CI* and *DI*, respectively.

**Proposition 5 (CI).** *If  $\Delta \models (\beta \mid \alpha)[l_1, u_1]$  and  $\Delta \models (\gamma \mid \alpha)[l_2, u_2]$  then  $\Delta \models (\beta \wedge \gamma \mid \alpha)[l_3, u_3]$ , where  $l_3 = \max\{0, l_1 + l_2 - 1\}$  and  $u_3 = \min\{u_1, u_2\}$ .*

**Proof.** Consider an arbitrary probability function  $P$  such that  $P \models \Delta$ . Then, we have  $P(\alpha \wedge \beta \wedge \gamma) \leq \min\{P(\alpha \wedge \beta), P(\alpha \wedge \gamma)\} \leq \min\{P(\alpha) \cdot u_1, P(\alpha) \cdot u_2\} = P(\alpha) \cdot \min\{u_1, u_2\}$ , which proves the upper bound.

For the lower bound, first note that  $P(\alpha \wedge \beta) = P(\alpha \wedge \beta \wedge \gamma) + P(\alpha \wedge \beta \wedge \neg\gamma)$  and  $P(\alpha \wedge \gamma) = P(\alpha \wedge \beta \wedge \gamma) + P(\alpha \wedge \neg\beta \wedge \gamma)$ . Adding both equations and reordering terms gives

$$2 \cdot P(\alpha \wedge \beta \wedge \gamma) = P(\alpha \wedge \beta) + P(\alpha \wedge \gamma) - P(\alpha \wedge \beta \wedge \neg\gamma) - P(\alpha \wedge \neg\beta \wedge \gamma). \quad (2)$$

CN implies that  $\Delta \models (\neg\beta \mid \alpha)[1 - u_1, 1 - l_1]$  and  $\Delta \models (\neg\gamma \mid \alpha)[1 - u_2, 1 - l_2]$ . Therefore,  $P(\alpha \wedge \beta \wedge \neg\gamma) \leq P(\alpha \wedge \neg\beta) \leq (1 - l_2) \cdot P(\alpha)$  and  $P(\alpha \wedge \neg\beta \wedge \gamma) \leq P(\alpha \wedge \neg\beta) \leq (1 - l_1) \cdot P(\alpha)$ . Using these inequalities and our assumptions  $\Delta \models (\beta \mid \alpha)[l_1, u_1]$  and  $\Delta \models (\gamma \mid \alpha)[l_2, u_2]$  in (2), we get

$$\begin{aligned} 2 \cdot P(\alpha \wedge \beta \wedge \gamma) &\geq l_1 \cdot P(\alpha) + l_2 \cdot P(\alpha) - (1 - l_2) \cdot P(\alpha) - (1 - l_1) \cdot P(\alpha) \\ &\geq (2 \cdot l_1 + 2 \cdot l_2 - 2) \cdot P(\alpha). \end{aligned}$$

Dividing by 2 gives us  $P(\alpha \wedge \beta \wedge \gamma) \geq (l_1 + l_2 - 1) \cdot P(\alpha)$ .  $\square$

**Proposition 6 (DI).** *If  $\Delta \models (\beta \mid \alpha)[l_1, u_1]$  and  $\Delta \models (\gamma \mid \alpha)[l_2, u_2]$  then  $\Delta \models (\beta \vee \gamma \mid \alpha)[l_3, u_3]$ , where  $l_3 = \max\{l_1, l_2\}$  and  $u_3 = \min\{1, u_1 + u_2\}$ .*



**Proof.** CN implies that  $\Delta \models (\neg\beta \mid \alpha)[1 - u_1, 1 - l_1]$  and  $\Delta \models (\neg\gamma \mid \alpha)[1 - u_2, 1 - l_2]$ . From this, we can use CI to infer that  $\Delta \models (\neg\beta \wedge \neg\gamma \mid \alpha)[l'_3, u'_3]$ , where  $l'_3 = \max\{0, 1 - u_1 - u_2\}$  and  $u'_3 = \min\{1 - l_1, 1 - l_2\}$ . Using the fact that  $\neg(\neg\beta \wedge \neg\gamma) \equiv \beta \vee \gamma$  and applying CN again implies that  $\Delta \models (\beta \vee \gamma \mid \alpha)[l_3, u_3]$ , where  $l_3 = 1 - \min\{1 - l_1, 1 - l_2\} = \max\{l_1, l_2\}$  and  $u_3 = 1 - \max\{0, 1 - u_1 - u_2\} = \min\{1, u_1 + u_2\}$ .  $\square$

Let us emphasize that CI and DI do not make use of any independency assumptions. To highlight the difference, we add a refinement of CI that assumes conditional independence between  $\beta$  and  $\gamma$ . Formally,  $\beta$  and  $\gamma$  are conditionally independent given  $\alpha$  iff  $P(\beta \wedge \gamma \mid \alpha) = P(\beta \mid \alpha) \cdot P(\gamma \mid \alpha)$ . This generalizes the basic notion of independence between two random variables. Under this assumption, we get simpler lower and upper bounds that correspond to the products of the individual lower and upper bounds. We call this refinement *ICI* in the following, where the I stands for the independency assumption that has been added to CI.

**Proposition 7 (ICI).** *If  $\beta$  and  $\gamma$  are conditionally independent given  $\alpha$ ,  $\Delta \models (\beta \mid \alpha)[l_1, u_1]$  and  $\Delta \models (\gamma \mid \alpha)[l_2, u_2]$  then  $\Delta \models (\beta \wedge \gamma \mid \alpha)[l_3, u_3]$ , where  $l_3 = l_1 \cdot l_2$  and  $u_3 = u_1 \cdot u_2$ .*

**Proof.** Consider an arbitrary probability function  $P$  such that  $P \models \Delta$ . If  $P(\alpha) = 0$ , the claim is trivially true because all terms in the equation defining the satisfaction relation become 0.

Hence, we can assume  $P(\alpha) > 0$ . Then, we have

$$\begin{aligned} P(\alpha \wedge \beta \wedge \gamma) &= P(\alpha) \cdot P(\beta \wedge \gamma \mid \alpha) \\ &= P(\alpha) \cdot P(\beta \mid \alpha) \cdot P(\gamma \mid \alpha) \\ &= \frac{P(\alpha \wedge \beta) \cdot P(\alpha \wedge \gamma)}{P(\alpha)} \\ &\leq \frac{u_1 \cdot P(\alpha) \cdot u_2 \cdot P(\alpha)}{P(\alpha)} \\ &= (u_1 \cdot u_2) \cdot P(\alpha). \end{aligned}$$

The lower bound follows analogously.  $\square$

The classical modus ponens is a natural tool for classical argumentation. Recall that modus ponens states that if  $\alpha$  and  $\alpha \rightarrow \beta$  are true, then  $\beta$  must be true as well. Therefore, another interesting inference rule for probabilistic argumentation is a *probabilistic modus ponens* that we abbreviate by *PMP*.

**Proposition 8 (PMP).** *If  $\Delta \models (\alpha)[l_1, u_1]$  and  $\Delta \models (\beta \mid \alpha)[l_2, u_2]$ , then  $\Delta \models (\beta)[l_3, u_3]$  where  $l_3 = l_1 \cdot l_2$  and  $u_3 = \min\{1, u_1 \cdot u_2 + 1 - l_1\}$ .*

**Proof.** Consider an arbitrary probability function  $P$  such that  $P \models \Delta$ . Then, for the lower bound, we have

$$\begin{aligned} P(\beta) &= P(\alpha \wedge \beta) + P(\neg\alpha \wedge \beta) \\ &\geq l_2 \cdot P(\alpha) + 0 \\ &\geq l_2 \cdot l_1. \end{aligned}$$

For the upper bound, we have

$$\begin{aligned} P(\beta) &= P(\alpha \wedge \beta) + P(\neg\alpha \wedge \beta) \\ &\leq u_2 \cdot P(\alpha) + P(\neg\alpha) \\ &\leq u_2 \cdot u_1 + (1 - l_1). \end{aligned}$$

Since we also have  $P(\beta) \leq 1$ , we can conclude  $P \models (\beta)[l_1 \cdot l_2, \min\{1, u_1 \cdot u_2 + 1 - l_1\}]$ .  $\square$

Note that we obtain the classical modus ponens for  $l_1 = u_1 = l_2 = u_2 = 1$ . Then the statement is that if  $\Delta \models (\alpha)[1]$  and  $\Delta \models (\beta \mid \alpha)[1]$ , then  $\Delta \models (\beta)[1]$ . Note also that if  $\Delta \models (\alpha)[0]$ , the claim becomes that  $\Delta \models (\beta)[0, 1]$ , which seems intuitively reasonable because the rule "if  $\alpha$  then  $\beta$ " is not applicable and so there is nothing that we could conclude. We give some more examples to give a better intuition for the probabilistic modus ponens.

**Example 11.** Consider the knowledgebase  $\Delta_1 = \{(a)[0.9], (b | a)[0.9]\}$ . Then PMP allows us to conclude that  $\Delta_1 \models (b)[0.81, 0.91]$ . As this example illustrates, we can end up with an interval probability even when the original probabilities were point probabilities. Intuitively, this is because our assumptions do not say anything about the probability of  $b$  when  $a$  is false. To see that the interval cannot be tightened further, we give examples for probability functions that satisfy all three conditionals.  $P_l$  with  $P_l(\{a, b\}) = 0.81$ ,  $P_l(\{a\}) = 0.09$  and  $P_l(\emptyset) = 0.1$  is an example for the lower bound, that is, we have  $P_l \models \Delta$  and  $P_l \models (b)[0.81]$ .  $P_u$  with  $P_u(\{a, b\}) = P_l(\{a, b\}) = 0.81$ ,  $P_u(\{a\}) = P_l(\{a\}) = 0.09$  and  $P_l(\{b\}) = 0.1$  is an example for the upper bound, that is, we have  $P_u \models \Delta$  and  $P_u \models (b)[0.91]$ .

**Example 12.** Consider the knowledgebase  $\Delta_2 = \{(a)[0.9], (b | a)[0.8]\}$ , where we decreased the probability of  $(b | a)$ . Then PMP allows us to conclude that  $\Delta_2 \models (b)[0.72, 0.82]$ . While the probability naturally decreases, let us note that the width of the interval remains the same.

**Example 13.** Consider the knowledgebase  $\Delta_3 = \{(a)[0.8], (b | a)[0.9]\}$ . We started again from  $\Delta_1$ , but decreased the probability of  $(a)$  this time. Then PMP allows us to conclude that  $\Delta_3 \models (b)[0.72, 0.92]$ . Note that, in this case, the interval also becomes wider because there is a bigger chance now that  $a$  is false (in which case we cannot say anything about the probability of  $b$ ).

As the previous examples illustrate, our probability intervals can become wider when applying PMP. While this is formally desirable because the uncertainty in our conclusions increases, it can be a practical problem because the intervals can become meaningless. Since a general statement about the evolution of the intervals along a reasoning chain becomes very technical and difficult to understand, we just illustrate the issue with an example.

**Example 14.** Consider the knowledgebase  $\Delta_4 = \{(a)[0.9], (b | a)[0.9], (c | b)[0.9], (d | c)[0.9]\}$ . Using, the first two conditionals, we can conclude as before that  $\Delta_4 \models (b)[0.81, 0.91]$ . This derived conditional together with the third conditional in  $\Delta_4$  then allows us to conclude that  $\Delta_4 \models (c)[0.73, 1]$  (probabilities rounded to two digits). We can then again use this derived conditional together with the fourth conditional in  $\Delta_4$  to conclude that  $\Delta_4 \models (d)[0.66, 1]$ .

As the example illustrates, the intervals along reasoning chains become wider and wider. The growth of the interval depends on the uncertainty in our conditionals. If all conditionals are deterministic (probability 0 or 1), we basically apply the classical modus ponens and do not introduce any uncertainty at all. Roughly speaking, the closer the probabilities are to 0.5 (complete uncertainty), the faster the intervals will grow.

Naturally, we can improve the guarantees by making additional assumptions. The following *generalized probabilistic modus ponens* adds knowledge about  $\beta$  when  $\alpha$  is false. We abbreviate it by *GPMP*.

**Proposition 9 (GPMP).** *If  $\Delta \models (\alpha)[l_1, u_1]$ ,  $\Delta \models (\beta | \alpha)[l_2, u_2]$  and  $\Delta \models (\beta | \neg\alpha)[l_3, u_3]$  then  $\Delta \models (\beta)[l_4, u_4]$  where  $l_4 = l_1 \cdot l_2 + (1 - u_1) \cdot l_3$  and  $u_4 = \min\{1, u_1 \cdot u_2 + u_3 \cdot (1 - l_1)\}$ .*

**Proof.** Consider an arbitrary probability function  $P$  such that  $P \models \Delta$ . Then, for the lower bound, we have

$$\begin{aligned} P(\beta) &= P(\alpha \wedge \beta) + P(\neg\alpha \wedge \beta) \\ &\geq l_2 \cdot P(\alpha) + l_3 \cdot P(\neg\alpha) \\ &\geq l_2 \cdot l_1 + l_3 \cdot (1 - u_1). \end{aligned}$$

For the upper bound, we have

$$\begin{aligned} P(\beta) &= P(\alpha \wedge \beta) + P(\neg\alpha \wedge \beta) \\ &\leq u_2 \cdot P(\alpha) + u_3 \cdot P(\neg\alpha) \\ &\leq u_2 \cdot u_1 + u_3 \cdot (1 - l_1). \end{aligned}$$

Since we also have  $P(\beta) \leq 1$ , we can conclude  $P \models (\beta)[l_1 \cdot l_2 + (1 - u_1) \cdot l_3, \min\{1, u_1 \cdot u_2 + u_3 \cdot (1 - l_1)\}]$ .  $\square$

We call GPMP the generalized PMP because we can always assume that  $\Delta \models (\beta | \neg\alpha)[0, 1]$ . In this case, GPMP just corresponds to PMP. The interesting case is when we actually have non-trivial guarantees for  $(\beta | \neg\alpha)$ .

**Example 15.** Consider the knowledgebase  $\Delta_5 = \{(a)[0.9], (b | a)[0.9], (b | \neg a)[0.9]\}$ , which is similar to  $\Delta_1$ , but now also contains the knowledge that  $b$  is very likely even if  $a$  is false. Then GPMP allows us to conclude that  $\Delta_5 \models (b)[0.9]$ . This makes intuitively sense, since  $a$  is either true or false and in both cases the probability of  $b$  must be 0.9.

**Example 16.** To add some uncertainty, consider the knowledgebase  $\Delta_6 = \{(a)[0.9], (b | a)[0.9], (b | \neg a)[0.6]\}$ , where  $b$  is now less likely when  $a$  is false. In this case, GPMP allows us to conclude that  $\Delta_6 \models (b)[0.87]$ .

In the previous examples, we were always able to derive point probabilities from point probabilities. This is no coincidence, but follows immediately from GPMP.

**Corollary 1** (GPMP for point probabilities). *If  $\Delta \models (\alpha)[p_1]$ ,  $\Delta \models (\beta | \alpha)[p_2]$  and  $\Delta \models (\beta | \neg\alpha)[p_3]$  then  $\Delta \models (\beta)[p_4]$  where  $p_4 = p_1 \cdot p_2 + (1 - p_1) \cdot p_3$ .*

Another useful rule for classical argumentation is Modus Tollens. Recall that Modus Tollens states that if  $\alpha \rightarrow \beta$  is true and  $\beta$  is false, then  $\alpha$  must be false as well. We can also generalize the Modus Tollens slightly. The corresponding inference rule is called *probabilistic modus tollens* and abbreviated by *PMT* in the following.

**Proposition 10** (PMT). *If  $\Delta \models (\beta | \alpha)[l_2, u_2]$  and  $\Delta \models (\beta)[l_3, u_3]$ , then  $\Delta \models (\alpha)[0, u_1]$  where  $u_1 \leq \min\{u_2, u_3\} + (1 - l_2)$ .*

**Proof.** Consider an arbitrary probability function  $P$  such that  $P \models \Delta$ . We have  $P(\alpha) = P(\alpha \wedge \beta) + P(\alpha \wedge \neg\beta)$ . We can derive two upper bounds for  $P(\alpha \wedge \beta)$ . First, we have  $P(\alpha \wedge \beta) \leq u_2 \cdot P(\alpha) \leq u_2$  because  $\Delta \models (\beta | \alpha)[l_2, u_2]$ . Second, we have  $P(\alpha \wedge \beta) \leq P(\beta) \leq u_3$  because  $\Delta \models (\beta)[l_3, u_3]$ . Hence,  $P(\alpha) \leq \min\{u_2, u_3\} + P(\alpha \wedge \neg\beta)$ . CN allows us to conclude that  $\Delta \models (\neg\beta | \alpha)[1 - u_2, 1 - l_2]$ . Therefore,  $P(\alpha) \leq \min\{u_2, u_3\} + 1 - l_2$  and  $P \models (\alpha)[0, \min\{u_2, u_3\} + 1 - l_2]$ .  $\square$

We obtain the classical Modus Tollens for  $l_2 = u_2 = 1$  and  $l_3 = u_3 = 0$ . In this case, the statement is that  $P \models (\alpha)[0]$ . In general, if we have a strong positive rule (probability  $l_2$  close to 1) and its consequence has low probability, then the premise must also have a low probability.

Another rule that is frequently used in probabilistic reasoning is Bayes' rule. It allows us to switch the condition and conclusion of a conditional if the individual marginal probabilities are known. A typical application is medical diagnosis, where often statistics exists about the frequency of symptoms when patients suffer from a particular disease. However, medical doctors typically want to reason in the other direction: given symptoms, what is the probability that the patient suffers from a particular disease? We can design a corresponding inference rule that we abbreviate by *BR*.

**Proposition 11** (BR). *If  $\Delta \models (\beta | \alpha)[l_1, u_1]$ ,  $\Delta \models (\alpha)[l_2, u_2]$  and  $\Delta \models (\beta)[l_3, u_3]$  with  $l_3 > 0$ , then  $\Delta \models (\alpha | \beta)[l_4, u_4]$ , where  $l_4 = \frac{l_1 \cdot l_2}{u_3}$  and  $u_4 = \frac{u_1 \cdot u_2}{l_3}$ .*

**Proof.** Consider an arbitrary probability function  $P$  such that  $P \models \Delta$ . For the lower bound, we use the fact that  $P(\beta) \leq u_3$ , that is,  $\frac{P(\beta)}{u_3} \leq 1$ . We have

$$\begin{aligned} P(\alpha \wedge \beta) &\geq l_1 \cdot P(\alpha) \\ &\geq l_1 \cdot P(\alpha) \cdot \frac{P(\beta)}{u_3} \\ &\geq \frac{l_1 \cdot l_2}{u_3} \cdot P(\beta). \end{aligned}$$

Note that the fraction is well-defined because  $u_3 \geq l_3 > 0$ . For the upper bound we use the fact that  $\frac{P(\beta)}{l_3} \geq 1$  and get symmetrically

$$\begin{aligned} P(\alpha \wedge \beta) &\leq u_1 \cdot P(\alpha) \\ &\leq u_1 \cdot P(\alpha) \cdot \frac{P(\beta)}{l_3} \\ &\leq \frac{u_1 \cdot u_2}{l_3} \cdot P(\beta). \quad \square \end{aligned}$$

**Example 17.** Consider a test for a disease. Suppose the sensitivity of the test is 99%, and so for someone with the disease, the test is positive with probability 0.99, and suppose the specificity of the test is 98% and so for someone without the disease, the test is negative with probability 0.98. Also suppose that the prevalence of the disease is one in two hundred (i.e. the probability that someone has the disease is 0.005). And suppose that the probability of a positive test result is 0.02. We represent this by the following conditionals, where  $d$  denotes patient has disease,  $p/\neg p$  denotes positive/negative test result (for simplicity, we assume that the test is never inconclusive).

$$(p|d)[0.99] \quad (\neg p|\neg d)[0.98] \quad (d)[0.005] \quad (p)[0.02]$$

We can apply *BR* to derive the probability that a patient with a positive/negative test result has the disease. Applying *BR* to  $(p|d)[0.99]$ ,  $(d)[0.005]$  and  $(p)[0.02]$  yields  $(d|p)[0.2475]$ . That is, only around 25% of those who test positive have the disease (i.e. the positive predictive value is poor).

To understand the consequences of a negative test result, we first apply *CN* to our assumptions  $(p)[0.02]$  and  $(d)[0.005]$  to derive  $(\neg p)[0.98]$  and  $(\neg d)[0.995]$ . Then we can apply *BR* to the derived conditionals and our assumption  $(\neg p|\neg d)[0.98]$  to derive  $(\neg d|\neg p)[0.995]$ . That is, people who test negative almost never have the disease (i.e. the negative predictive value is very good).

In some applications, it can also be useful to have a rule that allows relaxing conditionals by increasing their probability intervals. One application of this is merging inconsistent beliefs of the form  $(F)[p_1]$ ,  $(F)[p_2]$ ,  $p_1 < p_2$  by replacing them with the conditional  $(F)[p_1, p_2]$  that relaxes both. We call the rule *Interval Relaxation* and abbreviate it by *IR*.

**Proposition 12 (IR).** *If  $\Delta \models (\beta | \alpha)[l_1, u_1]$ , then  $\Delta \models (\beta | \alpha)[l_2, u_2]$  for all  $l_2, u_2$  such that  $0 \leq l_2 \leq l_1 \leq u_1 \leq u_2 \leq 1$ .*

**Proof.** Consider an arbitrary probability function  $P$  such that  $P \models \Delta$ . Then, we have  $P(\alpha) \cdot l_2 \leq P(\alpha) \cdot l_1 \leq P(\alpha \wedge \beta) \leq P(\alpha) \cdot u_1 \leq P(\alpha) \cdot u_2$ . Hence,  $P \models (\beta | \alpha)[l_2, u_2]$ .  $\square$

### 3.3. Proof systems for probabilistic argumentation

We summarize our inference rules in Table 1.

Our soundness results from the previous section imply that we can use them to build up sound proof systems. The following definition summarizes some basic terminology and notation.

**Definition 10 (Proof System, Proof,  $\vdash$ ).**

1. A proof system  $PS = (Ax, R)$  for conditionals consists of a set  $Ax$  of conditionals whose elements are called *axioms* and a set  $R$  of (*inference*) *rules* for conditionals.
2. Let  $PS = (Ax, R)$  be a proof system for conditionals, let  $\Delta$  be a set of conditionals and let  $\phi$  be a conditional. A proof for  $\phi$  in  $PS$  using the assumptions in  $\Delta$  is a sequence  $\phi_1, \dots, \phi_n$  of conditionals such that  $\phi_n = \phi$  and for all  $i = 1, \dots, n$ , we have
  - $\phi_i \in Ax$  or
  - $\phi_i \in \Delta$  or
  - there is a rule  $\frac{\psi_1, \dots, \psi_l}{\phi_i} \in R$  such that  $\{\psi_1, \dots, \psi_l\} \subseteq \{\phi_1, \dots, \phi_{i-1}\}$ .
3. If there is a proof for  $\phi$  in  $PS$  using  $\Delta$ , we write  $\Delta \vdash_{PS} \phi$ . If  $\Delta = \emptyset$ , we just write  $\vdash_{PS} \phi$ .

In the following, we will use the following axiom set:

$$Ax = \{(F)[0, 1] \mid F \in \mathcal{L}(\mathcal{A})\} \cup \{(G|F)[0, 1] \mid F, G \in \mathcal{L}(\mathcal{A})\}. \quad (3)$$

The axioms basically state that all (conditional) probabilities are between 0 and 1. They are obviously sound. In the previous section, we also showed that our inference rules are sound. Therefore, we have the following guarantee.

**Theorem 1 (Sound Proof Systems).** *Every proof system that uses the axioms (3) and any subset of the inference rules from Table 1 is sound.*

In the following example, we illustrate the use of axioms.

**Example 18.** Consider a propositional language over  $\mathcal{A} = \{a, b\}$ . Suppose that  $\Delta = \{(a)[0.5]\}$  and that we are interested in the probability of  $a \wedge b$ . The only inference rules that allow us to infer something about the conjunction of two formulae are *CI* and *ICI*. If we do not want to make any independency assumptions, we can use only *CI*. Without axioms, we would not be able to infer anything. However, we can use the axiom  $(b)[0, 1]$ . Then applying *CI* to  $(a)[0.5]$  and  $(b)[0, 1]$  yields  $(a \wedge b)[0, 0.5]$ .

In order to build up arguments from a set of assumptions automatically, we can consider proof systems of increasing complexity to build up increasingly complex arguments. In the remainder of this paper, we will consider the following examples of increasingly complex proof systems:

- **SIMPLE** =  $(Ax, \{PMP\})$ ,
- **STAR** =  $(Ax, \{PMP, CI\})$ ,

**Table 1**  
Summary of Inference Rules. We write  $\beta \perp \gamma \mid \alpha$  for the assumption that  $\beta$  is conditionally independent of  $\gamma$  given  $\alpha$  (ICI).

Name	Rule	Remarks
CN	$\frac{(\beta \mid \alpha)[l, u]}{(\neg\beta \mid \alpha)[1 - u, 1 - l]}$	
CE	$\frac{(\beta \wedge \gamma \mid \alpha)[l, u]}{(\beta \mid \alpha)[l, 1]}$	
DE	$\frac{(\beta \vee \gamma \mid \alpha)[l, u]}{(\beta \mid \alpha)[0, u]}$	
CI	$\frac{(\beta \mid \alpha)[l_1, u_1] \quad (\gamma \mid \alpha)[l_2, u_2]}{(\beta \wedge \gamma \mid \alpha)[l_3, u_3]}$	$l_3 = \max\{0, l_1 + l_2 - 1\},$ $u_3 = \min\{u_1, u_2\}$
DI	$\frac{(\beta \mid \alpha)[l_1, u_1] \quad (\gamma \mid \alpha)[l_2, u_2]}{(\beta \vee \gamma \mid \alpha)[l_3, u_3]}$	$l_3 = \max\{l_1, l_2\},$ $u_3 = \min\{1, u_1 + u_2\}$
ICI	$\frac{(\beta \mid \alpha)[l_1, u_1] \quad (\gamma \mid \alpha)[l_2, u_2]}{(\beta \wedge \gamma \mid \alpha)[l_3, u_3]}$	assuming $\beta \perp \gamma \mid \alpha,$ $l_3 = l_1 \cdot l_2,$ $u_3 = u_1 \cdot u_2$
PMP	$\frac{(\alpha)[l_1, u_1] \quad (\beta \mid \alpha)[l_2, u_2]}{(\beta)[l_3, u_3]}$	$l_3 = l_1 \cdot l_2,$ $u_3 = \min\{1, u_1 \cdot u_2 + 1 - l_1\}$
GPMP	$\frac{(\alpha)[l_1, u_1] \quad (\beta \mid \alpha)[l_2, u_2] \quad (\beta \mid \neg\alpha)[l_3, u_3]}{(\beta)[l_4, u_4]}$	$l_4 = l_1 \cdot l_2 + (1 - l_1) \cdot l_3,$ $u_4 = \min\{1, u_1 \cdot u_2 + u_3(1 - l_1)\}$
GPMP for point probabilities	$\frac{(\alpha)[p_1] \quad (\beta \mid \alpha)[p_2] \quad (\beta \mid \neg\alpha)[p_3]}{(\beta)[p_4]}$	$p_4 = p_1 \cdot p_2 + (1 - p_1) \cdot p_3$
PMT	$\frac{(\beta \mid \alpha)[l_2, u_2] \quad (\beta)[l_3, u_3]}{(\alpha)[0, u_1]}$	$u_1 \leq \min\{u_2, u_3\} + (1 - l_2)$
BR	$\frac{(\beta \mid \alpha)[l_1, u_1] \quad (\alpha)[l_2, u_2] \quad (\beta)[l_3, u_3]}{(\alpha \mid \beta)[l_4, u_4]}$	assuming $l_3 > 0,$ $l_4 = \frac{l_1 \cdot l_2}{u_3},$ $u_4 = \frac{u_1 \cdot u_2}{l_3}$
IR	$\frac{(\beta \mid \alpha)[l_1, u_1]}{(\beta \mid \alpha)[l_2, u_2]}$	for all $l_2, u_2$ such that $0 \leq l_2 \leq l_1 \leq u_1 \leq u_2 \leq 1$

- BRAN = (Ax, {PMP, CI, BR}).

The SIMPLE proof system provides a very simple way of drawing conclusions from conditionals using the probabilistic modus ponens. STAR extends SIMPLE with conjunction introduction, which is useful for combining atomic pieces of knowledge. BRAN extends STAR with the ability to undertake Bayesian reasoning.

### 3.4. Computing proofs

Before moving to probabilistic argumentation, let us look at some interesting computational problems that occur in applications of our proof systems. Our focus is on problems that are useful for probabilistic argumentation and that can be solved in a more comprehensible way than when using the linear optimization approach.

One advantage of a proof system is that we can generate proofs that can be checked by the user to understand the inference. One very basic problem is thus to automatically verify that a proof generated by our system is correct.

**Definition 11 (Proof Verification).** Given a proof system PS, a set of conditionals  $\Delta$  and a sequence of conditionals  $\phi_1, \dots, \phi_n$ , the problem of deciding if  $\phi_1, \dots, \phi_n$  is a proof for  $\phi_n$  in PS using  $\Delta$  is called *Proof Verification*.

The algorithm shown in Fig. 1 can solve the verification problem in polynomial time.

```

Algorithm Verify( $(Ax, R), \Delta, (\phi_1, \dots, \phi_n)$ ):
  for  $i=1$  to  $n$ :
    if  $\phi_i \notin (Ax \cup \Delta)$ :
      if there is no rule in  $R$  that allows inferring  $\phi_i$  from  $\{\phi_1, \dots, \phi_{i-1}\}$  :
        return false
  return true

```

**Fig. 1.** Verification algorithm that decides in polynomial-time if  $\phi_1, \dots, \phi_n$  is a proof for  $\phi_n$ .

**Proposition 13.** For each proof system that uses the axioms defined in (3) and inference rules from Table 1, Proof Verification can be solved in polynomial time. The worst-case runtime is  $O((|Ax| + |\Delta|) \cdot n + |R| \cdot n^{k+1})$ , where  $k$  is the largest number of assumptions in a rule.

**Proof.** Consider the Algorithm shown in Fig. 1. It basically goes through the sequence of conditionals and verifies that they follow from the axioms, assumptions or rules using the previous conditionals. If one conditional does not follow, the algorithm returns false. So it is correct by definition.

Checking if a conditional corresponds to an axiom or an assumption can be done in time  $O(|Ax| + |\Delta|)$ . If this is not the case, in the worst-case, we have to check for every inference rule if it can be used to derive  $\phi_n$  from  $\phi_1, \dots, \phi_{n-1}$ . For this, we have to check at most  $\binom{n}{k} = O(n^k)$  possible instantiations of the assumptions. Doing this for all rules results in  $O(|R| \cdot n^k)$  time. The overall runtime is then  $O(\sum_{i=0}^n (|Ax| + |\Delta| + |R| \cdot n^k)) = O((|Ax| + |\Delta|) \cdot n + |R| \cdot n^{k+1})$ , where  $k$  is the largest number of assumptions in a rule. Dependent on the selected rules, we have  $k \in \{1, 2, 3\}$ .  $\square$

For example, we have  $k = 2$  for SIMPLE and STAR and  $k = 3$  for BRAN.

An important question is then, how can we find proofs? We can think about the search space as a layered directed graph where the first layer consists of the axioms and assumptions. The next layer is then formed by applying inference rules to the conditionals that occur in previous layers such that at least one conditional is from the immediately preceding layer. Edges indicate which conditionals have been used for the inference. However, in principle, it is possible that the same conditional can be derived from two different rules. To avoid ambiguity, we therefore consider a labeled graph, where edges are labeled with the rule that has been applied.

**Definition 12 (Proof Graph).** Given a proof system  $PS = (Ax, R)$  and a set of conditionals  $\Delta$  (the assumptions), the associated proof graph is a layered, labeled directed graph  $(V, E, L)$ ,  $V = \bigcup_{i=0}^{\infty} V_i$ ,  $L : E \rightarrow R$  that is inductively defined as follows:

1.  $V_0 = \Delta \cup Ax$ ,
2. For all  $k \in \mathbb{N}_0$ ,  $V_{k+1}$  is defined as the set of all conditionals  $\phi = (\beta \mid \alpha)[l, u]$  such that we have

**Uniqueness**  $\phi \notin \bigcup_{j=0}^k V_j$  and

**Derivability** there is a rule  $\frac{\psi_1, \dots, \psi_m}{\phi}$  in  $R$  such that  $\psi_i \in \bigcup_{j=0}^k V_j$  for  $1 \leq i \leq m$ .

3. For every rule  $r : \frac{\psi_1, \dots, \psi_m}{\phi}$  that has been used to add a conditional  $\phi$  to  $V_{k+1}$ , we add corresponding edges  $\{(\psi_i, \phi) \mid 1 \leq i \leq m\}$  to  $E$  and label all these edges with  $r$ .

We define the depth of the proof graph as the index of the last non-empty layer, that is, as  $\sup\{i \in \mathbb{N}_0 \mid V_i \neq \emptyset\}$ .

By construction, the proof graph does not contain duplicates of conditionals (*Uniqueness*) and conditionals occur in order of their derivability (*Derivability*).

For every conditional  $\phi \in V$  in the proof graph, we let  $JR(\phi) = \{r \in R \mid \exists(\psi, \phi) \in E : L((\psi, \phi)) = r\}$  denote the rules that can be used to derive  $\phi$ . By going backwards from  $\phi$  following the edges that justify  $\phi$  and its predecessors, we can construct one or multiple proofs for  $\phi$ . We let  $\text{Proofs}(\phi)$  denote the set of all proofs for  $\phi$  that can be constructed in this way. Note that  $\text{Proofs}(\phi) = \{\phi\}$  (the only proof for  $\phi$  in the proof graph is  $\phi$  itself) if and only if  $\phi \in V_0$ .

Let us make some simple observations about the structure of the proof graph that follow from the construction.

**Proposition 14.**

1. *Non-triviality of derived conditionals:* If  $(\beta \mid \alpha)[l, u] \in V_i$  for  $i > 0$ , then  $l > 0$  or  $u < 1$  (or both).
2. *Layer-connectivity:* If  $\phi \in V_{i+1}$ , then there is a  $\psi \in V_i$  such that  $(\psi, \phi) \in E$ .

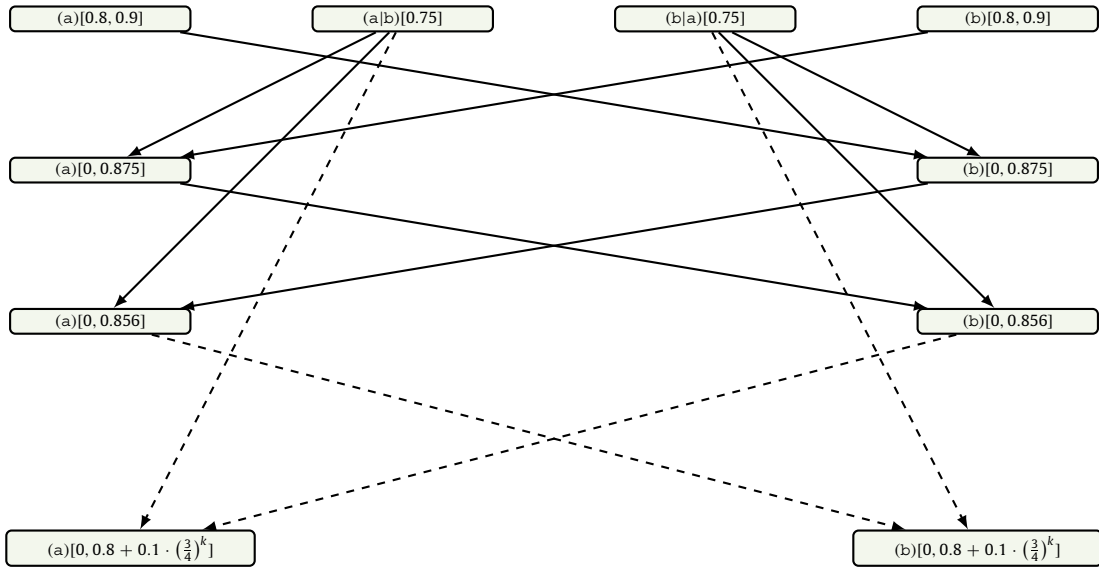


Fig. 2. Infinite subgraph of the proof graph for Example 19.

**Proof.** 1. If  $l = 0$  and  $u = 1$ , then  $(\beta \mid \alpha)[0, 1]$  is an axiom and therefore contained in  $V_0$ . By Uniqueness, it cannot occur in  $V_i$  for  $i > 0$ .

2. For the sake of contradiction, assume that there is no such  $\psi \in V_l$ . Then there must be an inference rule  $\frac{\psi_1, \dots, \psi_l}{\phi}$  such that  $\psi_i \in \bigcup_{j=0}^{l-1} V_j$  for  $1 \leq i \leq l$ . But then  $\phi$  is derivable in an earlier layer and we must have  $\phi \in V_j$  for some  $j \leq l$  contradicting Uniqueness.  $\square$

The first item states that only the first layer can contain trivial conditionals (because they are axioms). As we illustrated in Example 18, they can be useful in proofs, but should not be derived. The second item states that every conditional in a new layer must use some new information that has been added in the previous layer.

If the proof graph was always finite, it would be easy to generate upper bounds for the complexity of finding a proof. Unfortunately, this is not necessarily the case. Let us first note that the first layer of the proof graph is already infinitely large. This is because there is an infinite number of propositional formulae and we have an axiom for every formula that states that the probability is between 0 and 1. Subsequent layers can be infinitely large as well dependent on the inference rules that we consider. For example, *IR* allows deriving an infinite (even uncountable) number of inferences from every non-trivial inference in our proof graph. For example, consider a derived conditional  $(\beta \mid \alpha)[l, u]$  such that  $l > 0$ . Then, *IR* allows deriving all conditionals in the (uncountable) set  $\{(\beta \mid \alpha)[l', u] \mid 0 < l' < l\}$ . Similarly, rules like *CI* and *DI* that extend formulae can result in an infinite number of inferences. For example, consider again a derived conditional  $(\beta \mid \alpha)[l, u]$  with  $l > 0$ . Then, for every propositional formula  $\gamma$ , *DI* and the axiom  $(\gamma \mid \alpha)[0, 1]$  allow inferring  $(\beta \vee \gamma \mid \alpha)[l, 1]$ .

What is perhaps more surprising is that the proof graph can also be infinitely deep even if we do not apply any axioms and use only PMP as an inference rule. We illustrate this in the following example.

**Example 19.** Consider a propositional language over  $\mathcal{A} = \{a, b\}$ . Suppose that

$$\Delta = \{(a)[0.8, 0.9], (b)[0.8, 0.9], (a \mid b)[0.75], (b \mid a)[0.75]\}.$$

Note that we have completely symmetric knowledge about  $a$  and  $b$ . Therefore, everything that we can derive about  $a$  can be derived symmetrically for  $b$ .

$V_0$  contains  $\Delta$ . We can apply PMP to  $(a)[0.8, 0.9]$  and  $(b \mid a)[0.75]$  to derive  $(b)[0, 0.875]$ . Symmetrically, we can derive  $(a)[0, 0.875]$  and add both to  $V_1$ .

In  $V_1$ , we can now use  $(a)[0, 0.875]$  and  $(b \mid a)[0.75]$  to derive  $(b)[0, 0.85625]$ . Symmetrically, we can derive  $(a)[0, 0.85625]$  and add both to  $V_2$ .

In this way, we are creating a sequence of conditionals  $(a)[0, u_k]$  for  $V_k$ , where  $u_0 = 0.9$  and  $u_k = 0.75 \cdot u_{k-1} + 0.2$  for  $k > 0$ . As we explain below, this sequence converges to 0.8. However, it will never actually take this value and so the proof graph is infinitely deep because every subsequent layer contains a tighter upper bound for  $a$  (and  $b$ ). We illustrate the infinite subgraph in Fig. 2.

For completeness, we explain why the sequence converges to 0.8 without ever actually taking this value. This part can be skipped and is not important for the following. One can show by induction that  $u_k = 0.9 \cdot 0.75^k + 0.2 \cdot \sum_{i=0}^{k-1} 0.75^i$ . Note that

$S_k = \sum_{i=0}^{k-1} 0.75^i$  is a geometric series and therefore  $S_k = \frac{1-0.75^k}{1-0.75} = 4 \cdot (1-0.75^k)$ . Hence,  $u_k = 0.9 \cdot 0.75^k + 0.2 \cdot 4 \cdot (1-0.75^k) = 0.75^k \cdot (0.9 - 0.8) + 0.8 = 0.8 + 0.1 \cdot 0.75^k$ . Since  $0.1 \cdot 0.75^k$  is strictly monotonically decreasing,  $u_k$  is strictly monotonically decreasing. Furthermore,  $u_k > 0.8$  for all  $k \in \mathbb{N}$  and  $\lim_{k \rightarrow \infty} u_k = 0.8$ .

We can apply a linear programming solver to confirm that the knowledgebase in the previous example entails (a)[0.8]. However, there does not exist a proof for this statement in SIMPLE, even though we can come arbitrarily close. The fact that a proof does not exist is not surprising since our proof systems are not complete. However, it is interesting to note that statements that can be proved can have arbitrarily long proofs. In order to derive the conditional (a)[0,  $u_k$ ] in our previous example, we need  $2k - 1$  inference steps. This means, in particular, that “short” proofs (e.g., with length polynomial in the size of the knowledge base) for a conditional may not exist even if it can be proved in the proof system. However, let us note that, by construction, the proofs in the proof graph are never unnecessarily verbose. That is, when removing any assumptions in the proof, it is no longer a proof.

**Proposition 15.** *For every conditional  $\phi$  in the proof graph and for every proof  $\psi_1, \dots, \psi_n, \phi$  in  $\text{Proofs}(\phi)$ , there does not exist another proof  $\psi'_1, \dots, \psi'_n, \phi$  in  $\text{Proofs}(\phi)$  such that  $\{\psi'_1, \dots, \psi'_n\} \subsetneq \{\psi_1, \dots, \psi_n\}$ .*

**Proof.** The claim follows by induction over the depth  $d$  of  $\phi$ . If  $d = 0$ , then  $\phi$  must be an axiom or an assumption in  $\Delta$  and the only proof is  $\phi$  itself.

For the induction step, suppose that the claim is true for all  $d \leq N$  and consider a conditional  $\phi$  at depth  $N + 1$ . For the sake of contradiction, assume that there is another proof  $\psi'_1, \dots, \psi'_n, \phi$  in  $\text{Proofs}(\phi)$  such that  $\{\psi'_1, \dots, \psi'_n\} \subsetneq \{\psi_1, \dots, \psi_n\}$ . Let  $k$  be the smallest index such that  $\{\psi'_1, \dots, \psi'_k\} = \{\psi_1, \dots, \psi_k\}$  and  $\{\psi'_1, \dots, \psi'_{k+1}\} \neq \{\psi_1, \dots, \psi_{k+1}\}$ . Then  $\psi'_1, \dots, \psi'_k, \psi'_{k+1}$  is a proof for  $\psi'_{k+1}$ . By induction assumption, there can be no extension of this proof in the proof graph. Hence, we must have  $\psi'_{k+1} = \psi_{k+1}$  and therefore  $\{\psi'_1, \dots, \psi'_{k+1}\} = \{\psi_1, \dots, \psi_{k+1}\}$ , which is a contradiction.  $\square$

In order to build up probabilistic arguments from a given set of assumptions, we can consider a restricted forward search, where we start from the assumptions and generate new conditionals successively by applying selected inference rules. This corresponds to generating a subset of the proof graph. Axioms should not be generated explicitly, but only be used implicitly when they allow making a non-trivial inference as in Example 18. When applying forward search, we may also avoid creating longer and longer conditionals by excluding inference rules like *CN*, *CI* and *DI* and *ICI* or using them only in restricted ways. The way we restrict their use depends on the application and the structure of our assumptions  $\Delta$ . For example, if  $\Delta$  contains only Horn-like conditionals with a single literal in the head and a (possibly empty) conjunction of literals in the body, then we could allow only inferring conditionals whose head and body is a (possibly empty) conjunction of literals.

While forward search is well suited to build up arguments bottom-up, sometimes we may be interested in deriving a non-trivial bound for a particular qualitative conditional ( $\beta|\alpha$ ), that is, finding a probability interval  $[l, u] \subseteq [0, 1]$  such that  $\Delta \vdash_{\text{ps}} (\beta|\alpha)[l, u]$ . In this case, backward search is better suited. That is, we try to find an inference rule that allows us to infer something interesting about ( $\beta|\alpha$ ) and continue recursively with the premises of the rule until all non-instantiated conditionals have been replaced with assumptions or axioms, or we find that we cannot derive anything interesting.

However, as the proof graphs can, in general, be infinitely wide and deep, heuristics are needed for both forward and backward search. One natural idea is to order the inference rules by their comprehensibility and relevance for the application. One way to do this is to start from a minimal proof systems and to extend it incrementally if more inferences are needed. For example, if we want to use BRAN in our application, we could start finding proofs for SIMPLE, then go to STAR and then to BRAN if no proof can be found in the simpler proof systems.

## 4. Probabilistic argumentation

We now turn to probabilistic argumentation. In the following subsections, we consider a notion of a probabilistic argument, and how we can construct them using a proof system. We then consider types of argument graph and properties of them.

### 4.1. Arguments

From now on, we consider an argument as a set of consistent probability statements (premises) and an inference from the those premises (a claim).

**Definition 13.** Let  $\Phi \subseteq \mathcal{L}(\mathcal{A})$ , and  $\psi \in \mathcal{L}(\mathcal{A})$ .  $\langle \Phi, \psi \rangle$  is a **protoargument** iff (1)  $\Phi \models \psi$  and (2)  $\text{Sat}(\Phi) \neq \emptyset$ .  $\langle \Phi, \psi \rangle$  is an **argument** iff (1)  $\langle \Phi, \psi \rangle$  is a protoargument and (2) there is no  $\Phi'$  s.t.  $\Phi' \subset \Phi$  and  $\Phi' \models \psi$ .

So a protoargument has premises that are sufficient to entail the claim and the premises are consistent, and an argument is a protoargument that also ensures that there is no subset of premises that already entails the claim.



**Example 20.** If  $\{(a|b)[p], (b)[q]\} \subseteq \Delta$ , then *PMP for point probabilities* implies that the following is a protoargument, where  $r = \min\{1, p \cdot q + 1 - q\}$ .

$$\langle\{(a|b)[p], (b)[q]\}, (a)[p \cdot q, r]\rangle$$

Clearly, both  $(a|b)[p]$  and  $(b)[q]$  are necessary to derive the claim, so the protoargument is also an argument.

**Example 21.** If  $\{(a|b)[0.8], (b)[0.7]\} \subseteq \Delta$ , then, as in the previous example, *PMP for point probabilities* implies that the following is an argument.

$$\langle\{(a|b)[0.8], (b)[0.7]\}, (a)[0.56, 0.86]\rangle$$

Let us note that every node in a proof graph corresponds to a protoargument.

**Proposition 16.** Let  $PS = (Ax, R)$  be a sound proof system and let  $\Delta$  be a satisfiable set of conditionals. For every node  $\phi$  in the associated proof graph  $(V, E, l)$ , every proof  $(\phi_1, \dots, \phi_n) \in \text{Proofs}(\phi)$  ( $\phi = \phi_n$ ) in the graph corresponds to a protoargument  $\langle\{\phi_1, \dots, \phi_{n-1}\}, \phi\rangle$ .

**Proof.** The definition of a proof graph and soundness of  $PS$  imply that  $(\phi_1, \dots, \phi_n)$  is a valid proof for  $\phi$ , that is,  $\{\phi_1, \dots, \phi_{n-1}\} \models \phi$ . Furthermore, satisfiability of  $\Delta$  and soundness of  $PS$  imply that  $\{\phi_1, \dots, \phi_{n-1}\}$  is still satisfiable. Hence,  $\langle\{\phi_1, \dots, \phi_{n-1}\}, \phi\rangle$  is a protoargument.  $\square$

Conceptually, it is possible that  $\text{Proofs}(\phi)$  contains two proofs such that one is a shorter version of the other (using different inference rules). In this case, the longer version is a protoargument, but not an argument. In fact, even the shortest proof may not correspond to an argument because it is possible that our inference rules are not sufficient to find a minimal set of conditionals that (semantically) entail  $\phi$ . We illustrate this in the following example.

**Example 22.** Consider the knowledgebase  $\Delta = \{(a)[0.2], (\neg a | a)[0]\}$  and suppose that we use the proof system *SIMPLE* that contains only the inference rule *PMP*. Then the proof graph contains the proof  $(a)[0.2], (\neg a | a)[0], (\neg a)[0, 0.8]$ . However, while  $\langle\{(a)[0.2], (\neg a | a)[0]\}, (\neg a)[0, 0.8]\rangle$  is a protoargument, it is not an argument because  $(a)[0.2]$  already entails  $(\neg a)[0.8]$  and therefore also  $(\neg a)[0, 0.8]$ . Note that when we add the inference rule *CN*, the proof graph does indeed contain the shorter proof  $(a)[0.2], (\neg a)[0, 0.8]$  instead of the previous one.

From the perspective of an agent who is only aware of some inference rules, an argument may not be comprehensible if it is not derivable in the corresponding proof system. Therefore, we introduce arguments relative to a proof system.

**Definition 14.** Let  $\vdash_{PS}$  be the consequence relation for a sound proof system  $PS$ . Also let  $\Phi \subseteq \mathcal{L}(\mathcal{A})$ , and  $\psi \in \mathcal{L}(\mathcal{A})$ ,  $\langle\Phi, \psi\rangle$  is a **PS protoargument** iff (1)  $\Phi \vdash_{PS} \psi$  and (2)  $\text{Sat}(\Phi) \neq \emptyset$ .  $\langle\Phi, \psi\rangle$  is a **PS argument** iff (1)  $\langle\Phi, \psi\rangle$  is a PS protoargument and (2) there is no  $\Phi'$  s.t.  $\Phi' \subset \Phi$  and  $\Phi' \vdash_{PS} \psi$ .

Let us note that PS protoarguments are always protoarguments.

**Proposition 17.** For all sound proof systems  $PS$  and arguments  $A$ , if  $A$  is a PS protoargument, then  $A$  is a protoargument.

**Proof.** Let  $\vdash_{PS}$  be the consequence relation for  $PS$ . Assume  $A$  is a PS-protoargument. So  $\Phi \vdash_{PS} \psi$  and  $\text{Sat}(\Phi) \neq \emptyset$ . By soundness of  $PS$ , we have  $\Phi \models \psi$ . Hence,  $A$  is a protoargument.  $\square$

However, not every protoargument is a PS protoargument. For instance, in Example 22, the following is a protoargument (and also an argument), but not a *SIMPLE* protoargument.

$$\langle\{(a)[0.2]\}, (\neg a)[0, 0.8]\rangle$$

Let us also note that arguments and PS arguments are incomparable in the sense that none is a subset of the other. For instance, in Example 22,  $\langle\{(a)[0.2], (\neg a | a)[0]\}, (\neg a)[0, 0.8]\rangle$  is a *SIMPLE* argument, but not an argument and  $\langle\{(a)[0.2]\}, (\neg a)[0, 0.8]\rangle$  is an argument, but not a *SIMPLE* argument.

Proposition 15 implies that arguments for a given knowledgebase and proof system can be generated by building up the corresponding proof graph.

**Corollary 2.** Let  $PS = (Ax, R)$  be a sound proof system and let  $\Delta$  be a satisfiable set of conditionals. For every node  $\phi$  in the associated proof graph  $(V, E, l)$ , every proof  $(\phi_1, \dots, \phi_n) \in \text{Proofs}(\phi)$  ( $\phi = \phi_n$ ) in the graph corresponds to a PS protoargument  $\langle\{\phi_1, \dots, \phi_{n-1}\}, \phi\rangle$ .

We introduce some additional terminology that will be useful in the following. For an argument  $A = \langle \Phi, \psi \rangle$ , we refer to  $\Phi$  as the *support* of  $A$  and denote it by  $\text{Support}(A)$ . We refer to  $\psi$  as the *claim* of  $A$  and denote it by  $\text{Claim}(A)$ .

**Example 23.** For  $A = \langle \{(a|b)[0.8], (b)[0.7]\}, (a)[0.56, 0.86] \rangle$ ,  $\text{Support}(A) = \{(a|b)[0.8], (b)[0.7]\}$ , and  $\text{Claim}(A) = (a)[0.56, 0.86]$ .

A key advantage of using a proof system is that we can explain the inference of a claim of an argument from the support of the argument in terms of specific rules. If we use the linear programming approach, it can be difficult to understand how the claim follows from the support. Hence, using a proof system makes the argumentation process more transparent.

#### 4.2. Argument graphs

We assume that a knowledgebase can be inconsistent. Therefore, we may obtain arguments from the knowledgebase that have supports that are inconsistent with each other. Therefore, we need to consider counterarguments. We define an argument being counterargument to another argument in terms of unsatisfiability.

**Definition 15.** The following are types of **attack** where  $A$  and  $B$  are arguments.

- $A$  is a **rebuttal** of  $B$  if  $\text{Sat}(\{\text{Claim}(A), \text{Claim}(B)\}) = \emptyset$ .
- $A$  is a **direct undercut** of  $B$  if  $\exists \psi \in \text{Support}(B)$  s.t.  $\text{Sat}(\{\text{Claim}(A), \psi\}) = \emptyset$ .
- $A$  is an **undercut** of  $B$  if  $\exists \Gamma \subseteq \text{Support}(B)$  s.t.  $\text{Sat}(\{\text{Claim}(A)\} \cup \Gamma) = \emptyset$ .

We may refer to rebuttal by *rb*, to undercut by *un*, and to direct undercut by *du*.

So for a pair of arguments, an attack is either based on a conflict between the heads, or the head of one and the body of the other.

**Example 24.** The following are examples of counterarguments.

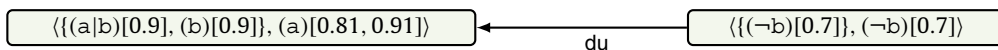
- $\langle \{(a|b)[0.8], (b)[0.7]\}, (a)[0.56, 0.86] \rangle$  is a rebuttal of  $\langle \{(-a|c)[0.9], (c)[0.9]\}, (-a)[0.81, 0.91] \rangle$  because the claims are not satisfiable together.
- $\langle \{(a)[0.5, 0.8]\}, (a)[0.5, 0.6] \rangle$  is a rebuttal of  $\langle \{(a)[0.7, 0.9]\}, (a)[0.7, 0.9] \rangle$  because the claims are not satisfiable together.
- $\langle \{(-b)[0.8]\}, (-b)[0.8] \rangle$  is a direct undercut of  $\langle \{(a|b)[0.9], (b)[0.9]\}, (a)[0.81, 0.91] \rangle$  because the claim of the attacker is not satisfiable with the premise  $(b)[0.9]$  in the attackee.
- $\langle \{(a|b)[0.2]\}, (a|b)[0.2] \rangle$  is an undercut of  $\langle \{(c|a)[1], (a|b)[1], (b)[1]\}, (a)[1] \rangle$  because the claim of the attacker is not satisfiable with the premises  $(a|b)[1]$  and  $(b)[1]$  in the attackee.
- $\langle \{(-b|d)[0.9], (d)[0.9]\}, (-b)[0.81, 0.91] \rangle$  is an undercut of  $\langle \{(a|b)[1], (b|c)[1], (c)[1]\}, (a)[1] \rangle$  because the claim of the attacker is not satisfiable with the premises  $\{(b|c)[1], (c)[1]\}$  in the attackee, since these premises imply  $(b)[1]$  which is not satisfiable with the claim  $(-b)[0.81, 0.91]$ .

Let us note the following relationships between the different attack relations.

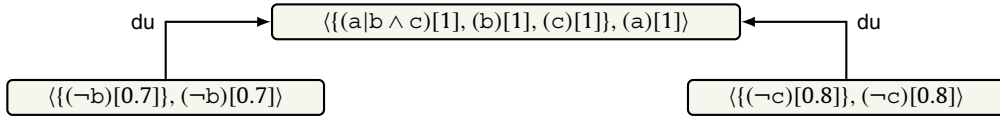
**Proposition 18.** Let  $A$  and  $B$  be arguments: (1) If  $A$  is a rebuttal of  $B$ , then  $A$  is an undercut of  $B$ ; (2) If  $A$  is a direct undercut of  $B$ , then  $A$  is an undercut of  $B$ .

Based on our different types of attacks, we can associate every knowledgebase  $\Delta$  consisting of probabilistic conditionals with an argument graph  $G = (\mathcal{N}, \mathcal{R})$ , where for each argument  $A \in \mathcal{N}$ ,  $\text{Support}(A) \subseteq \Delta$ , and  $(A, B) \in \mathcal{R}$  if there is an attack (rebuttal/(direct) undercut) from  $A$  to  $B$ . We do not impose the condition that for every subset of  $\Gamma \subseteq \Delta$ , there is an argument  $A \in \mathcal{N}$ , such that  $\text{Support}(A) = \Gamma$ . So a knowledgebase can be associated with more than one argument graph (e.g. Examples 25 and 28).

**Example 25.** The following is an argument graph constructed using SIMPLE arguments from  $\Delta = \{(a|b)[0.9], (b)[0.9], (-b)[0.7]\}$ . Using dialectical semantics (Definition 9), the argument at the leaf is the only acceptable argument in grounded, preferred, and stable, extensions.



**Example 26.** The following is an argument graph constructed using STAR arguments from  $\Delta = \{(a|b \wedge c)[1], (b)[1], (c)[1], (a)[1], (-b)[0.7], (-c)[0.8]\}$ . Using dialectical semantics (Definition 9), the arguments at the leaves are the acceptable arguments in grounded, preferred, and stable, extensions.



Given a knowledgebase, a type of argument (SIMPLE argument, STAR argument, etc), and a type of attack (rebuttal, undercut, direct undercut), we can use the following definition of an exhaustive argument graph which is an argument graph that contains every argument of the specified type from the knowledge and every attack of the specified type.

**Definition 16.** For proof system PS, let  $Args_{PS}(\Delta)$  be the set of PS arguments formed from  $\Delta$  defined as follows.

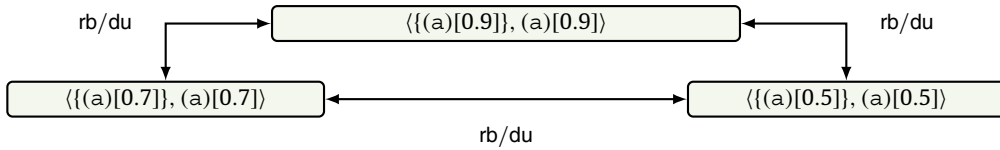
$$Args_{PS}(\Delta) = \{ \langle \Phi, \psi \rangle \mid \Phi \subseteq \Delta \text{ and } \Phi \neq \emptyset \text{ and } \Phi \vdash_{PS} \psi \text{ and there is no } \Phi' \subset \Phi \text{ s.t. } \Phi' \vdash_{PS} \psi \}$$

For a proof system PS, and  $\rho \in \{rb, du, un\}$ , let  $Attacks_{\rho}^{PS}(\Delta)$  be the set of attacks of type  $\rho$  formed from  $\Delta$  (i.e.  $Attacks_{\rho}^{PS}(\Delta) = \{A, B \in Args_{PS}(\Delta) \mid A \text{ is a } \rho \text{ attack of } B\}$ ). An **exhaustive graph** formed from  $\Delta$ , denoted  $G_{\rho}^{PS}(\Delta)$ , is the argument graph  $(Args_{PS}(\Delta), Attacks_{\rho}^{PS}(\Delta))$ .

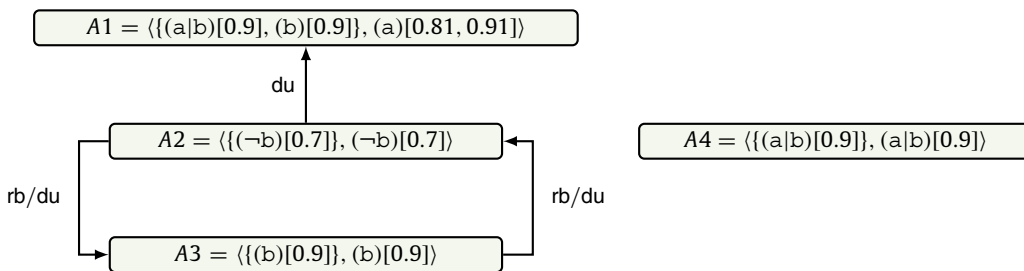
We assume that exhaustive graphs can be constructed by applying the inference rules for a proof system exhaustively. However, with some combinations of inference rules, and specific knowledgebases, there can be an infinite number of arguments (e.g., Example 19).

We illustrate exhaustive graphs with the following examples. Note, in the examples of exhaustive graphs, we will exclude arguments corresponding to axioms. So for example, we do not include  $\emptyset \vdash_{PS} (\phi \mid \psi)[0, 1]$ . We exclude them because they are not of relevance when considering conflicting evidence. They are tautologies and do not provide any useful information for the topic of discussion.

**Example 27.** For  $\Delta = \{(a)[0.9], (a)[0.7], (a)[0.5]\}$ , the following is the exhaustive argument graph using STAR arguments, and rebuttal, or undercut, or direct undercut.



**Example 28.** For  $\Delta = \{(a|b)[0.9], (b)[0.9], (\neg b)[0.7]\}$ , the following is the exhaustive argument graph using STAR arguments, and undercut, or direct undercut.



So the argument graph provides a transparent representation of the arguments and counterarguments whether they have produced exhaustively or through some selection process.

### 4.3. Analyzing argument graphs

In order to analyze the argument graphs, we can use dialectical semantics as we will discuss in Section 4.3.1 or we can use the epistemic approach as we will discuss in Section 4.3.2.

Whether we use the dialectical or epistemic approaches, there is a form of incompleteness when using a proof system PS (i.e. there may be inferences that are valid according to the semantics but that are not obtained by PS). This raises two issues that we would like to consider next.

The first issue is the *potential incompleteness of the argument graph* when using a proof system. When we construct an argument graph, then all the arguments in the graph are constructed using PS. So if PS is incomplete, then there may

be inferences that are valid according to the semantics but that are not obtained by a proof system PS. Hence there are arguments that can be obtained by the semantics but cannot be obtained as arguments when using PS. While a complete proof system would be interesting, we believe that even an incomplete system is useful for many applications. Let us first note that we can always perform complete reasoning by using the linear programming approach to solve the probabilistic satisfiability or entailment problem as highlighted in Proposition 1. Theoretically, it could be used to exhaustively test all potential arguments and to add them to the graph. However, since there is typically an (uncountably) infinite number of arguments, this is not a practical approach. Proof systems allow us to infer arguments in a more systematic manner by applying a forward search that successively extends the knowledge base with inferred conditionals. Since our proof systems are equipped with quite intuitive inference rules, the reasoning is, in particular, transparent for users with some background knowledge in probability theory. If the user is unsatisfied with the argumentation graph because it leaves an important question unanswered, we can still apply the linear programming approach in order to see if the knowledge base does allow entailing something interesting about the question (the derived interval is not the full probability interval  $[0, 1]$ ) and to add corresponding arguments to the graph.

Let us also note that when human agents consider a problem in terms of arguments, they are not normally exhaustive in terms of the arguments they consider. Similarly, when human agents enter into a discussion, they are not exhaustive in the arguments they exchange in their dialogue. Rather, they consider or exchange sufficient arguments with which to get an understanding of a topic, and thereby draw any required conclusions. In the context of the healthcare applications we consider in this paper, such as the case studies in Section 4.4, we envisage that the process of constructing an argument graph is an interactive process whereby the user chooses the arguments to put into the argument graph, perhaps with prompts, suggestions, or checks by the automated reasoning machinery.

The second issue concerns the *consistency of the extensions* that are obtained from an argument graph constructed using a proof system PS. When we have an argument graph, whether it is exhaustive or not, we obtain a subset of the arguments as an extension, using Dung's dialectical semantics, or using the epistemic approach. Given this extension, we may want to know whether the union of the claims of the arguments in the extension, or the union of the premises of the arguments in the extension, is consistent. In general, we do not rely on the structure of the argument graph to ensure consistency of the claims or premises of the extension. Rather, we see the primary role of the argument graph is to present arguments and counterarguments to a user so that they can see points of view in the knowledge, rather than to automatically check or resolve inconsistency. However, if the consistency of the claims or support of the arguments in an extension is not guaranteed to be consistent, then we employ consistency checking methods such as the one based on linear optimization explained in Proposition 1. We will expand on these issues in the following subsections.

#### 4.3.1. Analyzing argument graphs with a dialectical approach

Given an argument graph  $G$ , we can apply the definitions for dialectical semantics that we reviewed in Section 2.2. For instance, if we consider Example 28, then we have the empty set as the grounded extension, we have  $\{A1, A3, A4\}$  and  $\{A2, A4\}$  as preferred extensions, and we have  $\{A2, A4\}$  as the stable extension.

In order to investigate the nature of the dialectical semantics for analyzing the argument graphs, we introduce the following properties which capture two ways to describe the consistency of an extension. The first property holds if for each extension, the support of the arguments in the extension constitutes a consistent set, and the second property holds if for each extension, the claims of the arguments in the extension constitute a consistent set.

**Definition 17.** For an argument graph  $G$  and a dialectical semantics  $\sigma$ , such as grounded, preferred, or stable, the **dialectical support coherence (DSC)** and **dialectical claim coherence (DCC)** properties are defined as follows.

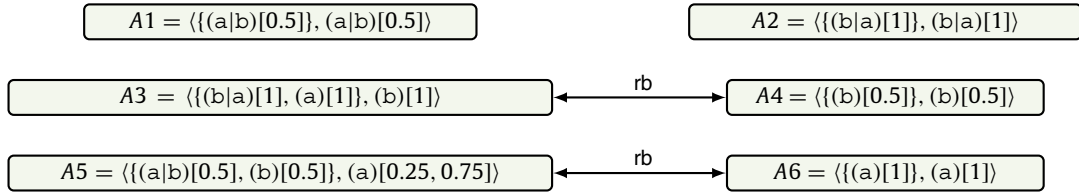
$$\text{for all } \Gamma \in \text{Extension}_\sigma(G), \text{ Sat}(\cup_{A \in \Gamma} \text{Support}(A)) \neq \emptyset \quad (\text{DSC})$$

$$\text{for all } \Gamma \in \text{Extension}_\sigma(G), \text{ Sat}(\cup_{A \in \Gamma} \text{Claim}(A)) \neq \emptyset \quad (\text{DCC})$$

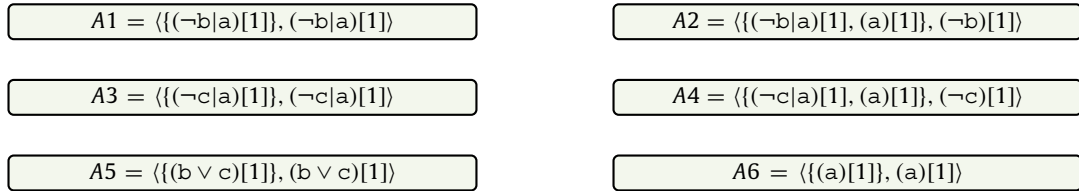
In the following, we consider examples to illustrate satisfaction and failure of the DSC and DCC properties.

**Example 29.** Consider the argument graph in Example 28. For this graph, and for each of grounded, preferred, and stable, semantics, the DSC and DCC properties hold.

**Example 30.** For  $\Delta = \{(a)[1], (b|a)[1], (a|b)[0.5], (b)[0.5]\}$ , the following is the exhaustive argument graph using SIMPLE arguments, and rebuttal. So there is a preferred extension where the union of the premises is  $\Delta$ , which is unsatisfiable, and so the DSC property does not hold, whereas for the ground extension, the DSC property does hold.



**Example 31.** For  $\Delta = \{(a)[1], (b \vee c)[1], (\neg b|a)[1], (\neg c|a)[1]\}$ , the following is the exhaustive argument graph using SIMPLE arguments, and rebuttal, undercut, or direct undercut. The preferred and grounded extension contains all the arguments, but the union of the premises is  $\Delta$ , which is unsatisfiable, and the union of the claims is  $\{(\neg b|a)[1], (\neg c|a)[1], (b \vee c)[1], (a)[1], (\neg b)[1], (\neg c)[1]\}$ , which is also unsatisfiable. So neither the DSC property nor the DCC property holds.



Despite the above examples, there are proof systems, with restrictions on the argumentation, for which we can show that the extensions are consistent. Since there are many proof systems we could consider, we illustrate this claim with the following straightforward result. For this, we restrict the arguments using the following definition: An argument  $A$  is a **literal argument** iff  $\text{Claim}(A)$  is of the form  $(\alpha)[u, v]$  and  $\alpha$  is a literal.

**Proposition 19.** *Let  $PS$  be the SIMPLE proof system, and let  $\rho$  be rebuttal. For a knowledgebase  $\Delta$ , if  $G$  is the subgraph of  $G_\rho^{\text{PS}}(\Delta)$  that contains all the literal arguments and the arcs involving those literal arguments, then the DCC property holds with grounded, preferred, or stable, semantics.*

**Proof.** The DCC property does not hold iff there is an extension  $\Gamma$  s.t.  $\text{Sat}(\cup_{A \in \Gamma} \text{Claim}(A)) = \emptyset$ . Each argument in  $G$  has a claim of the form  $(\alpha)[u, v]$  where  $\alpha$  is a literal. So  $\cup_{A \in \Gamma} \text{Claim}(A)$  is a set of literals. So the DCC property does not hold iff there is a subset  $\Gamma'$  of an extension such that  $\Gamma'$  contains two arguments, and  $\text{Sat}(\cup_{A \in \Gamma'} \text{Claim}(A)) = \emptyset$ . Since the attack relation is rebuttal, for each pair of arguments  $A$  and  $B$  in  $G$ , there is an arc from  $A$  to  $B$  iff  $\{\text{Claim}(A), \text{Claim}(B)\} \vdash \perp$ . So  $A$  and  $B$  cannot both be in the same extension. So there is no subset  $\Gamma'$  of an extension such that  $\Gamma'$  contains two arguments, and  $\text{Sat}(\cup_{A \in \Gamma'} \text{Claim}(A)) = \emptyset$ . So the DCC property holds.  $\square$

Another approach to ensuring consistency of an extension would be to use the approach of argumentation with sets of arguments [23,61]. For this, we can generalize the notion of attack (i.e. Definition 15) so that we allow attack relations to be from a set of arguments to an argument. For instance, the set of arguments  $\{A_1, \dots, A_n\}$  is a **set rebuttal** of an argument  $B$  if  $\text{Sat}(\{\text{Claim}(A_1), \dots, \text{Claim}(A_n), \text{Claim}(B)\}) = \emptyset$  and there is no subset  $\{A'_1, \dots, A'_m\} \subseteq \{A_1, \dots, A_n\}$  such that  $\text{Sat}(\{\text{Claim}(A'_1), \dots, \text{Claim}(A'_m), \text{Claim}(B)\}) = \emptyset$ . Then we can use the definitions for extensions given in [23,61] that are a direct generalization of the definitions given in Section 2.2. Using set rebuttal, the claims in an extension would be consistent, and using set undercut, the supports in an extension would be consistent, for any proof system  $PS$ . We leave the development of this alternative to future work.

To recap, in general, we do not rely on analyzing argument graphs (exhaustive or non-exhaustive) with the dialectical approach to ensure that the union of support of the arguments in an extension, or set of the claims of the arguments in the extension, is consistent (i.e. the DCC and DSC properties). As we can see from the above examples, DCC and DSC are not guaranteed to hold. However, for some specific cases of proof systems and argumentation, we can guarantee that analyzing exhaustive graphs with dialectical semantics does ensure consistency of claims or supports in the extensions, as illustrated by Proposition 19.

#### 4.3.2. Analyzing argument graphs with the epistemic approach

In the epistemic approach to probabilistic argumentation, a probability is assigned to each argument to denote the probability that it is acceptable [46,89]. For an argument  $A$ ,  $\pi(A)$  represents the degree of belief that  $A$  is acceptable. We assume that there is an underlying probability distribution over the set of arguments. In other words, there is a probability assignment to each element in the powerset of arguments such that the sum is 1, and then  $\pi(A)$  is the marginal distribution for  $\{A\}$ . The probability distribution can be assigned so that  $\pi$  is a function of the belief in the composition of the argument itself (for example, as a function of the belief in its premises, belief in its claim, and belief in the derivation of the claim from the premises), and the belief in the acceptability of other relevant arguments (e.g. directly/indirectly supporting or attacking).

How we might formalize this depends on the kinds of arguments we are dealing with and the kinds of application. We give one option for defining  $\pi$  later in this subsection.

The epistemic approach provides a finer grained assessment of an argument graph than given by Dung's definition of extensions. By adopting constraints on the distribution, the epistemic approach subsumes Dung's approach [44,89]. However, there is also a need for a non-standard view where we adopt alternative constraints on the distribution. For instance, we may wish to represent disbelief in arguments even when they are unattacked [68]. This might be because we disbelieve an argument but do not have the evidence for a counterargument.

Constraints on the probability function  $\pi$  can take account of the structure of the graph such as COH, RAT, and OPT below [46,89]. The COH and RAT constraints model the general requirement that, if belief in an argument is high, then the belief in an argument attacked by it should be low. While RAT captures a rather crisp version of the requirement, COH is a continuous version. The OPT constraint states the degree of belief in an argument should be bounded from below by one minus the sum of the beliefs in the attackers. There is a range of possible constraints [44,50,69], and a number of interrelationships identified between them (e.g. COH is a special case of RAT). So ensuring the probability function satisfies one or more of postulates such as these would mean that the attack relation is respected by the probability function.

**COH**  $\pi$  is *coherent* w.r.t.  $G$  if for all  $(A, B) \in \text{Arcs}(G)$  then  $\pi(A) \leq 1 - \pi(B)$ .

**RAT**  $\pi$  is *rational* w.r.t.  $G$  if for all  $(A, B) \in \text{Arcs}(G)$  then  $\pi(A) > 0.5$  implies  $\pi(B) \leq 0.5$ .

**OPT**  $\pi$  is *optimistic* w.r.t.  $G$  if  $\pi(A) \geq 1 - \sum_{B \text{ s.t. } (B,A) \in \text{Arcs}(G)} \pi(B)$  for every  $A \in \text{Nodes}(G)$ .

As explained above, there are various ways that a probability function  $\pi$  can be defined. Here, we will assume that the probability is a function of the belief in the premises being correct. We can therefore view this probability as a meta-level representation of uncertainty, and so it is orthogonal to the object-level uncertainty represented by the conditional statements in the premises. To obtain this, we start with the following definition of a meta-level probability function that assigns a probability value to each subset of the knowledgebase.

**Definition 18.** For a knowledgebase  $\Delta$ , a **meta-level probability function** for  $\Delta$  is  $\pi : \wp(\Delta) \rightarrow [0, 1]$  such that:  $\sum_{\Gamma \subseteq \Delta} \pi(\Gamma) = 1$ .

In the above definition, the mass assigned to the subsets sums to 1. We regard the assignment as the probability that the subset is the correct knowledgebase (i.e. the subset that contains the formulae that are correct). We then use this in the following definition for the probability of acceptability of an argument. It sums the probability assigned to the subsets of the knowledgebase that contain the premises of the argument. So the more mass assigned to subsets that contain the support of the argument, the higher the probability of acceptability of the argument.

**Definition 19.** For an argument  $A$  where  $\text{Support}(A) \subseteq \Delta$ , and a meta-level probability function  $\pi$  for  $\Delta$ , the **probability of acceptability** of  $A$  is

$$\pi(A) = \sum_{\Gamma \subseteq \Delta \text{ s.t. } \text{Support}(A) \subseteq \Gamma} \pi(\Gamma)$$

Given a meta-level probability function  $\pi$ , we can then identify the arguments in the graph that have a probability of acceptability greater than 0.5. This gives us the epistemic extension according to  $\pi$ .

**Definition 20.** For an argument graph  $G$ , and a meta-level probability function  $\pi$  for  $\Delta$ , the **epistemic extension** is

$$\text{Extension}_{\pi}(G) = \{A \in \text{Nodes}(G) \mid \pi(A) > 0.5\}$$

**Example 32.** For  $\Delta = \{(a|b)[0.9], (b)[0.9], (\neg b)[0.7]\}$ , consider the exhaustive argument graph in Example 28. Let the meta-level probability function  $\pi$  be specified as below. For sets  $\Gamma$  that are not listed below,  $\pi(\Gamma) = 0$ .

$$\pi(\{(b)[0.9]\}) = 0.1 \quad \pi(\{(a|b)[0.9], (b)[0.9]\}) = 0.2 \quad \pi(\{(\neg b)[0.7]\}) = 0.7$$

So  $\pi(A_1) = 0.2$ ,  $\pi(A_2) = 0.7$ , and  $\pi(A_3) = 0.3$ . So  $\text{Extension}_{\pi}(G) = \{A_2\}$ .

Next, we consider optional constraints on meta-level probability functions. The first constraint below ensures that no mass is assigned to inconsistent subsets.

**Definition 21.** For a knowledgebase  $\Delta$ , and a meta-level probability function  $\pi$  for  $\Delta$ , the **consistency constraint** is that for all  $\Gamma \subseteq \Delta$ , if  $\text{Sat}(\Gamma) = \emptyset$ , then  $\pi(\Gamma) = 0$ .

Whilst the above consistency constraint may be desirable, we do not want to always impose it since there are situations where we have some belief in an inconsistent set of formulae since we are unable to isolate a consistent subset that we believe. So in the worst case, we would assign a probability of 1 to the whole knowledgebase, and in the best case, we would have isolated a consistent subset of the knowledge to which we would assign a probability of 1. In between these extremes, we would assign mass to multiple sets, which would preferably be consistent, though may be inconsistent. When we construct arguments from a knowledgebase  $\Delta$ , we accept that  $\Delta$  could be inconsistent. Indeed, if it is consistent, then there will be no counterarguments from  $\Delta$ . Without the meta-level probability function, we treat all the formulae as equally likely. So the use of a meta-level probability function is just a refinement of this. Furthermore, we would imagine when an argument graph is constructed from  $\Delta$ , the meta-level probability function could be refined incrementally, perhaps in parallel with adding conditional statements to  $\Delta$  so that an increasingly informed analysis of the information is made, and that eventually the users can indeed isolate a consistent subset of the knowledge with a probability assignment of 1.

**Definition 22.** For a knowledgebase  $\Delta$ , and a meta-level probability function  $\pi$  for  $\Delta$ , the **entailment constraint** is: If  $\text{Sat}(\Gamma) \subseteq \text{Sat}(\Gamma')$ , then  $\pi(\Gamma) \leq \pi(\Gamma')$ , for  $\Gamma, \Gamma' \subseteq \Delta$ .

The above property ensures that, for example, belief in  $(a)[0, 1]$  should be greater than belief in  $(a)[0.4, 0.7]$  or belief in  $(a \vee b)[0.4, 0.7]$  should be greater than belief in  $(a)[0.4, 0.7]$ .

**Example 33.** Consider the meta-level probability function specified as below. This meta-level probability function satisfies the entailment constraint.

$$\begin{aligned}\pi(\{(b \vee c)[0.6, 0.9]\}) &= 0.14 \\ \pi(\{(b)[0.6, 0.9]\}) &= 0.09 \\ \pi(\{(b)[0.8, 0.9]\}) &= 0.07 \\ \pi(\{(\text{d}|e)[0.6, 0.9], (e)[0.5, 0.9]\}) &= 0.49 \\ \pi(\{(\text{d}|e)[0.6, 0.9], (\neg e)[0.6, 0.7]\}) &= 0.21\end{aligned}$$

**Proposition 20.** For a knowledgebase  $\Delta$ , and a meta-level probability function  $\pi$  for  $\Delta$ , the entailment constraint implies the following property: If  $\{\phi\} \models \psi$ , then  $\pi(\phi) \leq \pi(\psi)$ , for  $\phi, \psi \in \Delta$

**Proof.** Assume  $\{\phi\} \models \psi$ . So by definition of the entailment relation,  $\text{Sat}(\{\phi\}) \subseteq \text{Sat}(\{\psi\})$ . So the entailment constraint implies  $\pi(\phi) \leq \pi(\psi)$ .  $\square$

We now present two properties concerning the union of the support (respectively union of the claims) of arguments in the epistemic extension of an argument graph.

**Definition 23.** For an argument graph  $G$  and a meta-level probability function  $\pi$ , the **epistemic support coherence** (ESC) and **epistemic claim coherence** (ECC) properties are defined as follows.

$$\text{Sat}(\cup_{A \in \text{Extension}_{\pi}(G)} \text{Support}(A)) \neq \emptyset \quad (\text{ESC})$$

$$\text{Sat}(\cup_{A \in \text{Extension}_{\pi}(G)} \text{Claim}(A)) \neq \emptyset \quad (\text{ECC})$$

To illustrate these properties, we give two examples where the properties hold, and then an example where the property fails.

**Example 34.** From Example 28, consider  $\Delta = \{(a|b)[0.9], (b)[0.9], (\neg b)[0.7]\}$ . Let  $\pi$  be a meta-level probability function where  $\pi(\{(a|b)[0.9], (b)[0.9]\}) = 0.9$  and  $\pi(\{(\neg b)[0.7]\}) = 0.1$ . For the exhaustive argument graph  $G$  in Example 28,  $\text{Extension}_{\pi}(G) = \{A_1, A_3, A_4\}$ . So the ESC and ECC properties hold for this graph.

**Example 35.** From Example 30, consider  $\Delta = \{(a)[1], (b|a)[1], (a|b)[0.5], (b)[0.5]\}$ . Let  $\pi$  be a meta-level probability function where  $\pi(\{(a)[1], (b|a)[1]\}) = 0.6$  and  $\pi(\{(a|b)[0.5], (b)[0.5]\}) = 0.4$ . For the exhaustive argument graph  $G$  in Example 30,  $\text{Extension}_{\pi}(G) = \{A_2, A_3, A_6\}$ . So the ESC and ECC properties hold for this graph.

**Example 36.** Continuing Example 35, we consider an alternative meta-level probability function  $\pi'$  where  $\pi'(\{(a)[1], (b|a)[1], (a|b)[0.5], (b)[0.5]\}) = 1$ . For the exhaustive argument graph  $G$  in Example 30,  $\text{Extension}_{\pi}(G) = \{A_1, A_2, A_3, A_4, A_5, A_6\}$ . Because we have used a meta-level probability function that violates the consistency constraint (Definition 21), the ESC and ECC properties do not hold for this argument graph. Furthermore,  $\pi$  violates the COH and RAT constraints, and therefore it ignores the structure of the argument graph. So the choice of this meta-level probability function is questionable for this argument graph.

For the following results, we use the following definition where  $A$  is an argument and  $\pi$  is a meta-level probability function for  $\Delta$ :  $\text{Supporters}_\pi(A) = \{\Gamma \subseteq \Delta \mid \text{Support}(A) \subseteq \Gamma \text{ and } \pi(\Gamma) > 0\}$ .

**Proposition 21.** *For an argument graph  $G$  and a meta-level probability function  $\pi$ , if  $\pi$  satisfies the consistency constraint (Definition 21), then the epistemic support coherence (ESC) and epistemic claim coherence (ECC) properties hold.*

**Proof.** To show epistemic support coherence, we assume  $\text{Sat}(\cup_{A \in \text{Extension}_\pi(G)} \text{Support}(A)) = \emptyset$ , and derive a contradiction as follows. From  $\text{Sat}(\cup_{A \in \text{Extension}_\pi(G)} \text{Support}(A)) = \emptyset$ , there is one or more subsets  $\Theta$  of  $\cup_{A \in \text{Extension}_\pi(G)} \text{Support}(A)$  that is a minimal inconsistent subset (i.e.  $\text{Sat}(\Theta) = \emptyset$ , and for all  $\Theta' \subset \Theta$ ,  $\text{Sat}(\Theta') \neq \emptyset$ ). So for each minimal inconsistent subset  $\Theta$ , there are arguments  $A_i, A_j \in \text{Extension}_\pi(G)$ , and there are  $\phi_i, \phi_j \in \Theta$  such that  $\phi_i \in A_i$  and  $\phi_i \notin A_j$  and  $\phi_j \notin A_i$  and  $\phi_j \in A_j$ . Recall the consistency constraint holds. So for all  $\Gamma \subseteq \Delta$ , where  $\pi(\Gamma) > 0$ , if  $\text{Support}(A_i) \subseteq \Gamma$ , then  $\text{Support}(A_j) \not\subseteq \Gamma$ , and if  $\text{Support}(A_j) \subseteq \Gamma$ , then  $\text{Support}(A_i) \not\subseteq \Gamma$ . So  $\text{Supporters}_\pi(A_i) \cap \text{Supporters}_\pi(A_j) = \emptyset$ . So  $\sum_{\Gamma \in \text{Supporters}_\pi(A_i)} \pi(\Gamma) + \sum_{\Gamma \in \text{Supporters}_\pi(A_j)} \pi(\Gamma) \leq 1$ . So  $\pi(A_i) \leq 0.5$  or  $\pi(A_j) \leq 0.5$ . So this contradicts  $A_i$  and  $A_j$  being in  $\text{Extension}_\pi(G)$ . From this contradiction, we infer  $\text{Sat}(\cup_{A \in \text{Extension}_\pi(G)} \text{Support}(A)) \neq \emptyset$ . To show epistemic claim coherence, we assume epistemic support coherence holds. Since each argument  $A$  is such that  $\text{Support}(A) \vdash_{PS} \text{Claim}(A)$ , it is the case that  $\text{Sat}(\cup_{A \in \text{Extension}_\pi(G)} \text{Claim}(A)) \neq \emptyset$ .  $\square$

The following result shows that COH and RAT hold when  $\pi$  satisfies the consistency constraint. We would aim for the consistency constraint to hold, and so if it holds, we get the desirable properties of COH and RAT which ensure that the epistemic extension respects the structure of the graph (i.e. the extension does not contain both an attacker and attackee).

**Proposition 22.** *For an argument graph  $G$  and a meta-level probability function  $\pi$ , if  $\pi$  satisfies the consistency constraint (Definition 21), then  $\pi$  is coherent and rational.*

**Proof.** Assume that  $(A, B) \in \text{Arcs}(G)$ . So  $A$  attacks  $B$ , and so  $\text{Sat}(\text{Support}(A) \cup \text{Support}(B)) = \emptyset$ . But since  $A$  and  $B$  are arguments,  $\text{Sat}(\text{Support}(A)) \neq \emptyset$  and  $\text{Sat}(\text{Support}(B)) \neq \emptyset$ . So there are  $\phi, \psi$  such that  $\phi \in A$  and  $\phi \notin B$  and  $\psi \notin A$  and  $\psi \in B$ . Recall the consistency constraint holds. So for all  $\Gamma \subseteq \Delta$ , where  $\pi(\Gamma) > 0$ , if  $\text{Support}(A) \subseteq \Gamma$ , then  $\text{Support}(B) \not\subseteq \Gamma$ , and if  $\text{Support}(A) \subseteq \Gamma$ , then  $\text{Support}(B) \not\subseteq \Gamma$ . So  $\text{Supporters}(A) \cap \text{Supporters}(B) = \emptyset$ . So  $\sum_{\Gamma \in \text{Supporters}(A)} \pi(\Gamma) + \sum_{\Gamma \in \text{Supporters}(B)} \pi(\Gamma) \leq 1$ . So  $\pi(A) \leq 1 - \pi(B)$ . So COH holds. Since, COH implies RAT, RAT also holds.  $\square$

Example 32 shows that OPT does not hold in general (for instance,  $\pi(A_1) = 0.2$ ,  $\pi(A_2) = 0.7$ , and  $(A_2, A_1)$  holds, so  $\pi(A_1) \not\geq 1 - \pi(A_2)$ , and hence OPT fails). Whilst OPT is an interesting property for investigating argumentation systems, and in some applications it is a desirable property, we do not believe that it essential for our purposes in this paper. Indeed, in the epistemic approach, as discussed in [44], we want the flexibility to disbelieve an argument, even if there is not an attacker of it that is believed. For instance, if we are analyzing medical evidence, we may have an unattacked argument that is based on evidence that we do not believe, but we do not have evidence to construct a counterargument, and so we may choose to disbelieve the argument even though it is unattacked.

For applications of analyzing conflicting evidence, the epistemic approach allows for the representation of meta-level uncertainty, and this can then be used to identify the arguments that have a probability of acceptability above a certain threshold such as 0.5. So the individual arguments, counterarguments, and supporting arguments, capture the object-level uncertainty in the evidence (using the conditionals), and then the epistemic approach captures the meta-level uncertainty.

In order to ensure that the ESC and ECC properties hold, we can seek a meta-level probability function that satisfies the consistency constraint. Since we are dealing with inconsistent knowledge, we are likely to have an incomplete understanding of which formulae are correct (otherwise, we would be able to select a consistent subset with total confidence and eliminate the rest). From the incomplete understanding of the formulae, we might be able to identify some of the formulae that we either have strong belief or strong disbelief, and then select the maximum entropy distribution that is consistent with these beliefs.

#### 4.4. Case studies

For our case studies, we consider the adequacy of our proposals for managing evidence in healthcare via the following examples.

**Example 37.** Consider a scenario where we have three studies concerning patients with a specific disease and treatment. Two of the studies have similar results about the efficacy being 0.91 and 0.87 respectively whereas one study has a much lower value of 0.24. We use the following atoms.

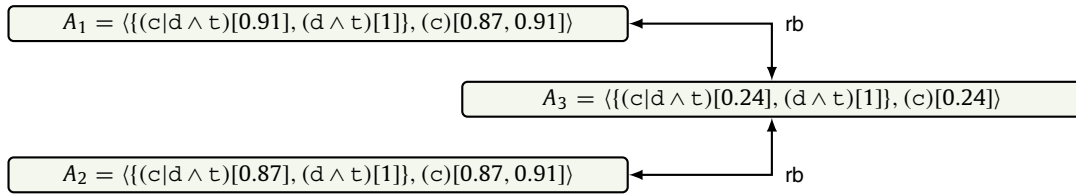
---

c	the patient is cured of the disease
d	the patient has the disease
t	the patient has the drug treatment

---



The argument graph below is obtained using a variant of the STAR consequence relation that also includes the IR inference rule. The premises of  $A_1$  imply  $(c)[0.91]$  and the premises of  $A_2$  imply  $(c)[0.87]$ . From these, we infer  $(c)[0.87, 0.91]$  using interval relaxation. So the claims of  $A_1$  and  $A_2$  both rebut the claim of  $A_3$ .



Also suppose the first two studies were reliable large studies whereas the third was a smaller less reliable study. Given that the first two studies were large reliable studies, the belief in the premises could be much higher than for the third study. For this, we use the following meta-level probability function  $\pi$ . Hence, the belief in the first two arguments is higher than the third.

$$\pi(\Gamma(\{(c|d \wedge t)[0.91], (c|d \wedge t)[0.87], (d \wedge t)[1]\})) = 0.9$$

$$\pi(\{(c|d \wedge t)[0.24], (d \wedge t)[1]\}) = 0.1$$

This example shows how clinical evidence can naturally be represented as conditionals, and these used in arguments to reason about efficacy of treatments. The premises and claim of each argument are explicitly presented, and the derivation of the claim from the premises is straightforward to identify. Furthermore, similar results from clinical studies can give multiple arguments with the same claim, and results that are substantially different can give rebutting arguments.

Note, in the above example, the meta-level probability function assigns non-zero mass to the first set and that set is inconsistent. Recall that the meta-level probability function assigns a value to indicate the probability that the subset of the knowledgebase is the correct subset. Whilst in general, it is desirable that mass is not assigned to an inconsistent set, we can see that there is only small difference in the two conditionals, and so the inconsistency could be viewed as minor. In future work, we will investigate using inconsistency measures to analyze inconsistent subsets, and set a threshold for tolerable inconsistency.

In our next example, we consider how we can use the meta-level probability function to capture a preference for arguments based on more specialized information.

**Example 38.** Consider a scenario where we have three studies concerning observable hypertension in people over 80 years of age. The first study is the prevalence in the population in this age group, the second study is prevalence in those that have a treatment for hypertension, and the third study is the prevalence in those that are treated with angiotensin-2 receptor blocker which is a common specific treatment for hypertension. We use the following atoms.

---

o	a random individual who is over 80 years
h	the blood pressure is above the normal range (hypertension is observable)
t	treated with a medicine for hypertension
a	treated with angiotensin-2 receptor blocker

---

We represent this scenario in Fig. 3. If we want to express a preference for the more specific information, we can use the meta-level probability distribution below. Hence, the belief in the third argument is highest, and it is the only argument in the epistemic extension.

$$\pi(\{(h|o)[0.72], (o)[1]\}) = 0.1$$

$$\pi(\{(h|o \wedge t)[0.25], (o \wedge t)[1]\}) = 0.2$$

$$\pi(\{(h|o \wedge a)[0.12], (o \wedge a)[1]\}) = 0.7$$

This example shows how different observational studies (studies that track classes of patient) give rebutting arguments. Each observational study can be naturally represented as a conditional. Furthermore, we see different studies are based on different criteria. By using a meta-level distribution, we can obtain the argument based on the most specific evidence.

In the following examples, we introduce the notion of supporting argument. This allows for larger arguments to be split into multiple subsidiary arguments. Furthermore, it allows for transparency in the probabilistic reasoning as we can break up the reasoning so that there is only one or two inference steps per argument.

**Definition 24.** For arguments  $A$  and  $B$ , if  $\text{Sat}(\text{Support}(A) \cup \text{Support}(B)) \neq \emptyset$ , and  $\text{Claim}(A) \in \text{Support}(B)$  then  $A$  is a **direct support** of  $B$ .

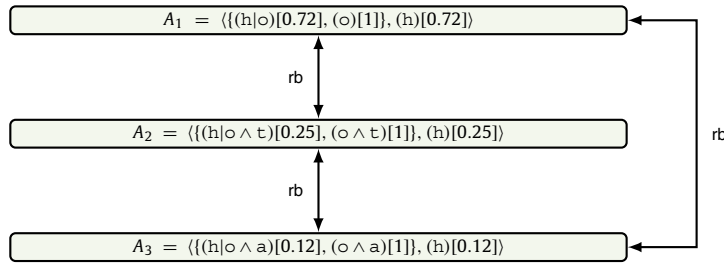


Fig. 3. Argument graph for Example 38.

When an argument  $A$  supports and an argument  $B$ , we call  $A$  a supporting argument for  $B$ . When we present argument graphs, we will use a dashed line as illustrated in the following example.

**Example 39.** Consider a scenario that involves identifying the diagnosis and appropriate treatment of appendicitis. We use the following atoms.

a	the patient has appendicitis
p	the patient has pain from appendix
e	the patient has eaten a large meal in the past 4 hours
pa	the patient has pain in the abdomen
i	the patient has indigestion
ia	the patient has an inflamed appendix
es	the patient has swelling around appendix in a scan
ea	the patient has enlarged appendix
ks	the patient has keyhole surgery
r	the patient has recovered in 2 weeks

We represent this scenario in Fig. 4. We explain the conditionals in the arguments as follows:  $A_1$  concerns the probability that a patient does not have pain from appendix when they have had a large meal and they are likely to be suffering from indigestion;  $A_2$  concerns the probability that a patient has an inflamed appendix when they have pain from the appendix;  $A_3$  concerns the probability that a patient has an enlarged appendix when they have swelling around appendix in a scan;  $A_4$  concerns the probability of the conjunction of probability of inflamed appendix and enlarged appendix;  $A_5$  concerns the probability that a patient has appendicitis given they have an inflamed appendix and enlarged appendix; And  $A_6$  concerns the probability that a patient has recovered within 2 weeks given that they have appendicitis, and they have had keyhole surgery. We assume a meta-level probability function over the formulae such that the belief in  $A_1$  is below 0.5, and the belief in the other arguments is above 0.5.

This example shows how we might use conditionals that come from observational studies, or from analysis of patient records, in order to capture supporting and attacking arguments relevant to a prognosis.

In the next example, we see an issue that arises when using conditional probabilities for considering statistical information about populations and epistemic information about individuals. Since we want to use argumentation as a means for bringing together both statistical and epistemic usage of probabilities, we need ways to deal with this multi-modal usage. [38] proposed a first-order probabilistic logic that allows reasoning about both types of probabilities, but many reasoning problems are undecidable in this language. Therefore, we use an extra-logic solution here.

Essentially, in our extra-logical solution, we will adopt a policy where we use different propositional letters to distinguish between using a probability to represent statistical information and using a probability to represent epistemic information. For instance, in the following example, we use  $pp$  to denote that a random individual tests positive, and  $p$  to denote that the patient we are concerned with tests positive. Similarly, we use  $pd$  to denote that a random individual has the disease, and  $d$  to denote that the patient we are concerned with has the disease. Note, there is nothing in particular in this policy about the population level proposition to start with  $pp$ . We can use any name for any proposition.

Next, we make a connection between the symbols so that we can for instance reason about conditionals concerning the population in general, and reason about conditionals concerning a specified individual. So for instance in the following example,  $A_1$  involves statistical reasoning to infer the claim  $(pd|pp)[0.495]$  which says that the probability that a randomly selected member of the population has the disease given that they have tested positive is 0.495, and  $A_2$  which involves empirical reasoning with the premise  $(d|p)[0.495]$  which says that the probability that the specific patient has the disease given that they have tested positive is 0.495. Since,  $A_1$  is a supporting argument to  $A_2$ , we make the connection explicit by adding a label to the arc between  $A_1$  and  $A_2$  in the graph specifying which atoms are connected going from the population

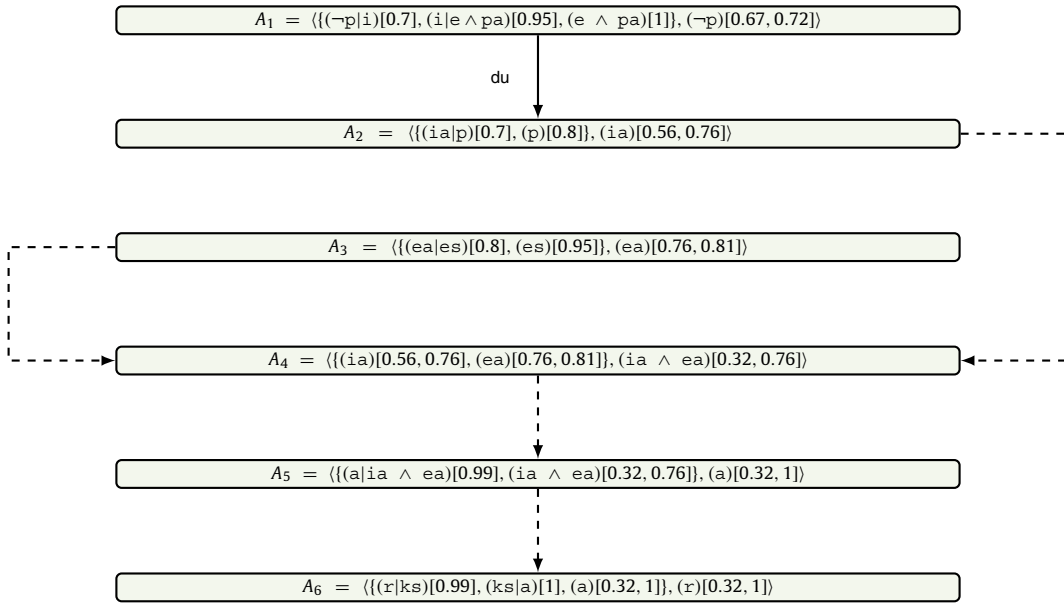


Fig. 4. Argument graph for Example 39.

level to the individual level. So for this example, we add the label composed of  $pd \mapsto d$  and  $pp \mapsto p$ . By using this explicit mapping, we clearly specify that the claim of  $A_1$  can be used as a premise in  $A_2$ .

In order to capture this connection, we expand the definition of support (as follows) so that it makes explicit how the symbols in the claim of the supporting argument are related to the symbols in the supported premise.

**Definition 25.** Let  $\Pi = \{\alpha_1 \mapsto \beta_1, \dots, \alpha_n \mapsto \beta_n\}$  be a set of mappings from a propositional atom to another propositional atom, and for a formula  $\phi$ , let  $\phi \circ \Pi$  denote the result of replacing each occurrence of  $\alpha_i$  in the formula  $\phi$  by  $\beta_i$  for each  $\alpha_i \mapsto \beta_i \in \Pi$ , and for a set of formulae  $\Phi$ , let  $\Phi \circ \Pi$  denote  $\{\phi \circ \Pi \mid \phi \in \Phi\}$ . For arguments  $A$  and  $B$ , if  $\text{Sat}((\text{Support}(A) \circ \Pi) \cup \text{Support}(B)) \neq \emptyset$ , and  $(\text{Claim}(A) \circ \Pi) \in \text{Support}(B)$  then  $A$  is an **instantiated support** of  $B$ .

This policy does not affect the individual arguments as they are a set of premises that imply the claim using the specific inference rules. However, the policy is in effect generalizing the definition of a support relationship. Essentially, for an argument  $A$  to support an argument  $B$ , instead of stipulating that the claim of  $A$  is a premise of  $B$ , we now stipulate that the formulae obtained by applying the mapping to the claim of  $A$  is a premise of  $B$ .

**Example 40.** Consider a diagnosis where we have a specific test and from this diagnosis and a treatment, we want to infer the probability that they will be cured. The sensitivity of the test is 99% and the probability of a positive test result is 0.2. Suppose that the prevalence of the disease is such that the probability that someone has the disease is 0.1. Also suppose that the probability someone is cured from the disease given the diagnosis and the treatment is 0.83.

The argument graph in Fig. 5 is obtained using the BRAN consequence relation, and it uses the following atoms.

pp	a random individual in the population tests positive
pd	a random individual in the population has the disease
p	the patient we are concerned with tests positive
d	the patient we are concerned with has the disease
t	the patient we are concerned with has the treatment
c	the patient we are concerned with is cured

This example shows how we can harness Bayesian reasoning in arguments about tests and diagnosis. Furthermore, it illustrates a way of connecting reasoning about specific individuals and reasoning about classes of individuals which can be important for argumentation with evidence in healthcare.

Whilst there are substantial proposals for first-order probabilistic logics that combine reasoning about classes of individuals and reasoning about specific individuals (see for example [7,38]), they involve more complex formalisms that go beyond what we require. They provide logics that combine statistical statements, and default statements, to give rich and

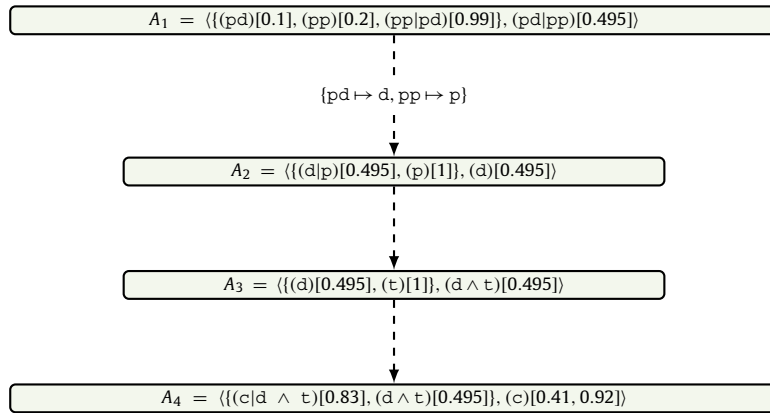


Fig. 5. Argument graph for Example 40.

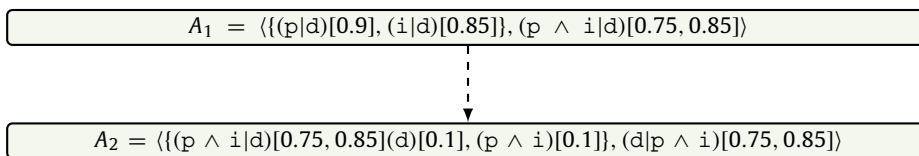
expressive formalisms. Furthermore, the reasoning problems in Halpern’s logic are undecidable. We have taken a different approach that is a policy to relate symbols as used in different arguments. This simpler approach is decidable and sufficient for our purposes. At the level of the argument graph, our approach allows for a transparent method for relating specific individuals to randomly selected individuals.

The use of Bayes rule is common in analyzing healthcare evidence. In the previous example, we considered how the BRAN consequence relation, which incorporates the BR proof rule, can be used for argumentation about diagnosis. In the next example, we consider using it in arguments about prognosis.

**Example 41.** Consider diagnosis scenario where we have a specific test and a specific prognostic indicator such as a pre-existing condition. Suppose the probability of a positive result given the person has the disease is 0.9, and the probability of the prognostic indicator given the person the disease is 0.7. Suppose that the prevalence of the disease is such that the probability that someone has the disease is 0.1. Also suppose the prevalence of having a positive test result and the prognostic indicator is 0.1. We use the following atoms.

$d$	a randomly selected patient has the disease
$p$	a randomly selected patient has a positive test result
$i$	a randomly selected patient has the prognostic indicator

The following is the argument graph obtained using the BRAN consequence relation.

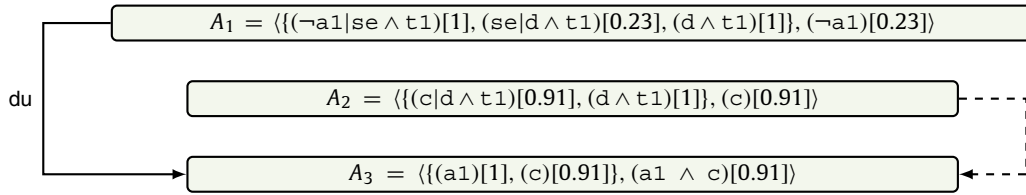


This example illustrates how we can use Bayesian reasoning in reasoning with both a clinical test and a prognostic indicator in order to construct an argument for a prognosis.

**Example 42.** Consider a scenario where we have two studies concerning patients with a specific disease and two treatments with the same specific side-effect. The first study is for  $T_1$  that shows a very high efficacy (0.91) but also a reasonably high probability of the side-effect (0.23), and the second study is for  $T_2$  that show a high efficacy (0.79) but a very low probability of the side-effect (0.0001). We use the following atoms.

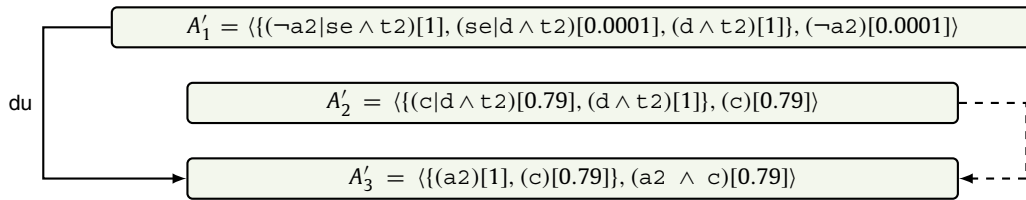
$c$	the patient is cured
$d$	the patient has the disease
$t_1$	the patient has drug treatment $T_1$
$t_2$	the patient has drug treatment $T_2$
$se$	the patient has a serious side-effect
$a_1$	treatment $T_1$ is assumed to be safe
$a_2$	treatment $T_2$ is assumed to be safe

The following is the argument graph for treatment  $T_1$  obtained using the STAR consequence relation and direct undercut.



Given the claim of  $A_1$  has a quite high probability that the treatment is not assumed to be safe (i.e.  $\neg a_1[0.23]$ ), then we may wish to regard the argument as having a high probability of acceptability. So we could assume a meta-level probability function over the formulae such that the belief in  $A_1$  is above 0.5, and the belief in  $A_2$  and  $A_3$  is below 0.5.

The following is the argument graph for treatment  $T_2$  obtained using the STAR consequence relation and direct undercut.



Given the claim of  $A'_1$  has a very low probability that the treatment is not assumed to be safe (i.e.  $\neg a_s[0.0001]$ ), then we may wish to regard the argument as having a low probability of acceptability. So we could assume a meta-level probability function over the formulae such that the belief in  $A'_1$  is below 0.5, and the belief in the other arguments is above 0.5.

So we have made explicit the supporting and counterarguments for each treatment. Then using the epistemic approach to probabilistic argumentation, we determine that the argument  $A_3$  is not acceptable (and so treatment  $T_1$  cannot be assumed to be safe as a cure) whereas the argument  $A'_3$  is acceptable (and so treatment  $T_2$  can be assumed to be safe as a cure).

This example illustrates how we can present arguments (supporting and attacking) concerning the benefits and problems for specific treatments. This makes explicit the evidence and how it is used.

Part of the aim of argumentation is to highlight important inconsistencies, and hence we might want to suppress minor inconsistencies. There are various ways we may wish to suppress minor inconsistencies, depending in the information we are dealing with. For instance, we could use IR which would give a single claim with a probability interval. However, we may wish to keep two conflicting claims, and thereby see the two points of view on the claim but ignore the attack between them. For instance, in the above example, we see that the difference in the probabilities of the inferences from  $A_1$  and  $A_2$  is small. In such a case, we may wish to ignore the inconsistencies, and therefore ignore the attack relationship. We will explore how we decide to suppress inconsistencies, and with what mechanisms, in future work.

As we explained in Section 4.3, we do not use dialectical semantics, or the epistemic approach to argumentation, to ensure that the union of the support of the arguments, or the union of the claims of the arguments, in an extension to be consistent. In other words, we do not assume that the DSC, DCC, ESC, or ECC properties hold in general, though, as we have shown, the DSC and DCC properties hold for some proof systems and forms of argumentation, and the ESC and ECC property hold if the meta-level probability function satisfies the consistency constraint. So when we want to ensure these properties hold, we use consistency checking. For the kinds of healthcare case studies we have considered in this section, when the ESC or ECC properties do not hold for an argument graph, it means that we need to investigate the knowledgebase (or seek more information, do further data analytics on patient records, or even do more experimental/clinical studies) in order to revise the meta-level probability function and/or knowledgebase. As an interim step, constructing an argument graph for which the ESC or ECC properties fail may be useful to help revise the meta-level probability function and perhaps to add to the knowledgebase.

## 5. Comparison with the literature

The proposal in this paper is related to work in abstract argumentation, structured argumentation, and probabilistic logics as we discuss in the following subsections.

### 5.1. Abstract argumentation

Two key approaches to probabilistic abstract argumentation are the constellations and the epistemic approaches [46]: In the constellations approach, there is uncertainty over the structure of the graph (e.g. [24,45,53]) and in the epistemic

approach there is uncertainty about whether an argument is believed to be acceptable [10,33,44,46,89]. A further approach is based on labellings for arguments using *in*, *out*, and *undecided*, from [15], augmented with *off* for arguments not occurring in the graph [82]. A probability distribution over labellings gives a form of probabilistic argumentation that overlaps with the constellations and epistemic approaches. These approaches provide some clear and intuitive options for conceptualizing uncertainty in abstract argumentation. However, the arguments are abstract, and therefore the structure of them is not considered.

An alternative to probabilistic abstract argumentation is graded and ranking-based semantics proposed for a number of argumentation frameworks [1,2,8,12,13,17,25,71,72,77,78,86]. An interesting aspect of these approaches is the proposal of a range of postulates for how the strength of an argument is influenced by the strength of other arguments. Many assign a value in the unit interval to arguments, though without a probabilistic interpretation. The postulates offer natural and intuitive behaviors for non-probabilistic interpretation, but often are not consistent with the laws of probability theory. Furthermore, since the arguments are abstract, there is no internal structure.

## 5.2. Probabilistic argumentation

At the structured level, Haenni [36] proposed an approach to probabilistic argumentation in which pros and cons are generated from a classical logic knowledgebase, and then a probability distribution over models of the language is used to assign a belief in each argument. Subsequently, this was generalized by Hunter to arbitrary argument graphs [46] in which various kinds of counterargument can be accommodated. Probabilistic extensions of ABA and ASPIC+ have been discussed in [16,24] and [79]. In other logic-based proposals, Verheij has combined probabilities with non-monotonic inference [93], and separately, he has combined qualitative reasoning in terms of reasons and defeaters (adapting Pollock's definitions [73]), with quantitative reasoning using argument strength, modeled as the conditional probability of the conclusions given the premises [94]. In another proposal, that extends the use of a probability distribution over possible worlds, Verheij [95] has presented a framework for argumentation that allows for the strength of logical arguments to be determined in terms of conditional probabilities.

The epistemic approach to probabilistic argumentation has been harnessed in a framework that allows for the generation of all possible arguments based in the language, then arguments are selected for presentation in an argument based on the probability distribution and the requirements for the graph (i.e. rather than exhaustively present all arguments, present those that are useful to a purpose according to some specified criteria) [47]. In all these proposals at the structured level (discussed in this section), the language is a form of propositional language, rather than a probabilistic language, the inference relation is for rules or propositional formulae, rather than inference for the probabilistic formulae. These proposals provide means for quantifying uncertainty in logical reasoning in structured arguments, but they do not provide the ability for representing and reasoning directly with conditional probability statements in structured arguments.

Other logics have been used in forms of probabilistic argumentation. In a rule-based system for dialogical argumentation, the belief in the premises of an argument is used to calculate the belief in the argument, though the nature of this belief is not investigated [85]. More recently, a probabilistic formulae of the form  $P(\alpha)\#v$  where  $\alpha$  is a propositional formula,  $\# \in \{<, \leq, =, \geq, >\}$ , and  $v \in [0, 1]$  has been used in a form of deductive argumentation [48]. This proposal also considered how sets of formulae could be relaxed to allow for the identification of a probability distribution for the epistemic approach to probabilistic argumentation to be used. However, this proposal does not provide the ability for representing and reasoning directly with conditional probability statements in structured arguments.

Conditional probabilities have been considered as an approach to quantifying argument strength but this work did not consider reasoning with conditions nor how they can be used in instantiating argument graphs [67]. More recently, conditional probabilities have been captured in a form of defeasible rule in defeasible logic programming which limits the reasoning to a form of modus ponens [21]. This enables uncertainty in logic programming to be quantified. However, it does not provide the richer reasoning with conditional probability statements that is required for instance for reasoning with evidence as discussed in this paper.

Conditional probabilities have also been used to interpret an attack by one argument on another [66]. The premise  $\psi$  attacks the claim  $\phi$  when  $P(\neg\phi \mid \psi) \geq t$  for some threshold  $t$ . Various properties of this interpretation are then investigated theoretically and with participants. However, this proposal does not consider how to construct arguments based on conditional probability statements.

## 5.3. Probabilistic logics

Our probabilistic argumentation approach is founded on probabilistic conditional logic [37,60]. The original propositional setting has been extended to more expressive logical languages [30,51,55] and to higher-order probabilities and nested conditionals [19,88]. State-of-the art reasoners are usually based on a linear programming formulation of the probabilistic entailment problem that uses column generation techniques [14,27,32]. As we mentioned before, a proof system for probabilistic conditional logic has been presented in [29]. The linear programming approach receives more attention because it is easier to implement. However, it has also been suggested to combine proof systems with linear programming approaches [43]. As we mentioned before, [29] presented 32 sound inference rules for probabilistic logic. In contrast to our work, the

rules have been designed with short proofs in mind, whereas our focus is on rules that are close to logical and basic probabilistic reasoning rules with focus on comprehensibility rather than generality and conciseness. Also, the soundness proofs of the rules in [29] have been left to the reader, whereas we give all soundness proofs for our rules in this paper. We could add some of the rules in [29] to our proof system to make it stronger. However, some of the rules are very specific and perhaps more interesting for a general logical setting than an argumentation setting where we are interested in modeling the reasoning of human experts like medical doctors or lawyers. For example, rule (i) in [29] has the form

$$\frac{(\alpha \mid \delta)[x, y] \quad (\alpha \vee \beta \mid \delta)[u, v] \quad (\alpha \wedge \beta \mid \delta)[w, z]}{(\beta \mid \delta)[s, t]}$$

where  $s = \max\{w, u - y + w\}$  and  $t = \max\{v, v - x + z\}$ . However, some of our rules also occur in [29]. Our rule *CN* corresponds to rule (iv), our axioms to rule (xvi), *IR* to rule (xix), *CI* to (xxv), *CE* to (xxvi), *DI* to (xxvii) and *DE* to (xxviii). As the rules in [29] are quite general, it may be possible to derive our remaining rules from them. However, we did not check this as our focus is not on finding a compact proof system, but one that is well suited for transparent probabilistic argumentation. For example, Bayes' rule is probably well known to most people applying probabilistic models so that a proof that uses *BR* is more comprehensible than one that derives the conclusion of *BR* using more primitive rules. What seems more interesting for our purpose are the rules from [29] that include classical logical assumptions, which may also be interesting for the argumentation setting and for reasoning about causality. However, we leave the discussion of these issues for future work.

Another (incomplete) set of inference rules has been presented in [70]. The motivation of this work is closer to ours in that it tries to capture human inference by these rules. The paper presents 14 inference rules. As opposed to our rules, these rules allow only point probabilities (rather than probability intervals) in the premises. Our rule *CI* corresponds to *And* in [70]. There is also an *Or*-rule in [70], but it allows connecting the conditions disjunctively rather than the consequents of conditionals. The authors present a *Probabilistic Modus Ponens* and *Probabilistic Modus Tollens*. While the former correspond to the special case of our *PMP* for point probabilities, the latter has the form

$$\frac{(\beta \mid \alpha)[p_2] \quad (\beta)[p_3]}{(\neg\alpha)[p_1, 1]}$$

where  $l_1 = \max\{\frac{1-p_2-p_3}{1-p_1}, \frac{p_1+p_2-1}{p_1}\}$ ,  $u_1 = 1$  for the *Probabilistic Modus Tollens*. Note that the reason that they derived a lower, while we derived an upper bound for *PMT* is simply that *PMT* speaks about the probability of  $\alpha$ , not about the probability of  $\neg\alpha$ . The remaining rules in [70] are inspired by classical logic rules as well and include probabilistic variants of Left Logical Equivalence, Right Weakening and Cautious Monotonicity. Generalizing these rules to interval probabilities and including them in our proof system is another interesting task for future work.

There has been some other work on restricted proof systems. Another *Generalized Modus Ponens* has been presented in [88]. As opposed to our *GPMP*, it generalizes *PMP* by allowing not only conditioning on a formula, but conditioning on another conditional. Adam's probability logic [22] allows to derive some statements about the probability of formulae based on classical rules that they are involved in and the probabilities of their premises. That is, given a classical rule,  $B_1, \dots, B_n \rightarrow H$ , one can derive some bounds on  $P(H)$  based on the probabilities of  $P(B_1), \dots, P(B_n)$ .

Proof systems have also been considered for probabilistic inference processes [81]. As opposed to our setting, where we reason over all probability distributions that satisfy our conditionals, an inference process picks a particular probability distribution. One popular example is choosing the distribution that maximizes entropy.

To the best of our knowledge, larger proof systems for probabilistic logic have not yet been applied to argumentation problems. However, let us note that a structured probabilistic argumentation approach that uses one variant of the *Probabilistic Modus Ponens* has recently been applied to planning under uncertainty [56].

## 6. Discussion

In this paper, we have proposed a set of sound inference rules for probabilistic logic that can be used within a framework for probabilistic argumentation. We can choose the set of inference rules we use according to the application. Each probabilistic argument provides a set of premises and a claim following from those premises. We have defined three types of attack relation (rebut, undercut, and direct undercut), and a type of support relation.

These proposals allow us to meet the requirements for argumentation with conditional probabilities that we presented in the introduction. We are able to provide transparent reasoning with conditional probabilities where each argument is based on one or two inference steps, and then sequences of inference step are captured by the support relation. This allows us to see how the more complex arguments are composed from simpler arguments, and to see which inference rules are used. We are also able to represent meta-level uncertainty over conditional probabilities using the epistemic approach to probabilistic argumentation, and hence we can represent and reason with both object-level and meta-level uncertainty in the computational argumentation.

Not all inconsistencies between conditionals are the same, which raises the question of how close do probabilities need to be for inconsistency to be void? For example,  $(a)[0.8]$  and  $(a)[0.81]$  are close and seem almost "consistent" together. Also if one proposition has a wide interval then it denotes lower confidence. For example, in the formula  $(a)[0, 0.8]$ , there

is a wide interval, and so there is little confidence in what the probability of  $a$  is. We can apply methods for measuring inconsistency in probabilistic logics [20,74,90] and inconsistency-tolerant probabilistic reasoning [4,59,75,80] to analyze and to deal with inconsistencies. From an argumentation viewpoint, we may also tolerate (i.e. ignore) an attack by one argument when the degree of inconsistency is low (e.g. when the entailed probabilities are different but close). In addition, we want to investigate robustness analysis by finding probability distributions that flip an epistemic extension (i.e. change the acceptable arguments), and we want to investigate how evidence can be aggregated so that arguments can combine evidence from multiple sources (e.g. if there are multiple randomized clinical trials involving the same treatments, and they have similar outcomes as represented by the conditionals, how can they be aggregated).

Also in future work, we want to extend the language so that we can consider Boolean combinations of conditionals (including propositional formulae that abbreviate conditionals of the form  $(\phi)[1]$ ) which will allow a richer range of evidence to be represented. We also want to extend the framework to allow representation of statements concerning utility, and allow for the construction of utility theoretic arguments. To do so, we can build up on earlier combinations of probabilistic logic and decision theory [5,65]. This could then offer a more sophisticated approach to argument-based decision making. Utilities appear to play an important role in some kinds of arguments, and even though there are a variety of papers that touch on utility theory in computational models of argument (e.g. [39,58,62,83–85]), how utilities are used within arguments remains to be captured in computational models of argument.

Also, in future work, we want to undertake further analysis of dialectical and epistemic semantics. For the former, we would like to investigate connections between extensions and maximal consistent subsets of an inconsistent knowledgebase following the analyzes for classical logic [96] and for the latter, we would like to investigate further properties for constraining probability assignments based on axioms for epistemic approach to probabilistic argumentation [44]. The latter might also involve introducing support (e.g. if  $A$  supports  $B$ , then  $P(A) \geq P(B)$ , because  $\text{Claim}(A) \in \text{Support}(B)$ ).

### Declaration of competing interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests: Anthony Hunter reports financial support was provided by The Alan Turing Institute. Nico Potyka reports financial support was provided by European Research Council.

### Data availability

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

### Acknowledgements

The research in this paper was partly supported by funding from the Alan Turing Institute for the *Interpretable and Explainable Deep Learning for Natural Language Understanding and Common Sense Reasoning* project and by funding from the European Research Council (ERC) under the European Union Horizon 2020 Research and Innovation Programme (grant agreement No. 101020934) for the *Argumentation-based Deep Interactive eXplanations* project.

### References

- [1] L. Amgoud, J. Ben-Naim, Ranking-based semantics for argumentation frameworks, in: *Proceedings of SUM'13*, in: *Lecture Notes in Artificial Intelligence*, vol. 8078, 2013, pp. 134–147.
- [2] L. Amgoud, J. Ben-Naim, Argumentation-based ranking logics, in: *Proceedings of the 14th International Conference on Autonomous Agents and Multi-agent Systems (AAMAS'15)*, 2015.
- [3] L. Amgoud, C. Cayrol, A reasoning model based on the production of acceptable arguments, *Ann. Math. Artif. Intell.* 34 (1–3) (2002) 197–215.
- [4] M. Adamcik, *Collective reasoning under uncertainty and inconsistency*, PhD thesis, Manchester Institute for Mathematical Sciences, The University of Manchester, 2014.
- [5] E. Acar, C. Meilicke, Multi-attribute decision making using weighted description logics, in: *Proceedings of Workshop on Logics for Reasoning About Preferences, Uncertainty, and Vagueness, (PRUV)*, in: *CEUR Workshop Proceedings*, vol. 1205, 2014, pp. 1–14, CEUR-WS.org.
- [6] A. Ahuja, W. Rödder, Project risk management by a probabilistic expert system, in: *Operations Research Proceedings*, Springer, 2003, pp. 329–334.
- [7] F. Bacchus, *Representing and Reasoning with Probabilistic Knowledge: A Logical Approach to Probabilities*, MIT Press, 1990.
- [8] E. Bonzon, J. Delobelle, S. Konieczny, N. Maudet, A comparative study of ranking-based semantics for abstract argumentation, in: *Proceedings of AAAI'16*, AAAI Press, 2016, pp. 914–920.
- [9] Ph. Besnard, A. Garcia, A. Hunter, S. Modgil, H. Prakken, G. Simari, F. Toni, Introduction to structured argumentation, *Argument Comput.* 5 (1) (2014) 1–4.
- [10] P. Baroni, M. Giacomin, P. Vicig, On rationality conditions for epistemic probabilities in abstract argumentation, in: *Proceedings of COMMA'14*, 2014, pp. 121–132.
- [11] E. Burnside, D. Rubin, R. Shachter, A Bayesian network for mammography, in: *Proceedings of the AMIA Symposium*, American Medical Informatics Association, 2000, p. 106.
- [12] P. Baroni, M. Romano, F. Toni, M. Aurisicchio, G. Bertanza, Automatic evaluation of design alternatives with quantitative argumentation, *Argument Comput.* 6 (1) (2015) 24–49.
- [13] P. Baroni, A. Rago, F. Toni, From fine-grained properties to broad principles for gradual argumentation: a principled spectrum, *Int. J. Approx. Reason.* 105 (2019) 252–286.
- [14] F.G. Cozman, L.F. di Ianni, Probabilistic satisfiability and coherence checking through integer programming, in: *Proceedings of ECSQARU'13*, in: *LNCS*, vol. 7958, Springer, 2013, pp. 145–156.



- [15] M. Caminada, D. Gabbay, A logical account of formal argumentation, *Stud. Log.* 93 (2009) 109–145.
- [16] K. Cyras, Q. Heinrich, F. Toni, Computational complexity of flat and generic assumption-based argumentation, with and without probabilities, *Artif. Intell.* 293 (2021) 103449.
- [17] C. Cayrol, M.-C. Lagasquie-Schiex, Graduality in argumentation, *J. Artif. Intell. Res.* 23 (2005) 245–297.
- [18] K. Cyras, T. Oliveira, A. Karamlou, F. Toni, Assumption-based argumentation with preferences and goals for patient-centric reasoning with interacting clinical guidelines, *Argument Comput.* 12 (2) (2021) 149–189.
- [19] G. De Bona, F.G. Cozman, M. Finger, Towards classifying propositional probabilistic logics, *J. Appl. Log.* 12 (3) (2014) 349–368.
- [20] G. De Bona, M. Finger, Measuring inconsistency in probabilistic logic: rationality postulates and Dutch book interpretation, *Artif. Intell.* 227 (2015) 140–164.
- [21] P. Dellunde, L. Godo, A. Vidal, Probabilistic argumentation: an approach based on conditional probability – a preliminary report, in: *Proceedings of JELIA'21*, in: *Lecture Notes in Computer Science*, vol. 12678, Springer, 2021, pp. 25–32.
- [22] L. Demey, B. Kooi, J. Sack, Logic and probability, in: Edward N. Zalta (Ed.), *The Stanford Encyclopedia of Philosophy*, 2019 edition, Metaphysics Research Lab, Stanford University, Summer 2019.
- [23] W. Dvorák, M. König, M. Ulbricht, St. Woltran, Rediscovering argumentation principles utilizing collective attacks, in: *Proceedings of KR'22*, 2022.
- [24] P. Dung, P. Thang, Towards (probabilistic) argumentation for jury-based dispute resolution, in: *Proceedings of COMMA'10*, IOS Press, 2010, pp. 171–182.
- [25] C. da Costa Pereira, A. Tettamanzi, S. Villata, Changing one's mind: erase or rewind? Possibilistic belief revision with fuzzy argumentation based on trust, in: *Proceedings of IJCAI'11*, 2011.
- [26] P.M. Dung, On the acceptability of arguments and its fundamental role in nonmonotonic reasoning, logic programming and n-person games, *Artif. Intell.* 77 (2) (1995) 321–358.
- [27] M. Finger, G. De Bona, Probabilistic satisfiability: logic-based algorithms and phase transition, in: *Proceedings of IJCAI'11*, 2011, pp. 528–533.
- [28] B. Fisher, J. Dignam, N. Wolmark, A. DeCillis, B. Emir, D. Wickerham, J. Bryant, M. Dimitrov, N. Abramson, J. Atkins, H. Shibata, L. Deschenes, R. Margolese, Tamoxifen and chemotherapy for lymph node-negative, estrogen receptor-positive breast cancer, *J. Natl. Cancer Inst.* 89 (1997) 1673–1682.
- [29] A.M. Frisch, P. Haddawy, Anytime deduction for probabilistic logic, *Artif. Intell.* 69 (1–2) (1994) 93–122.
- [30] J. Fisseler, First-order probabilistic conditional logic and maximum entropy, *Log. J. IGPL* 20 (5) (2012) 796–830.
- [31] O. Gladkova, C. DiMarco, R. Allen Harris, Argumentative meanings and their stylistic configurations in clinical research publications, *Argument Comput.* 6 (2015) 310–346.
- [32] G. Georgakopoulos, D. Kavvadias, C.H. Papadimitriou, Probabilistic satisfiability, *J. Complex.* 4 (1) (1988) 1–11.
- [33] D. Gabbay, O. Rodrigues, Probabilistic argumentation: an equational approach, *Log. Univers.* 9 (3) (2015) 345–382.
- [34] A. Hackshaw, *A Concise Guide to Clinical Trials*, Wiley Blackwell, 2009.
- [35] A. Hackshaw, *A Concise Guide to Observational Studies in Healthcare*, Wiley Blackwell, 2015.
- [36] R. Haenni, Modelling uncertainty with propositional assumptions-based systems, in: *Applications of Uncertainty Formalisms*, in: LNCS, vol. 1455, Springer, 1998, pp. 446–470.
- [37] T. Hailperin, *Boole's Logic and Probability: a Critical Exposition from the Standpoint of Contemporary Algebra, Logic and Probability Theory*, Elsevier, 1986.
- [38] J. Halpern, An analysis of first-order logics of probability, *Artif. Intell.* 46 (3) (1990) 311–350.
- [39] E. Hadoux, A. Hunter, Strategic sequences of arguments for persuasion using decision trees, in: *Proceedings of AAI'17*, AAI Press, 2017, pp. 1128–1134.
- [40] D. Heckerman, E. Horvitz, B. Nathwani, Toward normative expert systems: Part I the pathfinder project, *Methods Inf. Med.* 31 (02) (1992) 90–105.
- [41] P. Hansen, B. Jaumard, Probabilistic satisfiability, in: *Handbook of Defeasible Reasoning and Uncertainty Management Systems*, Springer, 2000, pp. 321–367.
- [42] X. Huang, J. Lin, D. Demner-Fushman, Evaluation of pico as a knowledge representation for clinical questions, in: *Proceedings of AIMA'06*, 2006, pp. 359–363.
- [43] P. Hansen, S. Perron, Merging the local and global approaches to probabilistic satisfiability, *Int. J. Approx. Reason.* 47 (2) (2008) 125–140.
- [44] A. Hunter, M. Thimm, Probabilistic reasoning with abstract argumentation frameworks, *J. Artif. Intell. Res.* 59 (2017) 565–611.
- [45] A. Hunter, Some foundations for probabilistic abstract argumentation, in: *Proceedings of COMMA'12*, IOS Press, 2012, pp. 117–128.
- [46] A. Hunter, A probabilistic approach to modelling uncertain logical arguments, *Int. J. Approx. Reason.* 54 (1) (2013) 47–81.
- [47] A. Hunter, Generating instantiated argument graphs from probabilistic information, in: *Proceedings of ECAI'20*, in: FAIA, vol. 325, IOS Press, 2020, pp. 769–776.
- [48] A. Hunter, Reasoning with inconsistent knowledge using the epistemic approach to probabilistic argumentation, in: *Proceedings of KR'20*, 2020, pp. 496–505.
- [49] A. Hunter, M. Williams, Aggregating evidence about the positive and negative effects of treatments, *Artif. Intell. Med.* 56 (2012) 173–190.
- [50] N. Käfer, C. Baier, M. Diller, C. Dubslaff, S. Gaggl, H. Hermanns, Admissibility in probabilistic argumentation, *J. Artif. Intell. Res.* 74 (2022).
- [51] G. Kern-Isberner, M. Thimm, Novel semantical approaches to relational probabilistic conditionals, in: *Proceedings of KR'10*, AAI Press, 2010.
- [52] C. Kah, L. Roberts, K. Shaffer, P. Haddawy, Construction of a Bayesian network for mammographic diagnosis of breast cancer, *Comput. Biol. Med.* 27 (1) (1997) 19–29.
- [53] H. Li, N. Oren, T. Norman, Probabilistic argumentation frameworks, in: *Proceedings of TAFE'11*, 2011.
- [54] T. Lukasiewicz, Probabilistic deduction with conditional constraints over basic events, *J. Artif. Intell. Res.* 10 (1999) 199–241.
- [55] T. Lukasiewicz, Probabilistic logic programming with conditional constraints, *ACM Trans. Comput. Log.* 2 (3) (2000) 289–339.
- [56] M. Morveli-Espinoza, J.C. Nieves, A. Possebom, C.A. Tacla, Dealing with incompatibilities among procedural goals under uncertainty, *Intel. Artif.* 22 (64) (2019) 47–62.
- [57] T. Mayer, S. Marro, E. Cabrio, S. Villata, Enhancing evidence-based medicine with natural language argumentative analysis of clinical trials, *Artif. Intell. Med.* 118 (2021) 102098.
- [58] P.-A. Matt, F. Toni, A game-theoretic measure of argument strength for abstract argumentation, in: *Proceedings of JELIA'08*, in: LNCS, vol. 5293, Springer, 2008, pp. 285–297.
- [59] D.P. Muiño, Measuring and repairing inconsistency in probabilistic knowledge bases, *Int. J. Approx. Reason.* 52 (6) (2011) 828–840.
- [60] N. Nilsson, Probabilistic logic, *Artif. Intell.* 28 (1995) 71–87.
- [61] S. Nielsen, S. Parsons, A generalization of Dung's abstract framework for argumentation: arguing with sets of attacking arguments, in: *Proceedings of ArgMAS '06*, in: LNCS, vol. 4766, Springer, 2006, pp. 54–73.
- [62] N. Oren, T.J. Norman, Arguing using opponent models, in: Peter McBurney, Iyad Rahwan, Simon Parsons, Nicolas Maudet (Eds.), *Proceedings of ArgMAS'09*, in: LNCS, vol. 6057, Springer, 2009, pp. 160–174.
- [63] N. Oren, T. Norman, A. Preece, Subjective logic and arguing with evidence, *Artif. Intell.* 171 (10–15) (2007) 838–854.
- [64] J.B. Paris, *The Uncertain Reasoner's Companion – A Mathematical Perspective*, Cambridge University Press, 1994.
- [65] N. Potyka, E. Acar, M. Thimm, H. Stuckenschmidt, Group decision making via probabilistic belief merging, in: *Proceedings of IJCAI'16*, IJCAI/AAAI Press, 2016, pp. 3623–3629.

- [66] N. Pfeifer, C. Fermüller, Probabilistic interpretations of argumentative attacks: logical and experimental foundations, in: Proceedings of WUPES'18, MatfyzPress Publishing House, 2018, pp. 141–152.
- [67] N. Pfeifer, On argument strength, in: Bayesian Argumentation, in: Synthese Library, vol. 362, Springer, 2012, pp. 185–193.
- [68] S. Polberg, A. Hunter, Empirical evaluation of abstract argumentation: supporting the need for bipolar and probabilistic approaches, *Int. J. Approx. Reason.* 93 (2018) 487–543.
- [69] S. Polberg, A. Hunter, M. Thimm, Belief in attacks in epistemic probabilistic argumentation, in: Proceedings of SUM'17, in: LNCS, vol. 10564, Springer, 2017, pp. 223–236.
- [70] N. Pfeifer, G.D. Kleiter, Framing human inference by coherence based probability logic, *J. Appl. Log.* 7 (2) (2009) 206–217.
- [71] F. Pu, J. Luo, Y. Zhang, G. Luo, Argument ranking with categoriser function, in: Proceedings of KSEM'14, in: Lecture Notes in Computer Science, vol. 8793, Springer, 2014, pp. 290–301.
- [72] F. Pu, J. Luo, Y. Zhang, G. Luo, Attacker and defender counting approach for abstract argumentation, in: Proceedings of CogSci'15, 2015, p. 1, cognitive-sciencesociety.org.
- [73] J. Pollock, *Cognitive Carpentry*, MIT Press, 1995.
- [74] N. Potyka, Linear programs for measuring inconsistency in probabilistic logics, in: Proceedings of KR'14, AAAI Press, 2014, pp. 568–578.
- [75] N. Potyka, Reasoning over linear probabilistic knowledge bases with priorities, in: Proceedings of SUM'15, in: LNCS, vol. 9310, Springer, 2015, pp. 121–136.
- [76] N. Potyka, Solving reasoning problems for probabilistic conditional logics with consistent and inconsistent information, PhD thesis, FernUniversität in Hagen, 2016.
- [77] N. Potyka, Continuous dynamical systems for weighted bipolar argumentation, in: Proceedings of KR'18, 2018, pp. 148–157.
- [78] N. Potyka, Interpreting neural networks as gradual argumentation frameworks, in: Proceedings of AAAI'21, 2021, pp. 6463–6470.
- [79] H. Prakken, Probabilistic strength of arguments with structure, in: Proceedings of KR'18, AAAI Press, 2018, pp. 158–167.
- [80] N. Potyka, M. Thimm, Probabilistic reasoning with inconsistent beliefs using inconsistency measures, in: Proceedings of IJCAI'15, AAAI Press, 2015, pp. 3156–3163.
- [81] J. Paris, Alena Vencovská, Proof systems for probabilistic uncertain reasoning, *J. Symb. Log.* 63 (3) (1998) 1007–1039.
- [82] R. Riveret, G. Governatori, On learning attacks in probabilistic abstract argumentation, in: Proceedings of AAMAS'16, 2016.
- [83] I. Rahwan, M. Madakkatel, J. Bonnefon, R. Awan, S. Abdallah, Behavioural experiments for assessing the abstract argumentation semantics of reinstatement, *Cogn. Sci.* 34 (8) (2010) 1483–1502.
- [84] I. Rahwan, P. Pasquier, L. Sonenberg, F. Dignum, A formal analysis of interest-based negotiation, *Ann. Math. Artif. Intell.* 55 (2009) 253–276.
- [85] R. Riveret, A. Rotolo, G. Sartor, H. Prakken, B. Roth, Success chances in argument games: a probabilistic approach to legal disputes, in: Proceedings of JURIX'07, IOS Press, 2007, pp. 99–108.
- [86] A. Rago, F. Toni, M. Aurisicchio, P. Baroni, Discontinuity-free decision support with quantitative argumentation debates, in: Chitta Baral, James Delgrande, Frank Wolter (Eds.), Proceedings of KR'16, AAAI Press, 2016, pp. 63–73.
- [87] I. Sassoon, N. Kökciyan, S. Modgil, S. Parsons, Argumentation schemes for clinical decision support, *Argument Comput.* 12 (3) (2021) 329–355.
- [88] G. Sanfilippo, N. Pfeifer, A. Gilio, Generalized probabilistic modus ponens, in: Proceedings of ECSQARU'17, Springer, 2017, pp. 480–490.
- [89] M. Thimm, A probabilistic semantics for abstract argumentation, in: Proceedings of ECAI'12, in: FAIA, vol. 242, IOS Press, 2012, pp. 750–755.
- [90] M. Thimm, Inconsistency measures for probabilistic logics, *Artif. Intell.* 197 (April 2013) 1–24.
- [91] Y. Tang, C.-W. Hang, S. Parsons, M. Singh, Towards argumentation with symbolic Dempster-Shafer evidence, in: Proceedings of COMMA'12, in: Frontiers in Artificial Intelligence and Applications, vol. 245, IOS Press, 2012, pp. 462–469.
- [92] J. Varghese, C. Beierle, N. Potyka, G. Kern-Isberner, Using probabilistic logic and the principle of maximum entropy for the analysis of clinical brain tumor data, in: Proceedings of CBMS'13, IEEE, 2013, pp. 401–404.
- [93] B. Verheij, Jumping to conclusions: a logico-probabilistic foundation for defeasible rule-based arguments, in: Proceedings of JELIA'12, in: LNCS, vol. 7519, Springer, 2012, pp. 411–423.
- [94] B. Verheij, Arguments and their strength: revisiting Pollock's anti-probabilistic starting points, in: Proceedings of COMMA'14, 2014.
- [95] B. Verheij, Proof with and without probabilities - correct evidential reasoning with presumptive arguments, coherent hypotheses and degrees of uncertainty, *Artif. Intell. Law* 25 (1) (2017) 127–154.
- [96] S. Vesic, Identifying the class of maxi-consistent operators in argumentation, *J. Artif. Intell. Res.* 47 (2013) 71–93.