Inconsistency Measurement in Probabilistic Logic

Glauber De Bona, Marcelo Finger, Nico Potyka, Matthias Thimm

March 2, 2018

Abstract

We survey the state of the art of inconsistency measurement in probabilistic logics. Compared to the setting of inconsistency measurement in classical logic, the incorporation of probabilistic assessments brings new challenges that have to be addressed by computational accounts to inconsistency measures. For that, we revisit rationality postulates for this setting and discuss the intricacies of probabilistic logics. We give an overview on existing measures and discuss their compliance with the rationality postulates. Finally, we discuss the relationships of inconsistency measures for probabilistic logic with Dutch books from economics.

1 Introduction

In this chapter, we will focus on measuring inconsistency in probabilistic logics. As opposed to classical knowledge bases, we enrich formulas with probabilities. A formula with probability 1 is supposed to be true (like a classical true formula), a formula with probability 0 is supposed to be false (like a classical negated true formula). Probabilities between 0 and 1 express our uncertainty about the truth state. In particular, we are often interested in conditional probabilities.

Example 1. Suppose we want to design an expert system for medical decision support. A group of medical experts is asked about their beliefs about the relationships between particular diseases and corresponding symptoms. Let us assume that the experts state the following beliefs about disease d and symptoms s_1, s_2:

- the probability of a patient with disease d exhibiting both symptom s_1 and symptom s_2 is at least 60%;
- the probability of a patient with disease d exhibiting symptom s_1 but not symptom s_2 is at least 50%;

- the probability of a patient with disease d exhibiting symptom s_1 is at most 80%.

Taken together, the experts' beliefs are inconsistent: according to the first two items, the probability of symptom s_1, given disease d, should be at least 110%. We have to adapt the beliefs in order to restore consistency. How should we proceed? Which pieces of information should be changed to restore consistency? Moreover, which pieces are to blame for the inconsistency, and to which degree? Once chosen which statement to change, should it be deleted or adapted by raising or lowering the probability in order to approximate consistency? These are the kind of questions an inconsistency measure for probabilistic logic can help to answer. $\square$

In the following, we will mainly focus on propositional probabilistic knowledge bases to keep things simple. Our logical language is similar to the ones considered in [28, 29, 24, 22]. Many ideas can be transfered to relational languages. We will give some examples and further references as we proceed.

We will start our discussion with a quick introduction to propositional probabilistic logics. Subsequently, we will introduce a collection of rationality postulates in Section 3. While many of these properties are direct translations from the classical setting, we will consider additional postulates that take the role of probabilities into account. One important postulate in this context is *Continuity*, which basically states that minor changes in probabilities should not yield major changes in the degree of inconsistency. Interestingly, Continuity is in conflict with some classical postulates. In Section 4, we will then discuss six approaches to measure inconsistency in probabilistic logics. We will start with measures that are inspired by classical approaches and then look at several measures that make stronger use of the probabilities in the knowledge bases. We compare these measures with respect to the postulates that they satisfy or violate. In Section 5, we will sketch some applications of inconsistency measures. We will briefly discuss how they can be used to repair inconsistent knowledge bases and to reason with knowledge bases that contain conflicts of different kinds.

2 Preliminaries

We consider a propositional language built up over a finite set of *atomic propositions (atoms)* $\mathcal{A} = \{a_1, \ldots, a_n\}$ in the usual way. That is, formulas are constructed inductively by connecting atomic propositions with logical connectives like $\neg, \wedge, \vee, \rightarrow$. $\mathcal{L}_{\mathcal{A}}$ denotes the set of all well-formed propositonal formulas over $\mathcal{A}$. Additionally, $\top$ denotes a tautology $a \vee \neg a$ for some $a \in \mathcal{A}$, and $\bot$ denotes a contradiction $\neg\top$.

A *possible world* w over $\mathcal{A}$ is a conjunction of $|\mathcal{A}| = n$ literals containing either a or $\neg a$ for each $a \in \mathcal{A}$. For instance, if $\mathcal{A} = \{a, b, c\}$, then $a \wedge b \wedge c$ and

$a \wedge b \wedge \neg c$ are two of the $2^3 = 8$ possible worlds over $\{a, b, c\}$. We denote by $\mathcal{W}_{\mathcal{A}} = \{w_1, \ldots, w_{2^n}\}$ the set of all possible worlds over $\mathcal{A}$. $w \in \mathcal{W}_{\mathcal{A}}$ *satisfies* $a \in \mathcal{A}$ ($w \models a$) iff a is a positive literal in w (a is not negated in w). For instance, $a \wedge b \wedge \neg c$ satisfies a and b but falsifies c. $\models$ can be extended to all $\varphi \in \mathcal{L}_{\mathcal{A}}$ recursively as usual. For instance, $w \models \neg\varphi$ iff $w \models \varphi$ does not hold, $w \models \varphi_1 \wedge \varphi_2$ iff $w \models \varphi_1$ and $w \models \varphi_2$ and so on.

A *probabilistic conditional* (or simply *conditional*) is an expression of the form $(\varphi|\psi)[\underline{q}, \bar{q}]$, where $\varphi, \psi \in \mathcal{L}_{\mathcal{A}}$ are propositional formulas and $\underline{q}, \bar{q} \in [0, 1] \cap \mathbb{Q}$ are rational numbers with $\underline{q} \leq \bar{q}$. Intuitively, $(\varphi|\psi)[\underline{q}, \bar{q}]$ says that "the probability that φ is true given that ψ is true lies within the interval $[\underline{q}, \bar{q}]$". $(\mathcal{L}_{\mathcal{A}} \mid \mathcal{L}_{\mathcal{A}})$ denotes the set of all conditonals over $\mathcal{A}$. A conditional $(\varphi|\psi)[q, q]$ with equal lower and upper bound is called *precise* and denoted by $(\varphi|\psi)[q]$ to improve readability. Similarly, if the condition is tautological, we write $(\varphi)[\underline{q}, \bar{q}]$ rather than $(\varphi|\top)[\underline{q}, \bar{q}]$ and call $(\varphi)[\underline{q}, \bar{q}]$ an *unconditional probabilistic assessment*.

A *probabilistic interpretation* $\pi : \mathcal{W}_{\mathcal{A}} \to [0, 1]$, with $\sum_j \pi(w_j) = 1$, is a probability mass over the set of possible worlds. Each probabilistic interpretation π induces a probability measure $P_\pi : \mathcal{L}_{\mathcal{A}} \to [0, 1]$ by means of $P_\pi(\varphi) = \sum_{w_j \models \varphi} \pi(w_j)$. Let $\mathcal{P}(\mathcal{A})$ be the set of all probabilistic interpretations $\pi : \mathcal{W}_{\mathcal{A}} \to [0, 1]$. A conditional $(\varphi|\psi)[\underline{q}, \bar{q}]$ is *satisfied* by π, also denoted by $\pi \models (\varphi|\psi)[\underline{q}, \bar{q}]$ iff $P_\pi(\varphi \wedge \psi) \geq \underline{q} P_\pi(\psi)$ and $P_\pi(\varphi \wedge \psi) \leq \bar{q} P_\pi(\psi)$. Note that when $P_\pi(\psi) > 0$, a probabilistic conditional $(\varphi|\psi)[\underline{q}, \bar{q}]$ is constraining the conditional probability of φ given ψ; but any π with $P_\pi(\psi) = 0$ trivially[1] satisfies the conditional $(\varphi|\psi)[\underline{q}, \bar{q}]$ (this semantics is adopted by Halpern [16], Frisch and Haddawy [13] and Lukasiewicz [24], for instance). For a conditional c we denote by $\mathsf{Mod}(c)$ the set of models of c, i. e., $\mathsf{Mod}(c) = \{\pi \mid \pi \models c\}$.

In order to take account of information that was presented multiple times, we regard a *knowledge base* as a finite multiset κ of probabilistic conditionals. Formally, a knowledge base κ is defined by a *multiplicity function* $M_\kappa : (\mathcal{L}_{\mathcal{A}} \mid \mathcal{L}_{\mathcal{A}}) \to \mathbb{N}$ such that $M_\kappa(C) > 0$ for only finitely many $C \in (\mathcal{L}_{\mathcal{A}} \mid \mathcal{L}_{\mathcal{A}})$. $M_\kappa(C)$ is the number of occurrences of C in κ. Let $\mathbb{K}$ denote the set of all knowledge bases. We write $C \in \kappa$ iff $M_\kappa(C) > 0$. For two knowledge bases κ_1, κ_2 we define multiset union $\cup$, multiset intersection $\cap$, and multiset difference $\setminus$ via

$$
\begin{aligned}
M_{\kappa_1 \cup \kappa_2}(C) &= M_{\kappa_1}(C) + M_{\kappa_2}(C) \\
M_{\kappa_1 \cap \kappa_2}(C) &= \min\{M_{\kappa_1}(C), M_{\kappa_2}(C)\} \\
M_{\kappa_1 \setminus \kappa_2}(C) &= \max\{0, M_{\kappa_1}(C) - M_{\kappa_2}(C)\}
\end{aligned}
$$

for all $C \in (\mathcal{L}_{\mathcal{A}} \mid \mathcal{L}_{\mathcal{A}})$. The cardinality of κ is $|\kappa| = \sum_{C \in (\mathcal{L}_{\mathcal{A}} | \mathcal{L}_{\mathcal{A}})} M_\kappa(C)$. If κ is such that M_κ is non-zero exactly for $C_1, \ldots, C_m$, we denote κ by $\{C_1 : M_\kappa(C_1), \ldots, C_m : M_\kappa(C_m)\}$.

[1] An approach that does not trivialize when $P_\pi(\psi) = 0$ can be found in [6].

Example 2. $\kappa = \{(a)[0.2] : 2, a[0.8] : 1\}$ denotes the knowledge base that contains two instances of the conditional $(a)[0.2]$, one instance of the conditional $a[0.8]$ and no other conditionals. We have

$$\begin{aligned}&\{(a)[0.2] : 2, (a)[0.8] : 1\} \cup \{(a)[0.8] : 1, (b)[0.5] : 1\}\\ =&\{(a)[0.2] : 2, (a)[0.8] : 2, (b)[0.5] : 1\}\end{aligned}$$

and

$$\{(a)[0.2] : 2, (a)[0.8] : 1\} \setminus \{(a)[0.2] : 1\} = \{(a)[0.2] : 1, (a)[0.8] : 1\}.$$

If κ is an ordinary set in the sense that $M_\kappa(C) = 1$ for all $C \in \kappa$, we omit the postfix : 1 to improve readability. For each knowledge base κ we assume some arbitrary but fixed *canonical enumeration* $\langle\kappa\rangle = (c_1, \dots, c_m)$ that represents κ as a sequence of its elements (including duplicates).

A *probabilistic interpretation* $\pi : \mathcal{W}_\mathcal{A} \to [0, 1]$ satisfies a knowledge base κ, denoted by $\pi \models \kappa$, if $\pi \models c$ for all $c \in \kappa$. We let $\mathsf{Mod}(\kappa) = \{\pi \mid \pi \models \kappa\}$ and call the elements of $\mathsf{Mod}(\kappa)$ the models of κ. A knowledge base κ is *consistent* (or *satisfiable*) if $\mathsf{Mod}(\kappa) \neq \varnothing$. κ is precise if all conditionals in κ are precise.

Knowledge bases κ_1, κ_2 are *extensionally equivalent*, denoted by $\kappa_1 \equiv^e \kappa_2$, if and only if $\mathsf{Mod}(\kappa_1) = \mathsf{Mod}(\kappa_2)$. Knowledge bases κ_1, κ_2 are *semi-extensionally equivalent* [45], denoted by $\kappa_1 \equiv^s \kappa_2$, if and only if there is a bijection $\rho_{\kappa_1,\kappa_2} : \kappa_1 \to \kappa_2$ such that $c \equiv^e \rho_{\kappa_1,\kappa_2}(c)$ for every $c \in \kappa_1$. This means that two knowledge bases κ_1 and κ_2 are semi-extensionally equivalent if we find a mapping between the conditionals of both knowledge bases such that each conditional of κ_1 is extensionally equivalent to its image in κ_2. Note that $\kappa_1 \equiv^s \kappa_2$ implies $\kappa_1 \equiv^e \kappa_2$ [45].

3 Rationality Postulates

There are many ways of measuring the inconsistency of a set of formulas in some formal language. For example, we may have an idiosyncratic measurement that maps every consistent set to -3.2 and every inconsistent set to 7. Besides the arbitrariness of those values, such a measurement does not allow one to express that one theory is "more inconsistent" than another. And yet, this is something one may want to express. For instance, the knowledge base $\kappa_1 = \{(p)[0.5], (\neg p)[0.5001]\}$ seems "less inconsistent" than $\kappa_2 = \{(p)[1], (\neg p)[1]\}$ because the probabilities in κ_2 need to be adjusted more drastically than those in κ_1. If we are mainly interested in measuring the inconsistency of a knowledge base, it is reasonable to postulate that every consistent knowledge base is associated to the same measurement, say 0; similarly, we would expect that the degree of inconsistency of κ_1 is lower than the one of κ_2. In the following, we will discuss a collection of *rationality postulates* that have been proposed for inconsistency measures. As we will see, not all of these postulates can be satisfied simultaneously.

To begin with, we introduce inconsistency measures for probabilistic logics and some additional terminology. Inconsistency measures for probabilistic knowledge bases are defined analogously to those for classical knowledge bases.

Definition 1. An inconsistency measure $\mathcal{I}$ is a function $\mathcal{I} : \mathbb{K} \to [0, \infty)$.

The value $\mathcal{I}(\kappa)$ for a knowledge base κ is called the *inconsistency value* of κ with respect to $\mathcal{I}$. Minimal inconsistent sets of probabilistic knowledge bases are defined analogously to their classical counterparts.

Definition 2. A set $\mathcal{M}$ of probabilistic conditionals is *minimal inconsistent* if $\mathcal{M}$ is inconsistent and every $\mathcal{M}' \subsetneq \mathcal{M}$ is consistent.

We let $\mathsf{MI}(\kappa)$ denote the set of the minimal inconsistent subsets of $\kappa \in \mathbb{K}$. Intuitively, the conditionals in a minimal inconsistent subset of a knowledge base are those that are responsible for an atomic conflict.

Example 3. Consider the knowledge base

$$\kappa = \{(a)[0.2] : 1, (a)[0.8] : 1, (a \wedge b)[0.6] : 1, (b \vee d)[1] : 1, (c)[0.5] : 1\}$$

Here, we have two minimal inconsistent subsets $\{(a)[0.2] : 1, (a)[0.8] : 1\}$ and $\{(a)[0.2] : 1, (a \wedge b)[0.6] : 1\}$.

Conditionals that do not take part in such a conflict are called *free*.

Definition 3. A probabilistic conditional $c \in \kappa$ is *free* in κ if and only if $c \notin \mathcal{M}$ for all $\mathcal{M} \in \mathsf{MI}(\kappa)$.

For a conditional or a knowledge base C let $\mathcal{A}(C) \subseteq \mathcal{A}$ denote the set of atoms appearing in C. A conditional is safe with respect to a knowledge base κ if it does not share any atoms with κ [45].

Definition 4. A probabilistic conditional $c \in \kappa$ is *safe* in κ if and only if $\mathcal{A}(c) \cap \mathcal{A}(\kappa \setminus \{c\}) = \varnothing$.

The notion of a free conditional is more general than the notion of a safe conditional [45].

Proposition 1. If c is safe in κ then c is free in κ.

Example 4. We continue Example 3. Here, $(b \vee d)[1] : 1$ is a free formula and $(c)[0.5] : 1$ is both free and safe.

We first consider a set of qualitative postulates from [45] that have direct counterparts for classical knowledge bases. κ, κ' denote knowledge bases and c a probabilistic conditional.

Consistency κ is consistent if and only if $\mathcal{I}(\kappa) = 0$

Monotonicity $\mathcal{I}(\kappa) \leq \mathcal{I}(\kappa \cup \{c\})$

Super-additivity If $\kappa \cap \kappa' = \varnothing$ then $\mathcal{I}(\kappa \cup \kappa') \geq \mathcal{I}(\kappa) + \mathcal{I}(\kappa')$

Weak independence If $c \in \kappa$ is safe in κ then $\mathcal{I}(\kappa) = \mathcal{I}(\kappa \setminus \{c\})$

Independence If $c \in \kappa$ is free in κ then $\mathcal{I}(\kappa) = \mathcal{I}(\kappa \setminus \{c\})$

Penalty If $c \in \kappa$ is not free in κ then $\mathcal{I}(\kappa) > \mathcal{I}(\kappa \setminus \{c\})$

Irrelevance of syntax If $\kappa_1 \equiv^s \kappa_2$ then $\mathcal{I}(\kappa_1) = \mathcal{I}(\kappa_2)$

MI-separability If $\mathsf{MI}(\kappa_1 \cup \kappa_2) = \mathsf{MI}(\kappa_1) \cup \mathsf{MI}(\kappa_2)$ and $\mathsf{MI}(\kappa_1) \cap \mathsf{MI}(\kappa_2) = \varnothing$ then $\mathcal{I}(\kappa_1 \cup \kappa_2) = \mathcal{I}(\kappa_1) + \mathcal{I}(\kappa_2)$

Normalisation $\mathcal{I}(\kappa) \in [0, 1]$

The property *consistency* demands that $\mathcal{I}(\kappa)$ takes the minimal value 0 if and only if κ is consistent. *Monotonicity* demands that $\mathcal{I}$ is non-decreasing under the addition of new information. *Super-additivity* strengthens this condition for disjoint knowledge bases. The properties *weak independence* and *independence* say that the inconsistency value should remain unchanged when adding "innocent" information. *Penalty* is the counterpart of *independence* and demands that adding inconsistent information increases the inconsistency value. *Irrelevance of syntax* states that the inconsistency value should not depend on the syntactic representation of conditionals. We use the equivalence relation $\equiv^s$ here since all inconsistent knowledge bases are equivalent with respect to $\equiv^e$. For an inconsistency measure $\mathcal{I}$, imposing irrelevance of syntax to hold in terms of $\equiv^e$ would yield $\mathcal{I}(\kappa) = \mathcal{I}(\kappa')$ for every two inconsistent knowledge bases κ, κ'. The property *MI-separability* states that determining the value of $\mathcal{I}(\kappa_1 \cup \kappa_2)$ can be split into determining the values of $\mathcal{I}(\kappa_1)$ and $\mathcal{I}(\kappa_2)$ if the minimal inconsistent subsets of $\kappa_1 \cup \kappa_2$ correspond to the disjoint union of those of κ_1 and κ_2. *Normalisation* states that inconsistency values should be bounded from above by one.

The following proposition states some relationships between these properties. The proof can be found in [45].

Proposition 2. Let $\mathcal{I}$ be an inconsistency measure and let κ, κ' be some knowledge bases.

1. If $\mathcal{I}$ satisfies *super-additivity* then $\mathcal{I}$ satisfies *monotonicity.*

2. If $\mathcal{I}$ satisfies *independence* then $\mathcal{I}$ satisfies *weak independence.*

3. If $\mathcal{I}$ satisfies *MI-separability* then $\mathcal{I}$ satisfies *independence.*

4. $\kappa \subseteq \kappa'$ implies $\mathsf{MI}(\kappa) \subseteq \mathsf{MI}(\kappa')$.

5. If $\mathcal{I}$ satisfies *independence* then $\mathsf{MI}(\kappa) = \mathsf{MI}(\kappa')$ implies $\mathcal{I}(\kappa) = \mathcal{I}(\kappa')$.

6. If $\mathcal{I}$ satisfies *independence* and *penalty* then $\mathsf{MI}(\kappa) \subsetneq \mathsf{MI}(\kappa')$ implies $\mathcal{I}(\kappa) < \mathcal{I}(\kappa')$.

Our qualitative properties do not take the crucial role of probabilities into account. In order to account for these we need some further notation. Let κ be a knowledge base. For $\vec{x} \in [0,1]^{2|\kappa|}$ we denote by $\kappa[\vec{x}]$ the knowledge base that is obtained from κ by replacing the probabilities of the conditionals in κ by the values in $\vec{x}$. More precisely, if $\langle \kappa \rangle = ((\varphi_1|\psi_1)[\underline{q}_1, \bar{q}_1], \ldots, (\varphi_n|\psi_n)[\underline{q}_n, \bar{q}_n])$ then $\langle \kappa[\vec{x}] \rangle = ((\varphi_1|\psi_1)[x_1, x_2], \ldots, (\varphi_n|\psi_n)[x_{2n-1}, x_{2n}])$ for $\vec{x} = \langle x_1, \ldots, x_{2n} \rangle \in [0,1]^{2n}$. Similarly, for a single probabilistic conditional $c = (\varphi|\psi)[\underline{q}, \bar{q}]$ and $x_1, x_2 \in [0,1]$ we abbreviate $c[x_1, x_2] = (\varphi|\psi)[x_1, x_2]$. The characteristic function of a knowledge base κ takes a probability vector $\vec{x} \in [0,1]^{2|\kappa|}$ and replaces the probabilities in κ accordingly. The formal definition makes use of the order on the probabilistic conditionals of a knowledge base that we discussed in Section 2.

Definition 5. Let $\kappa \in \mathbb{K}$ be a knowledge base. The function $\Lambda_\kappa : [0,1]^{2|\kappa|} \to \mathbb{K}$ with $\Lambda_\kappa(\vec{x}) = \kappa[\vec{x}]$ is called the *characteristic function* of κ.

The characteristic inconsistency function is composed of the characteristic function and an inconsistency measure and shows how different probability vectors $\vec{x} \in [0,1]^{|\kappa|}$ affect the inconsistency value.

Definition 6. Let $\mathcal{I}$ be an inconsistency measure and let $\kappa \in \mathbb{K}$ be a knowledge base. The function

$$\theta_{\mathcal{I},\kappa} : [0,1]^{2|\kappa|} \to [0,\infty)$$

with $\theta_{\mathcal{I},\kappa} = \mathcal{I} \circ \Lambda_\kappa$ is called the *characteristic inconsistency function* of $\mathcal{I}$ and κ.

The following property *continuity* [45] describes our main demand for continuous inconsistency measurement, i. e., a “slight” change in the knowledge base should not result in a “vast” change of the inconsistency value.

Continuity For any $\vec{y} \in [0,1]^{2|\kappa|}$, $\lim_{\vec{x} \to \vec{y}} \theta_{\mathcal{I},\kappa}(\vec{x}) = \theta_{\mathcal{I},\kappa}(\vec{y})$

The above property demands a certain *smoothness* of the behavior of $\mathcal{I}$. Given a fixed set of probabilistic conditionals this property demands that changes in the *quantitative* part of the conditionals trigger a continuous change in the inconsistency value. Note that we require the *qualitative* part of the conditionals, i. e. premises and conclusions of the conditionals, to be fixed. This makes this property not applicable for the classical setting. In the probabilistic setting satisfaction of this property is helpful for the knowledge engineer in restoring consistency. Observe that for every knowledge base $\kappa \in \mathbb{K}$ there is always a $\vec{x} \in [0,1]^{2|\kappa|}$ such that $\kappa[\vec{x}]$ is consistent, cf. [43].

Even though the property *continuity* is a natural requirement in the probabilistic setting, it is incompatible with two of our qualitative postulates.

Proposition 3. Let $\mathcal{I}$ be an inconsistency measure and let κ, κ' be some knowledge bases.

1. There is no $\mathcal{I}$ that satisfies *consistency*, *independence*, and *continuity*.
2. There is no $\mathcal{I}$ that satisfies *consistency*, *MI-separability*, and *continuity*.

The proof can be found in [9]. These incompatibility results suggest that, in order to drive the rational choice of an inconsistency measure for probabilistic knowledge bases, we must abandon at least one postulate among *consistency*, *independence* and *continuity* (recall that *MI-separability* entails *independence*). The property *consistency* seems to be indisputable since the least one can expect from an inconsistency measure is that it separates inconsistent from consistent cases, or some inconsistency from none. Therefore, we should give up either *independence* or *continuity*. A simple solution is to give up *independence* for its weaker version *weak independence* that is compatible with *consistency* and *continuity* [9].

In fact, there are more compelling reasons for giving up *independence* rather than *continuity*. The property *independence* was introduced based on the notion that minimal inconsistent subsets are the purest form of inconsistency [20], capturing all its causes in a knowledge base [19]. This notion can be traced back to the work of Reiter on the diagnosis problem [37] and to the standard AGM framework of belief revision [1], where minimal inconsistent subsets have a central role. Nevertheless, in the probabilistic case, minimal inconsistent subsets may fail to detect all causes of inconsistency, as the next example illustrates.

Example 5. Recall the situation in Example 1, formalized into the knowledge base $\kappa = \{(s_1 \wedge s_2)[0.6,1], (s_1 \wedge \neg s_2)[0.5,1], (s_1)[0,0.8]\}$. Suppose we want to schedule a meeting among the 3 different experts responsible for these assignments in order to reconcile them. To save resources, we plan to invite only the physicians whose probabilistic assessments are somehow causing the inconsistency. If minimal inconsistent subsets are supposed to capture such causes, the third physician would not be invited, for $\mathsf{MI}(\kappa) = \{\{(s_1 \wedge s_2)[0.6,1], (s_1 \wedge \neg s_2)[0.5,1]\}\}$, and $(s_1)[0,0.8]$ is free in κ. Suppose the expert who elicited the first conditional, $(s_1 \wedge s_2)[0.6,1]$, admits that the lower bound is rather high and relaxes it to $(s_1 \wedge s_2)[0.5,1]$, being compatible with $(s_1 \wedge \neg s_2)[0.5,1]$. Nonetheless, the updated knowledge base $\kappa' = \{(s_1 \wedge s_2)[0.5,1], (s_1 \wedge \neg s_2)[0.5,1], (s_1)[0,0.8]\}$ would still be inconsistent, for the first two conditionals imply $(s_1)[1]$, contradicting the third one. □

The arguments above indicate that we should give up *independence* rather than *continuity*. In this case, we can still demand *weak independence* consistently. The formulation of a weak form of continuity that could be consistent with *independence* seems a harder and less natural alternative.

4 Approaches

In the following we survey approaches to inconsistency measures for probabilistic logics from the literature.

4.1 Classical approaches

We start with existing approaches to inconsistency measurement for classical logic and adapt those to the probabilistic case, see also [45] where those measures were adapted to probabilistic conditional logic with precise probabilities. In particular, we have a look at the drastic inconsistency measure, the MI inconsistency measure, the MI^C inconsistency measure, and the η-inconsistency measure, see e. g. [20, 23] for the classical definitions. What these approaches have in common, due to their origin, is that they concentrate on the qualitative part of inconsistency rather than the quantitative part, i. e. the probabilities.

The simplest approach to define an inconsistency measure is by just differentiating whether a knowledge base is consistent or inconsistent.

Definition 7. Let $\mathcal{I}_{\text{drastic}} : \mathbb{K} \to [0, \infty)$ be the function defined as

$$\mathcal{I}_{\text{drastic}}(\kappa) = \begin{cases} 0 & \text{if } \kappa \text{ is consistent} \\ 1 & \text{if } \kappa \text{ is inconsistent} \end{cases}$$

for $\kappa \in \mathbb{K}$. The function $\mathcal{I}_{\text{drastic}}$ is called the *drastic inconsistency measure*.

The drastic inconsistency measure allows only for a binary decision on inconsistencies and does not quantify the severity of inconsistencies. One thing to note is that $\mathcal{I}_{\text{drastic}}$ is the upper bound for any inconsistency measure that satisfies *consistency* and *normalization*, i. e., if $\mathcal{I}$ satisfies *consistency* and *normalization* then $\mathcal{I}(\kappa) \leq \mathcal{I}_(\kappa)$ for every $\kappa \in \mathbb{K}$ [44].

The next inconsistency measure quantifies inconsistency by the number of minimal inconsistent subsets of a knowledge base.

Definition 8. Let $\mathcal{I}_{\text{MI}} : \mathbb{K} \to [0, \infty)$ be the function defined as

$$\mathcal{I}_{\text{MI}}(\kappa) = |\mathsf{MI}(\kappa)|$$

for $\kappa \in \mathbb{K}$. The function $\mathcal{I}_{\text{MI}}$ is called the MI *inconsistency measure*.

The definition of the MI inconsistency measure is motivated by the intuition that the more minimal inconsistent subsets the greater the inconsistency.

Only considering the number of minimal inconsistent subsets may be too simple for assessing inconsistencies in general. Another indicator for the severity of inconsistencies is the size of minimal inconsistent subsets. A large minimal inconsistent subset means that the inconsistency is distributed over a large number of conditionals. The more conditionals involved in an inconsistency the less severe the inconsistency is. Furthermore, a small minimal inconsistent subset means that the participating conditionals strongly represent contradictory information. The following inconsistency measure is from [20] and aims at differentiating between minimal inconsistent sets of different size.

Definition 9. Let $\mathcal{I}^C_{\mathsf{MI}} : \mathbb{K} \to [0, \infty)$ be the function defined as

$$\mathcal{I}^C_{\mathsf{MI}}(\kappa) = \sum_{\mathcal{M} \in \mathsf{MI}(\kappa)} \frac{1}{|\mathcal{M}|}$$

for $\kappa \in \mathbb{K}$. The function $\mathcal{I}^C_{\mathsf{MI}}$ is called the MI *inconsistency measure.*

Note that $\mathcal{I}^C_{\mathsf{MI}}(\kappa) = 0$ if $\mathsf{MI}(\kappa) = \varnothing$. The MI^Cinconsistency measure sums over the reciprocal of the sizes of all minimal inconsistent subsets. In that way, a large minimal inconsistent subset contributes less to the inconsistency value than a small minimal inconsistent subset.

The work [23] employs probability theory itself to measure inconsistency in classical theories by considering probability measures on classical interpretations. Those ideas can be extended for measuring inconsistency in probabilistic logics by considering probabilistic interpretations on probabilistic interpretations. Let $\hat{\pi} : \mathcal{P}(\mathcal{A}) \to [0, 1]$ be a probabilistic interpretation on $\mathcal{P}(\mathcal{A})$ such that $\hat{\pi}(\pi) > 0$ only for finitely many $\pi \in \mathcal{P}(\mathcal{A})$. Let $\mathcal{P}^2(\mathcal{A})$ be the set of those probabilistic interpretations. Then define the probability measure $P_{\hat{\pi}}$ analogously via

$$P_{\hat{\pi}}(c) = \sum_{\pi \in \mathcal{P}(\mathcal{A}), \pi \models c} \hat{\pi}(\pi) \tag{1}$$

for a conditional c. This means that the probability (in terms of $\hat{\pi}$) of a conditional is the sum of the probabilities of probabilistic interpretations that satisfy c. Note also that by restricting $\hat{\pi}$ to assign a non-zero value only to finitely many $\pi \in \mathcal{P}(\mathcal{A})$, the sum in (1) is well-defined.

Now consider the following definition of the η-inconsistency measure.

Definition 10. Let $\mathcal{I}_\eta : \mathbb{K} \to [0, \infty)$ be the function defined as

$$\mathcal{I}_\eta(\kappa) = 1 - \max\{\eta \mid \exists \hat{\pi} \in \mathcal{P}^2(\mathcal{A}) : \forall c \in \kappa : \hat{\pi}(c) \geq \eta\}$$

for $\kappa \in \mathbb{K}$. The function $\mathcal{I}_\eta$ is called the *η-inconsistency measure.*

The idea of the η-inconsistency measure is that it looks for the largest probability that can be consistently assigned to the conditionals of a knowledge base and defines the inconsistency value inversely proportional to this probability.

Example 6. We continue Example 5 and consider

$$\kappa = \{(s_1 \wedge s_2)[0.6, 1], (s_1 \wedge \neg s_2)[0.5, 1], (s_1)[0, 0.8]\}$$

Recall that $\mathsf{MI}(\kappa) = \{\{(s_1 \wedge s_2)[0.6, 1], (s_1 \wedge \neg s_2)[0.5, 1]\}\}$ and therefore

$$\begin{aligned}\mathcal{I}_{\text{drastic}}(\kappa) &= 1\\ \mathcal{I}_{\text{MI}}(\kappa) &= 1\\ \mathcal{I}^{C}_{\mathsf{MI}}(\kappa)&1/2\end{aligned}$$

Finally, consider $\hat{\pi} \in \mathcal{P}^2(\mathcal{A})$ such that $\hat{\pi}(\pi_1) = \hat{\pi}(\pi_2) = 1/2$ for

$$\begin{array}{llll}\pi_1(s_1 \wedge s_2) = 0.8 & \pi_1(\neg s_1 \wedge s_2) = 0.2 & \pi_1(s_1 \wedge \neg s_2) = 0 & \pi_1(\neg s_1 \wedge \neg s_2) = 0\\ \pi_2(s_1 \wedge s_2) = 0 & \pi_2(\neg s_1 \wedge s_2) = 0 & \pi_2(s_1 \wedge \neg s_2) = 0.8 & \pi_2(\neg s_1 \wedge \neg s_2) = 0\end{array}$$

Observe that

$$\begin{array}{ll}\pi_1 \models (s_1 \wedge s_2)[0.6, 1] & \pi_1 \models (s_1)[0, 0.8]\\ \pi_2 \models (s_1 \wedge \neg s_2)[0.5, 1] & \pi_2 \models (s_1)[0, 0.8]\end{array}$$

and therefore

$$\begin{aligned}\hat{\pi}((s_1 \wedge s_2)[0.6, 1]) &= 0.5\\ \hat{\pi}((s_1 \wedge \neg s_2)[0.5, 1]) &= 0.5\\ \hat{\pi}((s_1)[0, 0.8]) &= 1\end{aligned}$$

and thus $\hat{\pi}(c) \geq 0.5$ for all $c \in \kappa$. It can be easily seen that there cannot be some $\hat{\pi}'$ with $\hat{\pi}(c) > 0.5$ for all $c \in \kappa$ and therefore $\mathcal{I}_\eta(\kappa) = 0.5$.

The inconsistency measures discussed above were initially developed for inconsistency measurement in classical theories and therefore allow only for a "discrete" measurement. Hence, all of the above discussed inconsistency measures do not satisfy *continuity*. In particular, we have the following results [45].

Proposition 4. $\mathcal{I}_{\text{drastic}}$ satisfies *consistency*, *monotonicity*, *irrelevance of syntax*, *weak independence*, *independence*, and *normalisation*.

Proposition 5. $\mathcal{I}_{\mathrm{MI}}$ satisfies *consistency*, *monotonicity*, *super-additivity*, *irrelevance of syntax*, *weak independence*, *independence*, *MI-separability*, and *penalty*.

Proposition 6. $\mathcal{I}_{\mathsf{MI}}^{C}$ satisfies *consistency*, *monotonicity*, *super-additivity*, *irrelevance of syntax*, *weak independence*, *independence*, *MI-separability*, and *penalty*.

Proposition 7. $\mathcal{I}_{\eta}$ satisfies *consistency*, *monotonicity*, *irrelevance of syntax*, *weak independence*, *independence*, and *normalisation*.

However, satisfaction of *continuity* is crucial for an inconsistency measure in probabilistic logics in order to assess inconsistencies in a meaningful manner [45]. In the following, we continue with a survey of inconsistency measures that take the probabilities of conditionals into account and therefore address the postulate *continuity*.

4.2 Distance-based approaches

The measures presented in [45] rely on distance measures defined as follows.

Definition 11 (Distance measure)**.** A function $d : \mathbb{R}^n \times \mathbb{R}^n \to [0, \infty)$ (with $n \in \mathbb{N}^+$) is called a *distance measure* if it satisfies the following properties (for all $\vec{x}, \vec{y}, \vec{z} \in \mathbb{R}^n$):

1. $d(\vec{x}, \vec{y}) = 0$ if and only if $\vec{x} = \vec{y}$ (*reflexivity*)
2. $d(\vec{x}, \vec{y}) = d(\vec{y}, \vec{x})$ (*symmetry*)
3. $d(\vec{x}, \vec{y}) \leq d(\vec{x}, \vec{z}) + d(\vec{z}, \vec{y})$ (*triangle inequality*)

The simplest form of a distance measure is the *drastic distance measure* d_0 defined as $d_0(\vec{x}, \vec{y}) = 0$ for $\vec{x} = \vec{y}$ and $d_0(\vec{x}, \vec{y}) = 1$ for $\vec{x} \neq \vec{y}$ (for $\vec{x}, \vec{y} \in \mathbb{R}^n$ and $n \in \mathbb{N}^+$). A more interesting distance measure is the p-norm distance.

Definition 12. Let $n, p \in \mathbb{N}^+$. The function $d_p : \mathbb{R}^n \times \mathbb{R}^n \to [0, \infty)$ defined via

$$d_p(\vec{x}, \vec{y}) = \sqrt[p]{|x_1 - y_1|^p + \ldots + |x_n - y_n|^p}$$

for $\vec{x} = (x_1, \ldots, x_n), \vec{y} = (y_1, \ldots, y_n) \in \mathbb{R}^n$ is called the *p-norm distance*.

Special cases of the p-norm distance include the *Manhattan distance* (for $p = 1$) and the *Euclidean distance* (for $p = 2$).

Now we can define the "severity of inconsistency" in a knowledge base by the minimal distance of the knowledge base to a consistent one. As we are able to identify knowledge bases of the same qualitative structure in a vector space, we can employ distance measures for measuring inconsistency.

Definition 13. Let d be a distance measure. Then the function $\mathcal{I}_d : \mathbb{K} \to [0, \infty)$ defined via

$$\mathcal{I}_d(\kappa) = \inf\{d(\vec{x}, \vec{y}) \mid \kappa = \kappa[\vec{x}] \text{ and } \kappa[\vec{y}] \text{ is consistent}\} \tag{2}$$

for $\kappa \in \mathbb{K}$ is called the *d-inconsistency measure*.

The idea behind the d-inconsistency measure is that we look for a consistent knowledge base that both 1.) has the same qualitative structure as the input knowledge base and 2.) is as close as possible to the the input knowledge base. That is, if the input knowledge base is $\kappa[\vec{x}]$ we look at all $\vec{y} \in [0,1]^{2|\kappa|}$ such that $\kappa[\vec{y}]$ is consistent and $\vec{x}$ and $\vec{y}$ are as close as possible with respect to the distance measure d. While using the drastic distance measure gives us drastic inconsistency measure—that is $\mathcal{I}_{d_0} = \mathcal{I}_{\text{drastic}}$ [45]—using p-norms gives us genuinely novel inconsistency measures.

Example 7. We continue Example 5 and consider

$$\kappa = \{(s_1 \wedge s_2)[0.6, 1], (s_1 \wedge \neg s_2)[0.5, 1], (s_1)[0, 0.8]\}$$

Then we have

$$\begin{aligned} \mathcal{I}_{d_1}(\kappa) &= 0.3 \\ \mathcal{I}_{d_2}(\kappa) &\approx 0.55 \end{aligned}$$

In order to see $\mathcal{I}_{d_1}(\kappa) = 0.3$ observe that changing the interval of $(s_1 \wedge s_2)[0.6, 1]$ to $[0.5, 1]$ and of $(s_1)[0, 0.8]$ to $[0, 1]$ restores consistency. More precisely, the probabilistic interpretation π defined via

$$\begin{aligned} \pi(s_1 \wedge s_2) &= 0.5 \\ \pi(s_1 \wedge \neg s_2) &= 0.5 \\ \pi(\neg s_1 \wedge s_2) &= \pi(\neg s_1 \wedge \neg s_2) = 0 \end{aligned}$$

satisfies $(s_1 \wedge s_2)[0.5, 1]$, $(s_1 \wedge \neg s_2)[0.5, 1]$, and $(s_1)[0, 1]$. Note that we changed the probability interval of the first conditional by a value of 0.1 and that of the last interval by 0.2, yielding a sum of 0.3. It can be seen that no smaller modification yields a consistent knowledge base, so we have $\mathcal{I}_{d_1}(\kappa) = 0.3$.

In order to see $\mathcal{I}_{d_2}(\kappa) \approx 0.55$ consider the knowledge base κ' given via

$$\kappa' = \{(s_1 \wedge s_2)[0.5, 1], (s_1 \wedge \neg s_2)[0.4, 1], (s_1)[0, 0.9]\}$$

Observe κ' can be obtained from κ' by changing each probability interval by a value of 0.1, thus the distance between the probability intervals wrt. the Euclidean distance d_2 is $\sqrt{0.1^2 + 0.1^2 + 0.1^2} \approx 0.55$. It can also easily be seen that κ' is consistent and no other knowledge base with smaller Euclidean distance is consistent.

Taking results from [45, 9] into account we obtain the following picture regarding compliance with rationality postulates[2]. For reasons of simplicity,

[2]Note that [9] corrects some false claims from [45].

we only consider the instantiation of $\mathcal{I}_d$ with the p-norm distance d_p, see [45] for a more general treatment.

Proposition 8. Let $p \geq 1$. $\mathcal{I}_{d_p}$ satisfies *consistency*, *monotonicity*, *super-additivity* (only for $p = 1$), *irrelevance of syntax*, *weak independence*, and *continuity*.

For a variant of the above measure that also works with infinitesimal probabilities we refer to [26].

4.3 Violation-based approaches

The next family of inconsistency measures is based on the idea of measuring the violation of the numerical constraints that define the satisfaction relation for conditionals. For this purpose, it is convenient to represent the conditional satisfaction constraints in matrix notation [28, 21]. We assume an arbitrary but fixed order on $\mathcal{W}_\mathcal{A}$ (e. g., lexicographically) so that we can identify each probabilistic interpretation π with the vector $(\pi(w_1), \ldots, \pi(w_n))'$, $n = |\mathcal{W}_\mathcal{A}|$, where the worlds are enumerated according to our order.

Recall that $\pi \models (\varphi|\psi)[\underline{q}, \bar{q}]$ iff $P_\pi(\varphi \wedge \psi) \geq \underline{q} P_\pi(\psi)$ and $P_\pi(\varphi \wedge \psi) \leq \bar{q} P_\pi(\psi)$. Subtracting the right-hand-side of the inequalities and putting in the definition of P_π, we notice that these are linear inequalities over π. For instance, $P_\pi(\varphi \wedge \psi) \geq \underline{q} P_\pi(\psi)$ is true if and only if

$$
\begin{aligned}
0 &\leq P_\pi(\varphi \wedge \psi) - \underline{q} P_\pi(\psi) = P_\pi(\varphi \wedge \psi) - \underline{q}\big(P_\pi(\varphi \wedge \psi) + P_\pi(\neg\varphi \wedge \psi)\big) \\
&= (1 - \underline{q}) P_\pi(\varphi \wedge \psi) - \underline{q} P_\pi(\neg\varphi \wedge \psi) \\
&= (1 - \underline{q}) \sum_{w_j \models \varphi \wedge \psi} \pi(w_j) - \underline{q} \sum_{w_j \models \neg\varphi \wedge \psi} \pi(w_j) \\
&= \sum_{j=1}^{n} \big((1 - \underline{q}) \cdot 1_{\{\varphi \wedge \psi\}}(w_j) - \underline{q} \cdot 1_{\{\neg\varphi \wedge \psi\}}(w_j)\big) \cdot \pi(w_j) \\
&= a\,\pi,
\end{aligned}
$$

where the indicator function $1_{\{F\}} : \mathcal{W}_\mathcal{A} \to \{0, 1\}$ yields 1 if the argument satisfies the formula $F \in \mathcal{L}_\mathcal{A}$ and 0 otherwise, and a is a n-dimensional row vector whose i-th component is $a_j = (1 - \underline{q}) \cdot 1_{\{\varphi \wedge \psi\}}(w_j) - \underline{q} \cdot 1_{\{\neg\varphi \wedge \psi\}}(w_j)$. Note that $a\,\pi \geq 0$ is equivalent to $b\,\pi \leq 0$, where $b = -a$. We can write all constraints for our knowledge base in this form. Thus, given some knowledge base κ, we can arrange the vectors corresponding to the constraints in a matrix A_κ such that an interpretation π satisfies κ if and only if $A_\kappa\,\pi \leq 0$. We call A_κ the *constraint matrix of* κ.

If κ is inconsistent, the system of inequalities $A\,x \leq 0$ cannot be satisfied by a probability vector. However, we can relax the constraint to $A\,\pi \leq \epsilon$ for some non-negative vector ϵ. The minimum size of a vector that makes the system satisfiable is the *minimum violation value* of the knowledge base and can be understood as an inconsistency measure [30, 9].

Formally, the minimal violation value of a knowledge base is defined by an optimization problem. The problem is parameterized by a vector norm that measures the size of ϵ.

Definition 14 (Minimal Violation Value with respect to $\|.\|$)**.** Let κ be a knowledge base with corresponding constraint matrix A_κ (of size $m \times n$). Let $\|.\|$ be some continuous vector norm. The *minimal violation value of* κ *with respect to* $\|.\|$ is defined by

$$\begin{aligned} \min_{(x,\epsilon)\in\mathbb{R}^{n+m}} \quad & \|\epsilon\| \\ \textit{subject to} \quad & A_\kappa\, x \leq \epsilon, \\ & \sum_{i=1}^{n} x_i = 1, \\ & x \geq 0, \\ & \epsilon \geq 0. \end{aligned} \tag{3}$$

We denote the minimal violation value of κ by $\mathcal{I}_{\|.\|}(\kappa)$.

Proposition 9. $\mathcal{I}_{\|.\|}$ is an inconsistency measure and satisfies Consistency, Monotonicity, Weak Independence, Irrelevance of syntax and Continuity.

$\mathcal{I}_{\|.\|}$ satisfies neither Independence nor MI-Separability.

Proofs for these results can be found in [30, 9]. Measuring inconsistency by measuring the error in the system of inequalities is conceptually less intuitive than measuring the distance to a consistent knowledge base. However, it has some computational advantages as we discuss soon and still measures the degree of inconsistency continuously.

Example 8. Let $\kappa_\rho = \{(a)[\rho], (a)[1-\rho]\}$ for $0 \leq \rho \leq 0.5$. For instance, $\kappa_{0.5} = \{(A)[0.5], (A)[0.5]\}$ is consistent and $\kappa_0 = \{(A)[0], (A)[1]\}$ is inconsistent. Intuitively, the degree of inconsistency of κ_ρ should increase as ρ decreases from 0.5 to 0. We let $\mathcal{I}_1$, $\mathcal{I}_2$ and $\mathcal{I}_\infty$ denote the minimal violation measure with respect to the Manhattan norm $\|x\|_1 = \sum_{i=1}^n |x_i|$, Euclidean norm $\|.\|_2 = \sqrt{\sum_{i=1}^n x_i^2}$ and Maximum norm $\|x\|_\infty = \max\{|x_1|, \ldots, |x_n|\}$, respectively. Table 1 shows the corresponding inconsistency values.

Inspecting (3) shows that computing minimal violation values is a convex optimization problem (the objective function is convex and all constraints are linear). This has the principal advantage that we do not have to deal with non-global local optima and have polynomial runtime guarantees with respect to

	$\kappa_{0.5}$	$\kappa_{0.49}$	$\kappa_{0.4}$	$\kappa_{0.2}$	κ_{0}
$\mathcal{I}_1$	0	0.02	0.2	0.6	1
$\mathcal{I}_2$	0	0.014	0.141	0.424	0.707
$\mathcal{I}_\infty$	0	0.01	0.1	0.3	0.5

Table 1: Some minimal violation inconsistency values (Example 8).

the number of optimization variables and constraints. However, the number of optimization variables is still exponential in the number of atomic propositions (recall that each optimization variable corresponds to the probability of an interpretation).

We can do slightly better when using the Euclidean norm, which gives us a quadratic optimization problem. When using the Manhattan or the Maximum norm, the minimal violation value can actually be computed by a linear optimization problem. For the Manhattan norm $\|x\|_1 = \sum_{i=1}^{n} |x_i|$, we can immediately rewrite (3) as a linear program because the vector ϵ is constrained to be non-negative. Therefore, the absolute value function can be ignored.

Proposition 10. The minimal violation value $\mathcal{I}_{\|.\|_1}(\kappa)$ with respect to the Manhattan norm can be computed by solving the following linear program:

$$\begin{aligned} \min_{(x,\epsilon)\in\mathbb{R}^{n+m}} \quad & \sum_{i=1}^{m} \epsilon_i \\ \textit{subject to} \quad & A_\kappa\, x \leq \epsilon, \\ & \sum_{i=1}^{n} x_i = 1, \\ & x \geq 0, \\ & \epsilon \geq 0. \end{aligned} \tag{4}$$

For the Maximum norm $\|x\|_\infty = \max\{|x_1|, \ldots, |x_n|\}$, we can replace the relaxing vector ϵ with a scalar because we are only interested in the maximum component. Again, we can ignore the absolute value function due to the non-negativity of ϵ.

Proposition 11. The minimal violation value $\mathcal{I}_{\|.\|_\infty}(\kappa)$ with respect to the

Maximum norm can be computed by solving the following linear program:

$$\begin{aligned}
\min_{(x,\epsilon)\in\mathbb{R}^{n+1}} \quad & \epsilon \\
\textit{subject to} \quad & A_\kappa\, x \leq \epsilon \cdot \vec{1}, \\
& \sum_{i=1}^{n} x_i = 1, \\
& x \geq 0, \\
& \epsilon \geq 0,
\end{aligned} \tag{5}$$

where $\vec{1} \in \mathbb{R}^m$ denotes the column vector that contains only ones.

Let us relate these linear programs to the probabilistic satisfiability problem [14, 21], that is, to the problem of deciding whether a given knowledge base is consistent. This problem comes down to finding a solution $x \in \mathbb{R}^n$ of the following system of linear inequalities:

$$\begin{aligned}
& A_\kappa\, x \leq 0, \\
& \sum_{i=1}^{n} x_i = 1, \\
& x \geq 0.
\end{aligned}$$

A standard way to solve this system is to apply Phase 1 of the Simplex algorithm. This comes down to solving a linear program like

$$\begin{aligned}
\min_{(x,s)\in\mathbb{R}^{n+1}} \quad & s \\
\textit{subject to} \quad & A_\kappa\, x \leq 0, \\
& \sum_{i=1}^{n} x_i + s = 1, \\
& x \geq 0, \\
& \epsilon \geq 0,
\end{aligned} \tag{6}$$

Note that the vector $(0,1) \in \mathbb{R}^{n+1}$ is always a feasible vertex from which we can start the Simplex algorithm for this problem. κ is consistent iff the optimal solution $s^* = 0$. Notice that the structure of (6) is very similar to the structure of (5). We have the same number of optimization variables and constraints and there is only a minor difference in the first and second constraint. (4) does also have a very similar structure even though it adds a number of optimization variables that is linear in the size of the knowledge base. However, since n is exponential in the number of atoms of our language, this difference is usually negligible. In this sense, computing minimal violation measures for Manhattan and Maximum norm is barely harder than performing a probabilistic satisfiability test. Since inconsistency measures generalize probabilistic satisfiability tests, this is some evidence that these measures belong to the most efficient inconsistency measures. Note also that even though the Simplex algorithm has exponential runtime in the worst-case, it usually takes time linear in the number of optimization variables and quadratic in the number of constraints in practice [25]. In fact, due to the similarity between (6), (5) and (4), we can speed up computing minimal violation values for the Manhattan and Maximum norm by using similar techniques like for the probabilistic satisfiability problem. In particular, column generation techniques proved useful in this context [14, 18, 11, 7].

4.4 Dutch-book measures

In formal epistemology, there is an interest in measuring the incoherence of an agent whose beliefs are given as probabilities over propositions or previsions (expected values) for random variables — a *Bayesian* agent. If we have propositions from classical logic, the formalized problem at hand is exactly the one we are investigating. When the agent's degrees of belief are represented by a knowledge base, to measure the agent's incoherence is to measure the inconsistency of such a knowledge base. Schervish, Kadane and Seidenfeld [40, 41, 39] have proposed ways to measure the incoherence of an agent based on Dutch books.

Dutch books have been proposed by De Finetti as a foundation for probability theory [10]. Dutch book arguments are based on the agent's betting behavior induced by her degrees of belief, typically used to show her irrationality. These arguments rely on an operational interpretation of (imprecise) degrees of belief, in which their lower/upper bounds determines when the agent considers as fair to take part in a gamble, defined as follows:

Definition 15. A *gamble* on $\varphi|\psi$, with $\varphi, \psi \in \mathcal{L}_{\mathcal{A}}$, is an agreement between an agent and a bettor with two parameters, the *stake* $\lambda \in \mathbb{R}$ and the *relative price* $q \in [0, 1]$, stating that:

- the agent pays $\lambda \times q$ to the bettor if ψ is true and φ is false;
- the bettor pays $\lambda \times (1 - q)$ to the agent if ψ is true and φ is true;
- the gamble is called off, causing neither profit nor loss to the involved parts, if ψ is false.

A *Dutch book* is a set of gambles that will cause the agent a sure loss, no matter which possible world is the case. For instance, suppose an agent is willing to take part in two gambles, on φ and on $\neg\varphi$, both with stake 10 and relative price 0.6. No matter whether φ or $\neg\varphi$ is the case, the agent has to pay $10 \times 0.6 = 6$ to the bettor, while receiving only $10 \times 0.4 = 4$ back, which causes her a net loss of 2.

A central result of De Finetti's theory of probabilities lies in the fact that if an agent respects the laws of probabilities, no Dutch books are possible. That is, a Dutch book (sure loss) is possible only when the agent gambles are inconsistent with the laws or probability.

To relate the epistemic state of an agent to her vulnerability to Dutch books, we need a *willingness-to-gamble assumption*: if an agent believes that the probability of a proposition φ being true given that another proposition ψ is true lies within $[\underline{q}, \bar{q}]$, she finds acceptable (fair) gambles on $\varphi|\psi$ with stake $\lambda \geq 0$ and relative price $\underline{q}$ and gambles with stake $\lambda \leq 0$ and relative price $\bar{q}$. An agent *is vulnerable to (or exposed to)* a given Dutch book if she sees as fair, under the willingness-to-gamble assumption, each of the gambles in the Dutch book. We assume any set of gambles the agent sees as fair contains exactly two gambles on $(\varphi_i|\psi_i)$ per each conditional $(\varphi_i|\psi_i)[\underline{q}_i, \bar{q}_i] \in \kappa$, the base formalizing the agent's beliefs: one with stake $\underline{\lambda}_i \geq 0$ and relative price $\underline{q}_i$; and the other with stake $-\bar{\lambda}_i \leq 0$ and relative price $\bar{q}_i$. This is not restrictive, since gambles on the same $(\varphi_i|\psi_i)$ with the same relative price can be merged by summing the stakes, and the absence of a gamble is equivalent to a stake equal to zero. We can thus denote any set of gambles the agent finds acceptable simply by the absolute value of its stakes $\underline{\lambda}_1, \bar{\lambda}_1, \ldots, \underline{\lambda}_m, \bar{\lambda}_m \geq 0$, where $m = |\kappa|$.

If the set of probabilistic conditionals that represents an agent's epistemic state turns out to be inconsistent, then she is exposed to a Dutch book, and vice-versa [27]. In other words, an agent sees as fair a set of gambles that causes her a guaranteed loss if, and only if, the knowledge base codifying her (conditional) degrees of belief is inconsistent. In this way, Dutch book arguments were introduced to show that a set of degrees of belief must obey the axioms of probability and are a standard proof of incoherence (introductions to Dutch books and their relation to incoherence can be found in [42] and [10]). Hence, a natural approach to measuring an agent's degree of incoherence is through the magnitude of the sure loss she is vulnerable to. The intuition says that, the more incoherent an agent is, the greater the guaranteed loss that can be imposed on her through a Dutch book. Nevertheless, with no bounds on the stakes, such loss would also be unlimited for incoherent agents. To better understand the loss a Dutch book causes to an agent, we formalize it in the following.

Consider the knowledge base $\kappa = \{(\varphi_i|\psi_i)[\underline{q}_i, \bar{q}_i] | 1 \leq i \leq m\}$ representing an agent's epistemic state. Let $\underline{\lambda}_i, \bar{\lambda}_i \geq 0$ denote gambles on $(\varphi_i|\psi_i)$ the agent sees as acceptable, the first with relative price $\underline{q}_i$ and stake $\underline{\lambda}_i \geq 0$, the second with relative price $\bar{q}_i$ and stake $-\bar{\lambda}_i \leq 0$, for $1 \leq i \leq m$. A set of gambles the agent sees as fair can then be represented by the vector $\mathcal{G} = \langle \underline{\lambda}_1, \bar{\lambda}_1, \ldots, \underline{\lambda}_m, \bar{\lambda}_m \rangle$. If a possible world w_j is the case, the total net loss for the agent is

$$\ell_\kappa(\mathcal{G}, w_j) = \sum_{i=1}^{m} \bar{\lambda}_i(I_{\varphi_i \wedge \psi_i}(w_j) - \bar{q}_i I_{\psi_i}(w_j)) - \underline{\lambda}_i(I_{\varphi_i \wedge \psi_i}(w_j) - \underline{q}_i I_{\psi_i}(w_j)) \,.$$

Given the knowledge base κ representing an agent's epistemic state, the set of gambles $\mathcal{G}$ is a Dutch book if $\ell_\kappa(\mathcal{G}, w_j) < 0$ for all $w_j \in \mathcal{W}_\mathcal{A}$. When $\mathcal{G}$ is a Dutch book, the *sure loss* is defined as the amount the agent is guaranteed to lose, which is $\ell_\kappa^{sure}(\mathcal{G}) = \min_{w_j \in \mathcal{W}_\mathcal{A}} \ell_\kappa(\mathcal{G}, w_j)$. If $\mathcal{G}$ is a not a Dutch book, there is a possible world where the agent does not lose (maybe wins), then $\ell_\kappa^{sure}(\mathcal{G}) = 0$. For an arbitrary set of gambles $\mathcal{G}$ that the agent sees as fair, we thus define $\ell_\kappa^{sure}(\mathcal{G}) = \max\{\min_{w_j \in \mathcal{W}_\mathcal{A}} \ell_\kappa(\mathcal{G}, w_j), 0\}$. The maximum Dutch-book sure loss an agent is vulnerable to is given by $\max_\mathcal{G} \ell_\kappa^{sure}(\mathcal{G})$. To see that this maximization in unbounded, note that if $\ell_\kappa^{sure}(\mathcal{G}) = c > 0$ for some $\mathcal{G} = \langle \underline{\lambda}_1, \bar{\lambda}_1, \ldots, \underline{\lambda}_m, \bar{\lambda}_m \rangle$, then $\mathcal{G}' = \lambda \langle \underline{\lambda}_1, \bar{\lambda}_1, \ldots, \underline{\lambda}_m, \bar{\lambda}_m \rangle$ implies $\ell_{sure}(\mathcal{G}) = \lambda c > 0$ for any positive scalar $\lambda \in \mathbb{R}$. Consequently, any incoherent agent is vulnerable to an unbounded sure loss, and this quantity is not suitable to measure her incoherence.

Different strategies to measure incoherence as a finite Dutch-book loss are found in the Bayesian Statistics and formal epistemology literature. Schervish *et al.* propose a flexible formal approach to normalize the maximum sure loss generating a family of incoherence measures for upper and lower previsions on bounded random variables [41], which we adapt to our case. The simplest measures of this family arise when the sure loss is normalized by either the sum of the absolute values of the stakes, $\|\mathcal{G}\|_1 = \sum_i \underline{\lambda}_i + \bar{\lambda}_i \leq 1$, or their maximum, $\|\mathcal{G}\|_\infty = \max\{\underline{\lambda}_i, \bar{\lambda}_i \mid 1 \leq i \leq m\}$. We define the inconsistency measures $\mathcal{I}_{DB}^{sum} : \mathbb{K} \to [0, \infty)$ and $\mathcal{I}_{DB}^{max} : \mathbb{K} \to [0, \infty)$ on knowledge bases as these two incoherence measures on the corresponding agents represented by these knowledge bases [3]:

$$\mathcal{I}_{DB}^{sum}(\kappa) = \max_\mathcal{G} \frac{\ell_\kappa^{sure}(\mathcal{G})}{\|\mathcal{G}\|_1} \quad \text{and} \quad \mathcal{I}_{DB}^{max}(\kappa) = \max_\mathcal{G} \frac{\ell_\kappa^{sure}(\mathcal{G})}{\|\mathcal{G}\|_\infty}$$

Even though incoherence measures based, on one hand, on Dutch books proposed by the formal epistemology community and inconsistency measures based, on the other hand, on violations minimization proposed by Artificial Intelligence researchers may seem unrelated at first, they are actually two sides of the same coin. The linear programs that compute the maximum guaranteed loss an agent is exposed to are technically dual to those that minimize violations to measure inconsistency [9] [4]:

Theorem 1. For any $\kappa \in \mathbb{K}$, $\mathcal{I}_{DB}^{sum}(\kappa) = \mathcal{I}_{\|.\|_\infty}(\kappa)$ and $\mathcal{I}_{DB}^{max}(\kappa) = \mathcal{I}_{\|.\|_1}(\kappa)$.

[3] If $\|\mathcal{G}\|_1 = \|\mathcal{G}\|_\infty = 0$, then $\ell_\kappa^{sure}(\mathcal{G}) = 0$ and we define $\ell_\kappa^{sure}(\mathcal{G})/\|\mathcal{G}\|_1 = \ell_\kappa^{sure}(\mathcal{G})/\|\mathcal{G}\|_\infty = 0$.

[4] Nau [27] has already investigated this matter, mentioning some similar results.

Theorem 1 gives an operational interpretation, based on betting behavior, for the inconsistency measures $\mathcal{I}_{\|.\|_\infty}$ and $\mathcal{I}_{\|.\|_1}$. Naturally, the result also implies that $\mathcal{I}_{DB}^{sum}$ and $\mathcal{I}_{DB}^{max}$ hold the same properties as $\mathcal{I}_{\|.\|_\infty}$ and $\mathcal{I}_{\|.\|_1}$, respectively.

More elaborate measures proposed by Schervish *et al.* [41] normalize the sure loss by the amount the agent (or bettor) can possibly lose, either per gamble or in total. These quantities are called *escrows*. Let $\underline{q} = \langle \underline{q}_1, \dots, \underline{q}_m \rangle$, $\bar{q} = \langle \bar{q}_1, \dots, \bar{q}_m \rangle$ and $1_m = \langle 1 \dots 1 \rangle$ and be tuples with m elements. In a single gamble with stake $\underline{\lambda}_i > 0$ and relative price $\underline{q}_i$, the agent's (or the bettor's) escrow is $e^a_{\underline{i}} = \underline{\lambda}_i \times \underline{q}_i$ (or $e^b_{\underline{i}} = \underline{\lambda}_i \times (1 - \underline{q}_i)$). Conversely, in a gamble with stake $-\bar{\lambda}_i < 0$ and relative price $\bar{q}_i$, the agent's (or the bettor's) escrow is $e^a_{\bar{i}} = \bar{\lambda}_i \times (1 - \bar{q}_i)$ (or $e^b_{\bar{i}} = \bar{\lambda}_i \times \bar{q}_i$). Hence, if $\mathcal{G} = \langle \underline{\lambda}_1, \bar{\lambda}_1, \dots, \underline{\lambda}_m, \bar{\lambda}_m \rangle$ is a set of gambles, $e^a(\mathcal{G}) = \langle e^a_{\underline{1}}, e^a_{\bar{1}}, \dots, e^a_{\underline{m}}, e^a_{\bar{m}} \rangle$ (or $e^b(\mathcal{G}) = \langle e^b_{\underline{1}}, e^b_{\bar{1}}, \dots, e^b_{\underline{m}}, e^b_{\bar{m}} \rangle$) is the vector containing how much the agent (or the bettor) can lose per each gamble. Normalizing the sure loss by the maximum or the sum of this vector's elements yields inconsistency measures defined as:

$$\mathcal{I}_{DB}^{a,sum}(\kappa) = \max_{\mathcal{G}} \frac{\ell_\kappa^{sure}(\mathcal{G})}{\|e^a(\mathcal{G})\|_1} \quad ; \quad \mathcal{I}_{DB}^{a,max}(\kappa) = \max_{\mathcal{G}} \frac{\ell_\kappa^{sure}(\mathcal{G})}{\|e^a(\mathcal{G})\|_\infty}$$

$$\mathcal{I}_{DB}^{b,sum}(\kappa) = \max_{\mathcal{G}} \frac{\ell_\kappa^{sure}(\mathcal{G})}{\|e^b(\mathcal{G})\|_1} \quad ; \quad \mathcal{I}_{DB}^{b,max}(\kappa) = \max_{\mathcal{G}} \frac{\ell_\kappa^{sure}(\mathcal{G})}{\|e^b(\mathcal{G})\|_\infty}$$

These four inconsistency measures ($\mathcal{I}_{DB}^{a,sum}$, $\mathcal{I}_{DB}^{b,sum}$, $\mathcal{I}_{DB}^{a,max}$ and $\mathcal{I}_{DB}^{b,max}$) satisfy most of the rationality postulates [9]:

Proposition 12. $\mathcal{I}_{DB}^{a,sum}$, $\mathcal{I}_{DB}^{b,sum}$, $\mathcal{I}_{DB}^{a,max}$ and $\mathcal{I}_{DB}^{b,max}$ are well-defined, satisfy Consistency, Monotonicity and Weak Indepedence and are continuous for probabilities within $(0, 1)$. Furthermore, $\mathcal{I}_{DB}^{a,max}$ and $\mathcal{I}_{DB}^{b,max}$ satisfy Super-additivity, and $\mathcal{I}_{DB}^{a,sum}$ satisfies Normalization.

From their definition, one can see that $\mathcal{I}_{DB}^{a,sum}$, $\mathcal{I}_{DB}^{b,sum}$, $\mathcal{I}_{DB}^{a,max}$ and $\mathcal{I}_{DB}^{b,max}$ also satisfy Irrelevance of syntax, since the net loss of a gamble on $\varphi|\psi$ does not depend on how these formulas are written.

4.5 Fuzzy-logic measure

In [8], another inconsistency measure on probabilistic knowledge bases is proposed that makes use of Fuzzy concepts. The central notion of [8] is the *candidacy function*. A candidacy function is similar to a fuzzy set [15] as it assigns a degree of membership of a probabilistic interpretation belonging to the models of a knowledge base. More formally, a candidacy function $\mathfrak{C}$ is a function $\mathfrak{C} : \mathcal{P}(\mathcal{A}) \to [0, 1]$. A uniquely determined candidacy function $\mathfrak{C}_\kappa$ can be assigned to a (consistent or inconsistent) knowledge base κ as follows. For a probabilistic interpretation $\pi \in \mathcal{P}(\mathcal{A})$ and a set of probabilistic interpretations $S \subseteq \mathcal{P}(\mathcal{A})$ let $d(\pi, S)$ denote the distance of π to S with respect to the Euclidean norm, i. e., $d(\pi, S)$ is defined via

$$d(\pi, S) = \inf \left\{ \sqrt{\sum_{\omega \in \mathcal{W}_{\mathcal{A}}} (\pi(\omega) - \pi'(\omega))^2} \mid \pi' \in S \right\} \quad .$$

Let $h : \mathbb{R}^+ \to (0, 1]$ be a strictly decreasing, positive, and continuous log-concave function with $h(0) = 1$. Then the candidacy function $\mathfrak{C}^h_\kappa$ for a knowledge base κ is defined as

$$\mathfrak{C}^h_\kappa(\pi) = \prod_{c \in \kappa} h\left(\sqrt{2^{|\mathcal{A}|}} d(\pi, \mathsf{Mod}(\{c\}))\right)$$

for every $\pi \in \mathcal{P}(\mathcal{A})$. Note that the definition of the candidacy function $\mathfrak{C}^h_\kappa$ depends on the size of the signature $\mathcal{A}$. The intuition behind this definition is that a probabilistic interpretation π that is near to the models of each probabilistic conditional in κ gets a high candidacy degree wrt. $\mathfrak{C}^h_\kappa$. It is easy to see that it holds that $\mathfrak{C}^h_\kappa(\pi) = 1$ if and only if $\pi \in \mathsf{Mod}(()\kappa)$. Using the candidacy function $\mathfrak{C}^h_\kappa$ the inconsistency measure $\mathcal{I}^h_{\text{cand}}$ can be defined via

$$\mathcal{I}^h_{\text{cand}}(\kappa) = 1 - \max_{\pi \in \mathcal{P}(\mathcal{A})} \mathfrak{C}^h_\kappa(\pi) \tag{7}$$

for a knowledge base κ. The following results has been shown in [8].

Proposition 13. $\mathcal{I}^h_{\text{cand}}$ satisfies *consistency*, *monotonicity*, *continuity*, and *normalization*.

The function $\mathcal{I}^h_{\text{cand}}$ does not satisfy *super-additivity* as shown in [44].

Example 9. Let $\mathcal{A} = \{a_1, a_2\}$ be a propositional signature and let $\kappa_1 = \{(a_1)[1], (a_1)[0]\}$ and $\kappa_2 = \{(a_2)[1], (a_2)[0]\}$ be knowledge bases and let $\kappa = \kappa_1 \cup \kappa_2$. Note that both κ_1 and κ_2 are inconsistent and $\kappa_1 \cap \kappa_2 = \varnothing$. As $\mathcal{I}^h_{\text{cand}}$ is defined on the semantic level and does not take the names of propositions into account it follows that $\mathcal{I}^h_{\text{cand}}(\kappa_1) = \mathcal{I}^h_{\text{cand}}(\kappa_2)$. As the situations in κ_1 and κ_2 are symmetric and κ_i is symmetric with respect to $(a_i)[1]$ and $(a_i)[0]$ there are probabilistic interpretations π_i with $\mathcal{I}^h_{\text{cand}}(\kappa_i) = 1 - \mathfrak{C}^h_{\kappa_i}(\pi_i)$ for $i = 1, 2$ and

$$\begin{aligned} d(\pi_1, \mathsf{Mod}(\{(a_1)[1]\})) &= d(\pi_1, \mathsf{Mod}(\{(a_1)[0]\})) \\ &= d(\pi_2, \mathsf{Mod}(\{(a_2)[1]\})) \\ &= d(\pi_2, \mathsf{Mod}(\{(a_2)[0]\})) \end{aligned}$$

Let $x = d(\pi_1, \mathsf{Mod}(\{(a_1)[1]\}))$ and let $h^* : \mathbb{R}^+ \to (0, 1]$ be a strictly decreasing, positive, and continuous log-concave function with $h^*(0) = 1$ and

$h^*\left(\sqrt{2^{|\mathcal{A}|}}x\right) = 0.5$. Then it follows $\mathfrak{C}^{h^*}_{\kappa_1}(\pi_1) = 0.25$ and $\mathcal{I}^h_{\text{cand}}(\kappa_1) = 0.75$. In order to satisfy *super-additivity* $\mathcal{I}^h_{\text{cand}}$ must satisfy

$$\mathcal{I}^h_{\text{cand}}(\kappa) \geq \mathcal{I}^h_{\text{cand}}(\kappa_1) + \mathcal{I}^h_{\text{cand}}(\kappa_2) = 1.5$$

which is a contradiction since $\mathcal{I}^h_{\text{cand}}$ satisfies *normalization*.

On the other hand, $\mathcal{I}^h_{\text{cand}}$ complies with our notion of *irrelevance of syntax*[5]

Proposition 14. $\mathcal{I}^h_{\text{cand}}$ satisfies *irrelevance of syntax*.

Proof. Let κ_1 and κ_2 be such that $\kappa_1 \equiv^s \kappa_2$. Without loss of generality, assume $\kappa_1 = \{c_1, \ldots, c_n\}$ and $\kappa_2 = \{d_1, \ldots, d_n\}$ with $c_i \equiv d_i$ for $i = 1, \ldots, n$. It follows $\mathsf{Mod}(\{c_i\}) = \mathsf{Mod}(\{d_i\})$ for $i = 1, \ldots, n$ and therefore

$$d(\pi, \mathsf{Mod}(\{c_i\}) = d(\pi, \mathsf{Mod}(\{d_i\})$$

for every π and $i = 1, \ldots, n$. It follows $\mathfrak{C}^h_{\kappa_1} = \mathfrak{C}^h_{\kappa_2}$ and therefore the claim. □

4.6 Entropy measures

In [38] an inconsistency measure is presented that is based on the notion of *generalized divergence* which generalizes cross-entropy. Given vectors $\vec{y}, \vec{z} \in (0,1]^n$ with $\vec{y} = (y_1, \ldots, y_n)$ and $\vec{z} = (z_1, \ldots, z_n)$, the generalized divergence $D(\vec{y}, \vec{z})$ from $\vec{y}$ to $\vec{z}$ is defined

$$D(\vec{y}, \vec{z}) = \sum_{i=1}^{n} y_i \log_2 \frac{y_i}{z_i} - y_i + z_i$$

We abbreviate further

$$D^2(\vec{y}, \vec{z}) = D(\vec{y}, \vec{z}) + D(\vec{z}, \vec{y}) = \sum_{i=1}^{n} y_i \log_2 \frac{y_i}{z_i} + z_i \log_2 \frac{z_i}{y_i}$$

In [38], the measure $\mathcal{I}_{\text{gd}}$ is only defined for conditionals with point probabilities. So let $\kappa = \{c_1, \ldots, c_n\}$ and $c_i = (\psi_i | \varphi_i)[d_i]$ for $i = 1, \ldots, n$. Then the inconsistency measure $\mathcal{I}_{\text{gd}}$ is defined via

$$\begin{aligned}\mathcal{I}_{\text{gd}}(\kappa) = \min\{&D^2(\vec{y}, \vec{z}) \mid \pi \in \mathcal{P}(\mathcal{A}) \text{ and}\\ &y_i = (1 - d_i) P_\pi(\psi_i \varphi_i) \text{ and } z_i = d_i P_\pi(\neg\psi_i \wedge \varphi_i) \text{ for } i = 1, \ldots, n\}\end{aligned}$$

[5] In [44] it has been shown, however, that $\mathcal{I}^h_{\text{cand}}$violates a slightly different notion of irrelevance of syntax.

Let $\vec{y}^*, \vec{z}^*, P_{\pi^*}$ be some parameters such that $D^2(\vec{y}^*, \vec{z}^*)$ is minimal and $y_i^* = (1-d_i)P_{\pi^*}(\psi_i \varphi_i)$ and $z_i^* = d_i P_{\pi^*}(\neg\psi_i \wedge \varphi_i)$ are satisfied for $i = 1, \ldots, n$. Then it follows that

$$y_i^* - z_i^* = P_{\pi^*}(\psi_i \wedge \varphi_i) - d_i P_{\pi^*}(\varphi_i)$$

for $i = 1, \ldots, n$. Minimizing $D^2(\vec{y}, \vec{z})$ amounts to finding a probabilistic interpretation π^* such that $\vec{y}^*$ and $\vec{z}^*$ are as close as possible to each other with respect to D^2. In particular, if there is a π^* such that $y^* = z^*$ it follows that $P_{\pi^*}(\psi_i \wedge \varphi_i) - d_i P_{\pi^*}(\varphi_i) = 0$ and therefore $\pi \in \mathsf{Mod}((\psi_i|\varphi_i)[d_i])$ (for $i = 1, \ldots, n$), i.e, κ is consistent. Furthermore, the more y_i^* differs from z_i^* the more $P_{\pi^*}(\psi_i|\varphi_i)$ differs from d_i (for $i = 1, \ldots, n$). The measure $\mathcal{I}_{\text{gd}}$ is similar in spirit to $\mathcal{I}_d$ as they both minimize the distance of a knowledge base to a consistent one. However, the implementation of those measures is different as they use different distance measures. The following result has been shown in [44].

Proposition 15. The function $\mathcal{I}_{\text{gd}}$ satisfies *consistency*, *monotonicity*, *super-additivity*, *weak independence*, and *continuity*.

Another approach to inconsistency measuring based on entropy can be derived from works on consistency repairing in de Finetti's coherence setting [3, 4, 5]. In order to correct incoherent conditional probabilities, a discrepancy based on Kullback-Leibler divergence is minimized, which can be understood as an inconsistency measure. In the following, we adapt the approach of [3, 4, 5] to our semantics, keeping their focus on sets of conditionals with point probabilities.

Consider the knowledgebase $\kappa = \{c_1, \ldots, c_n\}$, with $c_i = (\psi_i|\varphi_i)[q_i]$ and $d_i \in (0, 1)$ [6] for $i = 1, \ldots, n$. The approach of [5] is based on the following scoring rule, which can be seen as evaluating the accuracy of a set of probabilistic conditionals in a possible world w:

$$S_{CRV}(\kappa, w) = \sum_{i=1}^{n} 1_{\{\varphi_i \wedge \psi_i\}}(w) \ln q_i + \sum_{i=1}^{m} 1_{\{\neg\varphi_i \wedge \psi_i\}}(w) \ln(1 - q_i)\,. \qquad (8)$$

Any probabilistic interpretation $\pi \in \mathcal{P}(\mathcal{A})$, which is a probability mass over the worlds $w \in W_{\mathcal{A}}$, defines an expected value $E_\pi(S_{CRV}(\kappa, w))$. For a fixed π, the point probabilities $\vec{q} = \langle q_1, \ldots, q_n \rangle$ that maximizes $E_\pi(S_{CRV}(\kappa[\vec{q}], w))$ are given by the vector $\vec{q}^{\,\pi} = \langle \frac{P_\pi(\varphi_1 \wedge \psi_1)}{P_\pi(\psi_1)}, \ldots, \frac{P_\pi(\varphi_n \wedge \psi_n)}{P_\pi(\psi_n)} \rangle \in [0, 1]^n$. The following discrepancy between a knowledge base κ and a probabilistic interpretation π gives the expected gap in accuracy, when measured by S_{CRV}, between the suboptimal $\kappa = \kappa[\vec{q}]$ and the maximally accurate $\kappa[\vec{q}^{\,\pi}]$:

[6] In [5], extreme probabilities (0 or 1) are avoided for technical reasons.

$$
\begin{aligned}
d_{CRV}(\kappa, \pi) &= E_{\pi}\Big(S_{CRV}(\kappa[\vec{q}^{\pi}], w) - S_{CRV}(\kappa, w)\Big) \\
&= \sum_{1 \le i \le n, P_{\pi}(\psi_i) > 0} P_{\pi}(\psi_i)\Big(q_i^{\pi} \ln \frac{q_i^{\pi}}{q_i} + (1 - q_i^{\pi}) \ln \frac{1 - q_i^{\pi}}{1 - q_i}\Big).
\end{aligned}
$$

This discrepancy directly yields an inconsistency measure for precise probabilistic knowledge bases. In [5], π must be such that $\pi(\bigvee_i \psi_i) = 1$, but we drop that restriction due to the semantics we are adopting.

$$
\mathcal{I}_{CRV}(\kappa) = \min\{d_{CRV}(\kappa, \pi) | \pi \in \mathcal{P}(\mathcal{A})\}.
$$

Proposition 16. $\mathcal{I}_{CRV}$ satisfies *c*onsistency, *m*onotonicity and *c*ontinuity.

5 Applications

The inconsistency measures introduced in the previous section give us a tool to analyze inconsistent knowledge bases. Our final goal is to reason over these knowledge bases in a sensible way. There are at least two ideas that we can consider for this purpose.

1. Repair the inconsistent knowledge base and apply classical probabilistic reasoning algorithms.

2. Apply paraconsistent reasoning algorithms that can deal with inconsistent knowledge bases.

The distance-based approaches in Section 4.2 are particularly well suited for repairing knowledge bases. In fact, when computing the inconsistency value of the knowledge base, we usually do so by finding a consistent knowledge base that minimizes the selected distance to the original knowledge base. However, there are some obstacles. First of all, if we consider only point probabilities, whether or not a unique closest consistent knowledge base exists depends on the selected norm. For instance, uniqueness is guaranteed for the Euclidean norm, but not for the Manhattan and Maximum norm. As a simple example, consider the knowledge base $\{(a)[0.2], a[0.6]\}$. With respect to the Manhattan norm, each repair $\{(a)[p] : 2\}$ with $p \in [0.2, 0.6]$ is minimal and the choice would be arbitrary without further assumptions.

Second, even if a unique solution exists, repairing the knowledge base means loss of information. To make this clear, consider the knowledge bases $\{(a)[0.4], a[0.6]\}$ and $\{(a)[0.1], a[0.9]\}$. Both knowledge bases have the unique minimal repair $\{(a)[0.5] : 2\}$ with respect to the Euclidean norm. The fact that the second knowledge base has a significantly higher variance is lost. If we think of the knowledge bases as representing the opinions of two different experts, 0.5 is close to both experts' opinion in the first knowledge base, but not in the second.

In cases like this, where we have to shift a huge amount of probability mass to repair the knowledge base, applying paraconsistent reasoning mechanisms can be a better choice. We can derive such reasoning mechanisms from the fuzzy- and violation-based approaches. The idea is to replace the models of a knowledge base with those probabilistic interpretations that are close enough to being a model. That is, if the knowledge base is consistent, we use the usual models to perform reasoning. If it is inconsistent, we use the probabilistic interpretations that are closest to being a model. For the fuzzy-based measures from Section 4.5, this means that we use those interpretations that maximize the candidacy value with respect to $\mathfrak{C}^h_\kappa$ [8]. For the violation-based measures from Section 4.3, we use those interpretations that minimize the violation of the knowledge base [30, 35]. We explain this approach in somewhat more detail for minimum violation measures.

As a first step, we define the *generalized models* of a knowledge base κ as the set of probabilistic interpretations that minimize the violation value (3) in Definition 14. More strictly speaking, we let

$$\mathsf{GMod}(\kappa) = \{\pi \in \mathcal{P}(|)(\pi, \epsilon) \textit{ minimizes } (3) \textit{ for some } \epsilon \in \mathbb{R}\}.$$

Intuitively, $\mathsf{GMod}(\kappa)$ contains those probability distributions that violate the knowledge base minimally. In particular, if κ is consistent, we have $\mathsf{Mod}(\kappa) = \mathsf{GMod}(\kappa)$. However, often we have some special conditionals that should not be violated at all. We call these conditionals *integrity constraints*. We assume that the integrity constraints are consistent. Now given a knowledge base κ and a set of integrity constraints IC, we define the corresponding generalized models as the set of probabilistic interpretations that satisfy IC and minimally violate κ.

Definition 16. Let κ, IC be knowledge bases such that IC is consistent. Let $\|.\|$ be some continuous vector norm. The set of *generalized models of* κ *with respect to* IC *and* $\|.\|$ is defined by

$$\mathsf{GMod}^{\|.\|}_{IC}(\kappa) = \{\pi \in \mathsf{Mod}(IC) \mid (\pi, \epsilon) \textit{ minimizes } (3) \textit{ for some } \epsilon \in \mathbb{R}\}.$$

If IC and $\|.\|$ are clear from the context or not important for the discussion, we will just write $\mathsf{GMod}(\kappa)$ to keep our notation simple. $\mathsf{GMod}(\kappa)$ is guaranteed to be non-empty and has some nice technical properties that allow us to reason as efficiently with generalized models as with classical models in many cases.

There are two major approaches to perform reasoning over consistent probabilistic knowledge bases. In both cases, our final goal is to answer conditional probabilistic queries, that is, to compute the conditional probability of a formula φ given another formula ψ. We denote such queries by $(\varphi \mid \psi)$. The first

Query	$\|.\|_1$	$\|.\|_2$	$\|.\|_\infty$
$(P \mid N)$	$[0.1, 0.9]$	$[0.376, 0.624]$	$[0.366, 0.633]$
$(P \mid Q)$	$[0.1, 0.9]$	$[0.533, 0.679]$	$[0.536, 0.689]$
$(P \mid R)$	$[0.1, 0.9]$	$[0.321, 0.467]$	$[0.314, 0.463]$
(N)	$[1, 1]$	$[0.801, 0.801]$	$[0.789, 0.789]$

Table 2: Generalized entailment results (rounded to 3 digits) for Nixon diamond with $IC = \varnothing$ (Example 10).

approach is to compute upper and lower bounds on the conditional probability of φ given ψ with respect to all models of κ [28, 17]. This approach is often referred to as the *probabilistic entailment problem*. The second approach is a two-stage process. We first select a best model that satisfies the knowledge base and then use this model to compute the conditional probability of φ given ψ [29, 22]. The 'best' model is determined by an evaluation function. For instance, we may be interested in maximizing entropy or minimizing some notion of distance to a prior distribution. We refer to this approach as the *model selection problem*. Both approaches can be easily generalized to inconsistent knowledge bases by just replacing the probabilistic interpretations that satisfy the knowledge base (the classical models) with those that minimally violate the knowledge base (the generalized models) [34, 35]. A detailed description and discussion of both approaches can be found in [36]. The following example illustrates how our generalization of the probabilistic entailment problem to inconsistent knowledge bases can be applied.

Example 10. Let us consider the *Nixon diamond*. We believe that quakers (Q) are usually pacifists (P) while republicans (R) are usually not. However, we know that Nixon (N) was both a quaker and a republican. Let us model our beliefs with the following knowledge base:

$$\kappa = \{(P \mid Q)[0.9], (P \mid R)[0.1], (N)[1], (Q \wedge R \mid N)[1]\}.$$

κ is inconsistent. For instance, its minimal violation value with respect to the Euclidean norm is $\mathcal{I}_{\|.\|_2}(\kappa) \approx 0.42$. Let us set $IC = \varnothing$ and ask for the probability that Nixon was a pacifist. Table 2 shows the result and some additional queries that show in which way the knowledge in κ has been relaxed.

We can in particular see that the Manhattan norm yields the most conservative results in the sense that it provides very large answer intervals. However, the answer is still bounded away from the trivial bounds 0 and 1. For the Euclidean and the Maximum norm, we maintain the knowledge that quakers are probably pacifists and that republicans are probably not (the probabilities are bounded away from 0.5). We also notice in Table 2 that the probability

Query	$\|\cdot\|_1$	$\|\cdot\|_2$	$\|\cdot\|_\infty$
$(P \mid N)$	$[0.1, 0.9]$	$[0.384, 0.615]$	$[0.376, 0.624]$
$(P \mid Q)$	$[0.1, 0.9]$	$[0.517, 0.615]$	$[0.520, 0.624]$
$(P \mid R)$	$[0.1, 0.9]$	$[0.384, 0.482]$	$[0.376, 0.481]$
(N)	$[1, 1]$	$[1, 1]$	$[1, 1]$

Table 3: Generalized entailment results (rounded to 3 digits) for Nixon diamond with $IC = \{(N)[1]\}$ (Example 10).

that the person under consideration is Nixon (N) has also been subject to change. However, since we have no doubts about Nixon's existence, we let $(N)[1]$ become an integrity constraint. That is, we now let $IC = \{(N)[1]\}$ and

$$\kappa = \{(P \mid Q)[0.9], (P \mid R)[0.1], (QR \mid N)[1]\}$$

Table 3 shows the new generalized entailment results.

The generalizations of both the probabilistic entailment problem and the model selection problem satisfy some interesting properties. Intuitively, these properties can be described as follows:

Consistency If $\kappa \cup IC$ is consistent, the generalized reasoning results coincide with the classical reasoning results.

Independence If some subset of $\kappa \cup IC$ is consistent, then generalized reasoning results that depend only on this subset coincide with the classical reasoning results.

Continuity If $\kappa \cup IC$ is topologically close to a consistent knowledge base, then the generalized reasoning results will be close to the classical reasoning results.

A thorough discussion of these properties and their exact preconditions can be found in [32, 36].

In several applications, we want to override general rules by more specific rules. This can be modeled by a knowledge base that is partitioned into subsets with different priorities. If we assume that each subset of the partition is consistent with given integrity constraints, we can consider another form of generalized probabilistic reasoning. Similar as before, we start with the models of the integrity constraints $\mathcal{M}_0$. We then select from $\mathcal{M}_0$ those models that minimally violate the conditionals with highest priority yielding a subset $\mathcal{M}_1$ of $\mathcal{M}_0$. We continue in this way, constructing $\mathcal{M}_{i+1}$ by selecting from $\mathcal{M}_i$ those models that minimally violate the conditionals with the next highest priority. A detailed description of this approach and its properties can be found in [31].

The following example illustrates how this approach can be used to generalize the probabilistic entailment problem to knowledge bases with priorities.

Example 11. We consider a probabilistic version of an access control policy scenario from [2]. Suppose we have different files and different users and want to automatically deduce the probability that a user has access to a file. If the probability is 1, we might grant access immediately, otherwise we might send a confirmation request to the system administrator. If the probability is very low, say smaller than 0.1, we might want to send a warning in addition.

We model this problem using a relational probabilistic language similar to [24, 12]. We build up formulas over a finite set of typed predicate symbols, a finite set of typed individuals and an infinite set of (typed) variables. We allow the usual logical connectives, but do not allow quantifiers.

We use the types *User* and *File* and the predicates *grantAccess*(*User*, *File*), *employee*(*User*), *exec*(*User*), *blacklisted*(*User*), *confidential*(*File*), where *exec* abbreviates *executive manager*. Let *alice* and *bob* be individuals of type *User* and let *file1*, *file2* be individuals of type *File*.

Our priority knowledge base has the form $\kappa = (\kappa_1, \kappa_2, \kappa_3, \kappa_4, \kappa_5, IC)$, where a higher index means higher priority. That is, κ_1 has the lowest and κ_5 has the highest priority (disregarding the integrity constraints IC that cannot be violated at all). The subsets of the knowledge base are defined as follows:

$$\begin{aligned}
\kappa_1 &= \{(\mathit{grantAccess}(U,F))[0], (\mathit{blacklisted}(U))[0.05]\} \\
\kappa_2 &= \{(\mathit{grantAccess}(U,F) \mid \mathit{employee}(U))[0.5], \\
&\quad\ (\mathit{blacklisted}(U) \mid \mathit{employee}(U))[0.01]\} \\
\kappa_3 &= \{(\mathit{grantAccess}(U,F) \mid \mathit{confidential}(F))[0]\} \\
\kappa_4 &= \{(\mathit{grantAccess}(U,F) \mid \mathit{exec}(U))[0.7], \\
&\quad\ (\mathit{blacklisted}(U) \mid \mathit{exec}(U))[0.001]\} \\
\kappa_5 &= \{(\mathit{exec}(\mathit{alice}))[1], (\mathit{employee}(\mathit{bob}))[1], (\mathit{confidential}(\mathit{file1}))[1]\} \\
IC &= \{(\mathit{employee}(U) \mid \mathit{exec}(U))[1], (\mathit{grantAccess}(U,F) \mid \mathit{blacklisted}(U)(F))[0]\}
\end{aligned}$$

On the first level, we define generic knowledge. If no knowledge is available, we do not want to grant access to anybody. Also, we make the assumption that it is rather unlikely that a user is blacklisted. On the second level, we increase the access probability and decrease the blacklist probability for employees. On level 3, we make an exception for confidential files. Afterwards, we further increase access probability and decrease blacklist probability for executive managers on level 4. The last level contains domain knowledge. We know that *alice* is an executive manager, *bob* is an employee and *file1* is confidential. Our integrity constraints state that executive managers are employees and that we do not grant access to blacklisted users.

We have the following rounded reasoning results when using the Euclidean norm to determine our strict priority models:

$$\begin{array}{lr} \mathit{grantAccess}(\mathit{alice}, \mathit{file1})[0.7] & \mathit{grantAccess}(\mathit{bob}, \mathit{file1})[0] \\ \mathit{grantAccess}(\mathit{alice}, \mathit{file2})[0.7] & \mathit{grantAccess}(\mathit{bob}, \mathit{file2})[0.5] \\ \mathit{blacklisted}(\mathit{alice})[0.0001] & \mathit{blacklisted}(\mathit{bob})[0.01]. \end{array}$$

The results make intuitively sense. For instance, the first query shows that for the executive manager *alice*, the access rule $(\mathit{grantAccess}(U, F) \mid \mathit{exec}(U))[0.7]$ with priority 4 has been applied, while $(\mathit{grantAccess}(U, F) \mid \mathit{confidential}(F))[0]$ with priority 3 and $(\mathit{grantAccess}(U, F))[0]$ with priority 1 have been ignored. Similarly, we can see that for the employee *bob* the rule $(\mathit{grantAccess}(U, F) \mid \mathit{confidential}(F))[0]$ with priority 3 applies because *file1* is confidential and *bob* is not an executive manager.

Generalized reasoning approaches can also be applied in multi-agent systems. For instance, in [33], multi-agent decision problems have been investigated where each agent has individual beliefs and utilities. Generalized Probabilistic Entailment can be used to derive group beliefs from the individual beliefs. Then expected utilities for the group can be computed from these group beliefs. Since this approach yields utility intervals rather than point utilities, one can define different preference relations. These approaches satisfy independence and continuity properties similar to the ones that we discussed after Example 10 and also satisfy some desirable social choice properties [33].

6 Summary

In this chapter, we gave an overview of approaches to measuring inconsistency in probabilistic logics. The most important property that distinguishes measures for probabilistic logics from measures for classical logics is *Continuity*. Continuity guarantees that minor changes of probabilities cannot result in major changes in the inconsistency value. This property seems highly desirable for analyzing inconsistencies in probabilistic logics because conflicts can be resolved by carefully adjusting probabilities in the knowledge base. However, as explained in Section 3, Continuity is actually in conflict with *Independence* and *MI-Separability* that have been considered for classical measures. As argued in Section 3, our position is that these properties should be given up for probabilistic knowledge bases in favor of *Continuity*.

In Section 4, we discussed different approaches for measuring inconsistency in probabilistic logics. The first class of measures was directly *adapted from inconsistency measures for classical logics*. While these measures are able to measure inconsistencies qualitatively, they do not take probabilities into account. *Distance-based measures* attempt to minimize the distance in probabilities from the original knowledge base to a consistent repair. They measure inconsistency continuously and usually yield a repair as a byproduct. However, they can be difficult to compute due to their non-convex nature. *Violation-based measures* attempt to find a better tradeoff between computational and analytic properties. To do so, they do not minimize the distance in probabilities directly, but try to minimize the error in the numerical constraints that correspond to the knowledge base. While this approach is less intuitive than minimizing the distance directly, it still measures inconsistency continuously and can be solved by convex programming techniques in general, and even via linear programming for two specific measures. These two measures are equivalent to some *measures based on Dutch books*, from the Bayesian philosophy/statistics community, which were then presented. Afterwards, we discussed a *measure based on fuzzy logic* that relies on assigning degrees of membership of probabilistic interpretation belonging to models of a knowledge base. Finally, we discussed *measures that rely on the notion of entropy*.

In Section 5, we sketched some applications of inconsistency measures for probabilistic logics in repairing and reasoning with inconsistent knowledge bases. Distance-based measures are well suited for repairing knowledge bases. Adapting the probabilities in the knowledge base in a minimal way seems to be the most intuitive way to repair inconsistent probabilistic knowledge bases. However, by replacing the inconsistent knowledge base with a repair, we may lose information about the variance in the information. So instead, we may want to infer probabilities directly from the inconsistent knowledge base. Violation-based measures are well suited for this purpose. By replacing the models of a knowledge base with those probability distributions that minimally violate the knowledge base, we can transfer reasoning approaches for consistent knowledge bases to inconsistent ones. As we discussed, these generalizations guarantee that classical reasoning results on the consistent part of the knowledge base remain unaffected (Independence) and that reasoning results over knowledge bases that are close to consistent knowledge bases are not too far from the classical reasoning results (Continuity).

References

[1] C.E. Alchourrón, P. Gärdenfors, and D. Makinson. On the logic of theory change: Partial meet contraction and revision functions. *The Journal of Symbolic Logic*, 50(02):510–530, 1985.

[2] Piero A Bonatti, Marco Faella, and Luigi Sauro. Adding default attributes to el++. In *AAAI*, 2011.

[3] A. Capotorti and G. Regoli. Coherent correction of inconsistent conditional probability assessments. In *Proceedings of IPMU'08*, pages 891–898, 2008.

[4] A. Capotorti, G. Regoli, and F. Vattari. On the use of a new discrepancy measure to correct incoherent assessments and to aggregate conflicting opinions based on imprecise conditional probabilities. In *Proceedings of the Sixth International Symposium on Imprecise Probability: Theories and Applications, ISIPTA*, 2009.

[5] A. Capotorti, G. Regoli, and F. Vattari. Correction of incoherent conditional probability assessments. *International Journal of Approximate Reasoning*, 51(6):718–727, 2010.

[6] G. Coletti and R. Scozzafava. *Probabilistic Logic in a Coherent Setting.* Trends in Logic, Vol. 15: Studia Logica Library. Kluwer Academic Publishers, 2002.

[7] Fabio Gagliardi Cozman and Lucas Fargoni Ianni. Probabilistic satisfiability and coherence checking through integer programming. In L. C. van der Gaag, editor, *Proc. Symbolic and Quantitative Approaches to Reasoning with Uncertainty, ECSQARU 2013*, volume 7958 of *Lecture Notes in Computer Science*, pages 145–156. Springer Berlin Heidelberg, 2013.

[8] Lionel Daniel. *Paraconsistent Probabilistic Reasoning.* PhD thesis, L'École Nationale Supérieure des Mines de Paris, December 2009.

[9] Glauber De Bona and Marcelo Finger. Measuring inconsistency in probabilistic logic: Rationality postulates and dutch book interpretation. *Artificial Intelligence*, 227:140–164, 2015.

[10] B. de Finetti. Theory of Probability, 1974.

[11] Marcelo Finger and Glauber De Bona. Probabilistic satisfiability: Logic-based algorithms and phase transition. In *IJCAI*, pages 528–533, 2011.

[12] Jens Fisseler. First-order probabilistic conditional logic and maximum entropy. *Logic Journal of IGPL*, 20(5):796–830, 2012.

[13] Alan M. Frisch and Peter Haddawy. Anytime Deduction for Probabilistic Logic. *Artificial Intelligence*, 69(1–2):93–122, 1994.

[14] George F. Georgakopoulos, Dimitris J. Kavvadias, and Christos H. Papadimitriou. Probabilistic satisfiability. *Journal of Complexity*, 4(1):1–11, 1988.

[15] Giangiacomo Gerla. *Fuzzy Logic: Mathematical Tools for Approximate Reasoning.* Trends in Logic. Springer-Verlag, 2001.

[16] Joseph Y. Halpern. An Analysis of First-Order Logics of Probability. *Artificial Intelligence*, 46:311–350, 1990.

[17] Pierre Hansen and Brigitte Jaumard. Probabilistic satisfiability. In Jrg Kohlas and Serafn Moral, editors, *Handbook of Defeasible Reasoning and Uncertainty Management Systems*, volume 5 of *Handbook of Defeasible Reasoning and Uncertainty Management Systems*, pages 321–367. Springer Netherlands, 2000.

[18] Pierre Hansen and Sylvain Perron. Merging the local and global approaches to probabilistic satisfiability. *International Journal of Approximate Reasoning*, 47(2):125 – 140, 2008.

[19] A. Hunter and S. Konieczny. Shapley inconsistency values. In *10th International Conference on Principles of Knowledge Representation and Reasoning (KR)*, pages 249–259, 2006.

[20] Anthony Hunter and Sébastien Konieczny. Measuring inconsistency through minimal inconsistent sets. In Gerhard Brewka and Jérôme Lang, editors, *Proceedings of the Eleventh International Conference on Principles of Knowledge Representation and Reasoning (KR'2008)*, pages 358–366, Sydney, Australia, September 2008. AAAI Press, Menlo Park, California.

[21] B. Jaumard, P. Hansen, and Marcus Poggi. Column generation methods for probabilistic logic. *ORSA - Journal on Computing*, 3(2):135–148, 1991.

[22] G. Kern-Isberner. *Conditionals in Nonmonotonic Reasoning and Belief Revision*. Springer, 2001.

[23] K. M. Knight. Measuring inconsistency. *Journal of Philosophical Logic*, 31:77–98, 2001.

[24] Thomas Lukasiewicz. Probabilistic deduction with conditional constraints over basic events. *J. Artif. Int. Res.*, 10(1):199–241, April 1999.

[25] J. Matousek and B. Gärtner. *Understanding and Using Linear Programming*. Universitext (1979). Springer, 2007.

[26] David Picado Muiño. Measuring and Repairing Inconsistency in Probabilistic Knowledge Bases. *International Journal of Approximate Reasoning*, 52(6):828–840, September 2011.

[27] R.F. Nau. Coherent assessment of subjective probability. Technical report, DTIC Document, 1981.

[28] Nils J. Nilsson. Probabilistic logic. *Artificial Intelligence*, 28:71–88, February 1986.

[29] J. B. Paris. *The Uncertain Reasoner's Companion – A Mathematical Perspective*. Cambridge University Press, 1994.

[30] Nico Potyka. Linear programs for measuring inconsistency in probabilistic logics. In *Proceedings of the 14th International Conference on Principles of Knowledge Representation and Reasoning (KR'14)*, 2014.

[31] Nico Potyka. Reasoning over linear probabilistic knowledge bases with priorities. In *The 9th International Conference on Scalable Uncertainty Management (SUM'15)*, 2015.

[32] Nico Potyka. *Solving Reasoning Problems for Probabilistic Conditional Logics with Consistent and Inconsistent Information.* PhD thesis, FernUniversität in Hagen, Hagen, 2015.

[33] Nico Potyka, Erman Acar, Matthias Thimm, and Heiner Stuckenschmidt. Group decision making via probabilistic belief merging. In *Proceedings of the 25th International Joint Conference on Artificial Intelligence (IJCAI'16)*, July 2016.

[34] Nico Potyka and Matthias Thimm. Consolidation of probabilistic knowledge bases by inconsistency minimization. In *Proceedings of the 21st European Conference on Artificial Intelligence (ECAI'14)*, pages 729–734, August 2014.

[35] Nico Potyka and Matthias Thimm. Probabilistic reasoning with inconsistent beliefs using inconsistency measures. In *Proceedings of the 24th International Joint Conference on Artificial Intelligence (IJCAI'15)*, July 2015.

[36] Nico Potyka and Matthias Thimm. Inconsistency-tolerant reasoning over linear probabilistic knowledge bases. *International Journal of Approximate Reasoning*, 2017 (to appear).

[37] R. Reiter. A theory of diagnosis from first principles. *Artificial intelligence*, 32(1):57–95, 1987.

[38] Wilhelm Rödder and Longgui Xu. Elimination of Inconsistent Knowledge in the Probabilistic Expert-System-Shell SPIRIT (in German). In B. Fleischmann, R. Lasch, K. Derigs, W. Domschke, and K. Riedler, editors, *Operations Research Proceedings: Selected Papers of the Symposium on Operations Research 2000*, pages 260–265. Springer-Verlag, 2001.

[39] M.J. Schervish, J.B. Kadane, and T. Seidenfeld. Measures of incoherence: How not to gamble if you must. In *Bayesian Statistics 7: Proceedings of the 7th Valencia Conference on Bayesian Statistics*, pages 385–402, 2003.

[40] M.J. Schervish, T. Seidenfeld, and J.B. Kadane. Two measures of incoherence: How not to gamble if you must. Technical report, Department of Statistics, Carnegie Mellon University, 1998.

[41] M.J. Schervish, T. Seidenfeld, and J.B. Kadane. Measuring incoherence. *Sankhyā: The Indian Journal of Statistics, Series A*, pages 561–587, 2002.

[42] A. Shimony. Coherence and the axioms of confirmation. *The Journal of Symbolic Logic*, 20(01):1–28, 1955.

[43] Matthias Thimm. Measuring Inconsistency in Probabilistic Knowledge Bases. In Jeff Bilmes and Andrew Ng, editors, *Proceedings of the Twenty-Fifth Conference on Uncertainty in Artificial Intelligence (UAI'09)*, pages 530–537. AUAI Press, June 2009.

[44] Matthias Thimm. *Probabilistic Reasoning with Incomplete and Inconsistent Beliefs*. Number 331 in Dissertations in Artificial Intelligence. IOS Press, December 2011.

[45] Matthias Thimm. Inconsistency measures for probabilistic logics. *Artificial Intelligence*, 197:1–24, April 2013.