International Journal of Approximate Reasoning 52 (2011) 828–840

Contents lists available at ScienceDirect

International Journal of Approximate Reasoning

journal homepage: www.elsevier.com/locate/ijar

Measuring and repairing inconsistency in probabilistic knowledge bases ☆

David Picado Muiño *

Institut für Diskrete Mathematik und Geometrie, Wiedner Hauptstrasse 8/10, 1040 Vienna, Austria

ARTICLE INFO

Article history:
Received 4 October 2010
Revised 18 December 2010
Accepted 18 February 2011
Available online 5 March 2011

Keywords:
Probabilistic knowledge bases
Probabilistic satisfiability
Inconsistency
Measures of inconsistency
CADIAG-2

ABSTRACT

In this paper we present a family of measures aimed at determining the *amount* of inconsistency in probabilistic knowledge bases. Our approach to measuring inconsistency is *graded* in the sense that we consider *minimal* adjustments in the degrees of certainty (i.e., probabilities in this paper) of the statements necessary to make the knowledge base consistent. The computation of the family of measures we present here, in as much as it yields an adjustment in the probability of each statement that restores consistency, provides the modeler with possible repairs of the knowledge base. The case example that motivates our work and on which we test our approach is the knowledge base of CADIAG-2, a well-known medical expert system.

1. Introduction

In the last few years the amount of literature dealing with aspects of inconsistency in knowledge bases has grown considerably and has become central in the field of databases and knowledge-based systems. In this paper we focus primarily on two aspects of inconsistency in probabilistic knowledge bases: the *evaluation* of inconsistency and possible *repair strategies* in the presence of inconsistency.

The evaluation of inconsistency in a database helps us understand it better. In particular, a measure of inconsistency allows us to determine *how reliable* the information contained in a database is and how this information could be used (for example, for inferential purposes) and even modified or adjusted to meet consistency. In this paper we present a family of measures aimed at quantifying the amount of inconsistency in probabilistic knowledge bases. Each measure in the family that we present here quantifies the amount of inconsistency by determining *how far* the knowledge base is from consistency based on a notion of distance given by a particular p-norm (which we call p-distance). This approach to quantifying inconsistency connects to possible repairs of the knowledge base in the sense that the computation of these measures yields an adjustment in the probability assignments that makes the knowledge base consistent.

Our approach to dealing with inconsistency is thus *graded* in the sense that we take into account the *amount* of adjustment of the degrees of certainty (i.e., probabilities) in the knowledge base. Alternative, non-graded approaches to measuring inconsistency in databases (not particularly defined for probabilistic databases) there are several. Among the most popular we have those based on the number and composition of the minimal inconsistent subsets contained in the database (see [9] or [10]). In connection with these approaches to evaluating inconsistency we have repair strategies based on the removal of statements that occur in *minimal inconsistent subsets* of the database (see for example [6,11] or [16] for repair strategies of this nature). Graded approaches to quantify inconsistency in probabilistic knowledge bases we know only [17] and, as will be made clear later, is essentially different to ours.

☆ Supported by the Vienna Science and Technology Fund (*WWTF*), Grant MA07-016.
* Tel.: +43 1 58801 10456.
E-mail address: picado@logic.at

0888-613X/$ - see front matter
doi:10.1016/j.ijar.2011.02.003

Our case example throughout the paper will be a coded version of a fragment of the knowledge base of CADIAG-2 (Computer Assisted DIAGnosis), a well-known rule-based expert system aimed at providing support in diagnostic decision making in the field of internal medicine. In fact, the work presented in this paper is mostly motivated by CADIAG-2.

As will be explained later in more detail the knowledge base of CADIAG-2 can be regarded, at least for consistency-checking and evaluation purposes, as probabilistic. An assumption that we believe to be implicit in CADIAG-2 and in most (if not all) probabilistic databases is the fact that the probabilities of the conditioning events in probabilistic conditional statements be strictly greater than zero. By assuming such fact we are excluding the possibility that a probability function *satisfy* [1] a conditional statement unless it assigns a probability strictly greater than zero to its conditioning event. In a large number of papers that deal with topics related to probabilistic satisfiability such an assumption is not taken (for example, in [17]).

Our paper is structured as follows: in Section 2 we introduce most of the notation we will be using throughout the paper and some preliminary definitions. In Section 3 we briefly describe the relevant features of the binary fragment of the knowledge base of CADIAG-2 that we will be using as our case example. Section 4 introduces and studies the family of measures of inconsistency announced above. In Section 5 we deal with some aspects of the computation of the inconsistency measures presented in Section 4 for the general case and, in Section 6, we focus on our case example and we measure the amount of inconsistency of the binary fragment of CADIAG-2's knowledge base by means of our inconsistency measures.

2. Preliminary definitions and notation

Throughout this paper we will be working with a finite propositional language $L = \{p_1, \ldots, p_l\}$, for some $l \in \mathbb{N}$, and will denote by *SL* its closure under boolean connectives:conjunction ($\wedge$), disjunction ($\vee$) and negation ($\neg$). Within the context of CADIAG-2 our language L will represent the set of medical entities occurring in the inference rules of the system. Such medical entities fall into two general types: symptoms, findings, signs and test results (to which we will commonly refer as *symptoms*) that form the subset $S \subset L$ on the one hand and therapies and diseases (to which we will commonly refer as *diseases*) that form the subset $D \subset L$ on the other.

We will use the symbols $\perp$ and $\top$ for classical *contradiction* and classical *tautology*, respectively.

Definition 1. *Let $\omega : SL \longrightarrow [0, 1]$. We say that ω is a probability function on L if the following two conditions hold, for all $\theta, \phi \in SL$:*

- *If $\models \theta$ then $\omega(\theta) = 1$.*
- *If $\models \neg(\theta \wedge \phi)$ then $\omega(\theta \vee \phi) = \omega(\theta) + \omega(\phi)$.* [2]

We can restrict probability functions to the set $[0, 1] \cap \mathbb{Q}$. Such probability functions will be called *rational*.

A probability distribution ω on L can be characterized by the values it assigns to the expressions of the form

$$\pm p_1 \wedge \cdots \wedge \pm p_l,$$

which we call *states* or *worlds*, where $+p$ and $-p$ stand for p and $\neg p$, respectively. We denote the set of states in L by W, which we assume to be ordered in some way, and define W_ϕ as follows, for $\phi \in SL$:

$$W_\phi = \{\alpha \in W \mid \alpha \models \phi\}.$$

Based on this characterization we can identify probability functions on L with 2^l-coordinate vectors in $\mathbb{D}_{2^l}$, with

$$\mathbb{D}_{2^l} = \left\{ (x_1, \ldots, x_{2^l}) \mid x_i \geq 0, \sum_i x_i = 1 \right\},$$

where the coordinate x_i stands for the probability assigned to the state $\alpha_i \in W$.

Sentences in *SL* can also be identified (up to semantical equivalence) with 2^l-coordinate vectors. For $\phi \in SL$ we can set the 2^l-coordinate vector $\vec{r}_\phi$ as follows: for each $i \in \{1, \ldots, 2^l\}$ and $\alpha_i \in W$, $r^i_\phi = 1$ if $\alpha_i \in W_\phi$ and $r^i_\phi = 0$ otherwise.

Let us set $\mathcal{K} = \{\phi|\theta \mid \phi, \theta \in SL\}$, the set of conditional statements in *SL*.

We define probability on conditional statements in *SL* from the notion of unconditional probability in the conventional way. For ω a probability function on L and $\phi|\theta \in \mathcal{K}$,

$$\omega(\phi|\theta) = \frac{\omega(\phi \wedge \theta)}{\omega(\theta)}.$$

[1] Precise definitions of all these concepts will be given later.

[2] Here (and throughout) $\models$ is classical entailment.

We will denote the collection of closed intervals contained in [0, 1] by $\mathfrak{I}$. Intervals of the form $[\eta, \eta] \in \mathfrak{I}$ (i.e., real point values in the interval [0, 1]) will normally be denoted by η itself.

For the next definition let us consider $\Delta \subseteq \mathcal{K}$.

Definition 2. *An assignment v on Δ is a map from Δ to $\mathfrak{I}$.*

The assignment v will be said to be *point valued* if $v(\phi|\theta) \in [0, 1]$ for all $\phi|\theta \in \Delta$.

We denote the set of assignments on Δ by V_Δ and set

$$\mathcal{PK} = \{[\Delta, v] \mid \Delta \subseteq \mathcal{K},\ \Delta \neq \emptyset,\ v \in V_\Delta\}.$$

We will sometimes write $[\Delta, v] \in \mathcal{PK}$ in extended form; i.e., as

$$\{v(\phi|\theta) = \Omega \mid \phi|\theta \in \Delta\},$$

with $\Omega \in \mathfrak{I}$ or, more in keeping with the notation later employed for the rules of CADIAG-2, as

$$\{\langle \phi|\theta, \Omega\rangle \mid \phi|\theta \in \Delta,\ v(\phi|\theta) = \Omega\}.$$

Let $[\Delta, v] \in \mathcal{PK}$ and $\Delta' \subset \Delta$. We denote by $v_{|\Delta'}$ the restriction of v on Δ'.

Consider now $[\Delta_1, v_1], [\Delta_2, v_2] \in \mathcal{PK}$ and assume that, if $\phi|\theta \in \Delta_1 \cap \Delta_2$, $v_1(\phi|\theta) = v_2(\phi|\theta)$. We define the assignment $v = v_1 + v_2$ on $\Delta_1 \cup \Delta_2$ as follows: $v(\phi|\theta) = v_1(\phi|\theta)$ for all $\phi|\theta \in \Delta_1$ and $v(\phi|\theta) = v_2(\phi|\theta)$ for all $\phi|\theta \in \Delta_2$.

Definition 3. *We say that the probability function ω on L satisfies $[\Delta, v] \in \mathcal{PK}$ (denoted $\models_\omega [\Delta, v]$) if, for all $\phi|\theta \in \Delta$, we have that $\omega(\theta) > 0$ and $\omega(\phi|\theta) \in v(\phi|\theta)$.*

In that sense we say that $[\Delta, v]$ is *(probabilistically) satisfiable* or *consistent* [3] if there exists a probability function ω on L that satisfies $[\Delta, v]$.

Definition 4. *We say that $[\Delta, v]$ is a minimal unsatisfiable set (or minimal inconsistent set) if $[\Delta, v]$ is not satisfiable and, for all $\Delta' \subset \Delta$, $[\Delta', v_{|\Delta'}]$ is satisfiable.*

In order to prove some results in this paper it will be useful to regard our language L as a collection of unary predicates or sets in a first-order language and SL the closure of predicates in L under boolean combinations. For the next (and last) two definitions of this section we will regard L as a set of unary predicates.

Definition 5. *An interpretation $\mathcal{I}$ of L is a pair $(D^{\mathcal{I}}, \nu^{\mathcal{I}})$, where $D^{\mathcal{I}}$ is a finite non-empty set and $\nu^{\mathcal{I}}$ is a map from $L \times D^{\mathcal{I}}$ to $\{0, 1\}$.*

A valuation $\nu^{\mathcal{I}}$ can be extended to elements in $SL \times D^{\mathcal{I}}$ in its conventional way.

Given an interpretation $\mathcal{I}$ of L, we will refer to the elements in $D^{\mathcal{I}}$ by latin characters $a, b, c, \ldots$

For what follows let $\mathcal{I}$ be an interpretation of L and $[\Delta, v] \in \mathcal{PK}$.

Definition 6. *$\mathcal{I}$ is a (probabilistic) model of $[\Delta, v]$ if, for all $\phi|\theta \in \Delta$, we have that*

$$\frac{\sum_{a \in D^{\mathcal{I}}} \nu^{\mathcal{I}}(\phi \wedge \theta, a)}{\sum_{a \in D^{\mathcal{I}}} \nu^{\mathcal{I}}(\theta, a)} \in v(\phi|\theta).$$

3. The knowledge base of CADIAG-2

The medical expert system CADIAG-2 consists of two main components: the inference engine and the knowledge base. The inference engine is based on methods of approximate reasoning in fuzzy set theory, in the sense of [18]. Formalizations and analyses of it can be found in, for example, [7] or [15]. The knowledge base of the system consists of a set of *IF-THEN* rules intended to represent relationships between distinct medical entities. The vast majority of them are binary (i.e., they relate single medical entities) and only such rules are considered in our paper. The one that follows is an example of a binary rule of CADIAG-2 (taken from [3]):

[3] We use both words interchangeably throughout the article.

IF suspicion of liver metastases by liver palpation
THEN pancreatic cancer
with *degree of confirmation* 0.3

The *degree of confirmation* refers, intuitively, to the degree to which the antecedent (i.e., '*suspicion of liver metastases by liver palpation*' in the example above) confirms the consequent (i.e., '*pancreatic cancer*' above). It is claimed in part of the literature on CADIAG-2, like for example in [1], that such degrees of confirmation can be understood as probabilities and the rules themselves as probabilistic conditional statements. In most of the literature on CADIAG-2 though, like for example in [2] or [3], an alternative, non-probabilistic interpretation for the degrees of confirmation is suggested. However, it is proved in [12] that the binary fragment of CADIAG-2's knowledge base is *satisfiable* in terms of the suggested non-probabilistic interpretation if and only if it is *probabilistically satisfiable*. Thus, at least for consistency-checking and evaluation purposes, one can regard CADIAG-2's binary knowledge base as probabilistic no matter which of these two possible interpretations is to be chosen.

We will represent binary rules in CADIAG-2 by pairs of the form $\langle q|p, \eta\rangle$, with $q|p \in \mathcal{K}$, $p, q \in L$ and the degree of confirmation $\eta \in [0, 1]$. We will denote the set consisting of such rules by $\Phi_{CB} \in \mathcal{PK}$.

We can classify rules in Φ_{CB} into three different types: rules in which both antecedent (or, in probabilistic terms, conditioning event) and consequent (or conditioned event) are medical entities in S (*symptom–symptom*), rules in which both antecedent and consequent are medical entities in D (*disease–disease*) and those in which the antecedent is a medical entity in S and the consequent an entity in D (*symptom–disease*).[4] The degree of confirmation in a rule of the first two types is a value in the set $\{0, 1\}$ and it is in this sense that we say that rules of these types are *classical*.

Table 1 shows in numbers the composition of Φ_{CB}.

Table 1
Composition of Φ_{CB}.

Number of symptoms	1761
Number of diseases	341
Number of *symptom–symptom* rules	720
Number of *disease–disease* rules	218
Number of *symptom–disease* rules	17573

4. Measuring inconsistency: *p*-distance

Our approach to measuring inconsistency for knowledge bases of the form $[\Delta, v] \in \mathcal{PK}$ is based on the quantification of the *minimal* adjustment that one needs to make on v in order for $[\Delta, v]$ to be satisfiable. Such an approach is similar in nature to that defined in [17]. The main difference between both approaches rests on the notion of satisfiability for probabilistic conditional statements: null probability for the conditioning event in a probabilistic statement is allowed in [17] and makes the statement satisfiable by default. Our approach becomes much more complex in comparison to [17] due mostly to our definition of satisfiability, which we believe is more natural and intuitive, and to a more general notion of minimal adjustments.

In order to quantify the minimal adjustment we will rely on the so-called *p-norm* which, for a vector $\vec{x} \in \mathbb{R}^m$, is given by

$$\|\vec{x}\|_p = \left(\sum_{i=1}^{m} |x_i|^p\right)^{\frac{1}{p}},$$

for $p \geq 1$ (not necessarily an integer). The most common p-norms are certainly those of order 1 and 2 (i.e., 1-norm and 2-norm). We also have

$$\|\vec{x}\|_\infty = \lim_{p\to\infty}\left(\sum_{i=1}^{m} |x_i|^p\right)^{\frac{1}{p}} = \max_i |x_i|$$

among the most common norms, which we will refer to as the ∞-norm and regard as a p-norm with $p = \infty$. In this context we will abuse notation slightly and will write $[1, \infty]$ for the range of possible values for p.

Let $\epsilon \in {}^*\mathbb{R}$ be a *non-zero infinitesimal* in the set of hyperreal numbers.[5] We will be working with the set

$$\mathcal{A} = \left\{t,\ t+\epsilon,\ \frac{1}{\epsilon}\ \middle|\ t \in \mathbb{R}\right\}.$$

[4] CADIAG-2's knowledge base formally contains values for conditional relations with a medical entity in D as the antecedent and a medical entity in S as the consequent. However, such rules are not used by CADIAG-2's inference mechanism and are not taken into account for our examples in this paper.

[5] That is to say, ϵ is such that $0 < \epsilon < \frac{1}{n}$ for all $n \in \mathbb{N}$.

The reason why we consider $\epsilon \in {}^*\mathbb{R}$, although it should not come at a surprise given the notion of satisfiability we are working with, will be made clear below. As is common, we will denote the standard (i.e., real) part of a finite number $a \in \mathcal{A}$ by $st(a)$ (for more on hyperreal numbers and, in general, nonstandard analysis see for example [8]).

We mention at this point that we are not interested in infinitesimal precision and thus we do not consider algebraic operations involving infinitesimals. The consideration of ϵ as an infinitesimal magnitude suffices for our purposes.

For what follows let $[\Delta, v] \in \mathcal{PK}$ and $\Gamma \subset \Delta$, with

$$\Delta' = \Delta - \Gamma = \{\phi_1|\theta_1, \ldots, \phi_k|\theta_k\},$$

for some $k \in \mathbb{N}$, and $v(\phi_i|\theta_i) = [\underline{\eta}_i, \overline{\eta}_i]$ for all $i \in \{1, \ldots, k\}$. Γ is intended to represent the set of conditional statements in Δ that are regarded as *correctly* evaluated by v and that should (arguably) not be considered when assessing the p-distance of $[\Delta, v]$.

We define $[\Delta, v_{\vec{x}}]$ to be the set

$$\{v_{\vec{x}}(\phi_i|\theta_i) = [\underline{\eta}_i - x_i, \overline{\eta}_i + x_{i+k}] \mid i \in \{1, \ldots, k\}\} \cup \{v_{\vec{x}}(\phi|\theta) = v(\phi|\theta) \mid \phi|\theta \in \Gamma\},$$

where x_i, x_{i+k} are positive real values satisfying the constraint

$$0 \leq \underline{\eta}_i - x_i \leq \overline{\eta}_i + x_{i+k} \leq 1,$$

for all $i \in \{1, \ldots, k\}$.

For the next definitions let $[\Delta, v] \in \mathcal{PK}$ and consider $[\Delta, v_{\vec{x}}]$ and $\Gamma \subset \Delta$ as above, with $\vec{x} \in \mathbb{R}^{2k}$ and $p \in [1, \infty]$.

Definition 7. *We define the real set $F_\Gamma^p([\Delta, v])$ as follows:*

$$F_\Gamma^p([\Delta, v]) = \{\|\vec{x}\|_p \mid [\Delta, v_{\vec{x}}] \text{ is satisfiable}\}.$$

Notice that $F_\Gamma^p([\Delta, v])$ is bounded below for all $p \in [1, \infty]$ and thus, if it is not empty, its infimum exists (i.e., $\inf(F_\Gamma^p([\Delta, v]))$ exists). Notice also that $\inf(F_\Gamma^p([\Delta, v]))$ corresponds to the p-norm distance between the cartesian product

$$\Pi_{\phi|\theta \in \Delta} v(\phi|\theta)$$

and the set given by the point-valued assignments u on Δ for which $[\Delta, u]$ is satisfiable.

Definition 8. *We define the p-distance (to consistency) of $[\Delta, v]$ with respect to the set Γ – denoted $DC_\Gamma^p([\Delta, v])$ – as follows:*

- $F_\Gamma^p([\Delta, v]) \neq \emptyset$ *and* $\min(F_\Gamma^p([\Delta, v])) = \inf(F_\Gamma^p([\Delta, v]))$ *then*

  $$DC_\Gamma^p([\Delta, v]) = \inf(F_\Gamma^p([\Delta, v])).$$

- *If* $F_\Gamma^p([\Delta, v]) \neq \emptyset$ *and* $\min(F_\Gamma^p([\Delta, v]))$ *does not exist then*

  $$DC_\Gamma^p([\Delta, v]) = \inf(F_\Gamma^p([\Delta, v])) + \epsilon.$$

- *If* $F_\Gamma^p([\Delta, v]) = \emptyset$ *then we set*

  $$DC_\Gamma^p([\Delta, v]) = \frac{1}{\epsilon}.$$

That $DC_\Gamma^p([\Delta, v])$ is well defined is clear. If $[\Gamma, v_{|\Gamma}]$ is unsatisfiable or there exists a statement of the form $\phi|\bot$ in Δ then we have that $F_\Gamma^p([\Delta, v]) = \emptyset$ and thus that $DC_\Gamma^p([\Delta, v]) = \frac{1}{\epsilon}$. If $[\Gamma, v_{|\Gamma}]$ is satisfiable and define Λ to be the set of probability distributions on L that satisfy $[\Gamma, v_{|\Gamma}]$ we have that each $\omega \in \Lambda$ with $\omega(\theta_i) > 0$ for all $i \in \{1, \ldots, k\}$ satisfies $[\Delta, v_{\vec{x}}]$ for some real values x_i, with $i \in \{1, \ldots, 2k\}$, and thus, if any such probability distribution existed, $F_\Gamma^p([\Delta, v])$ would not be empty. On the other hand, if there were no such probability distributions then $F_\Gamma^p([\Delta, v])$ would be empty, in which case we would have $DC_\Gamma^p([\Delta, v]) = \frac{1}{\epsilon}$.

We will write $DC^p([\Delta, v])$ instead of $DC_\emptyset^p([\Delta, v])$ whenever there is no set $\Gamma \subset \Delta$ with respect to which we are defining the p-distance of $[\Delta, v]$.

For our next definition let us consider again $[\Delta, v] \in \mathcal{PK}$, with $[\Delta, v_{\vec{x}}]$ and $\Gamma \subset \Delta$ as above.

Definition 9. *We define the set $R_\Gamma([\Delta, v])$ of repairs of $[\Delta, v]$ with respect to Γ as follows:*

$$R_\Gamma([\Delta, v]) = \{\vec{x} \in \mathbb{R}^{2k} \mid [\Delta, v_{\vec{x}}] \text{ is satisfiable}\}.$$

For the next definition let $\vec{x} \in \mathbb{R}^{2k}$ be a repair in $R_\Gamma([\Delta, v])$, for $k \in \mathbb{N}$.

Definition 10. *We say that $\vec{x}$ is p-optimal if $\|\vec{x}\|_p = DC^p_\Gamma([\Delta, v])$.*

4.1. Properties of DC^p

Throughout all this subsection let $[\Delta, v] \in \mathcal{PK}$, with $\Gamma \subset \Delta$ and

$$\Delta - \Gamma = \{\phi_1|\theta_1, \ldots, \phi_k|\theta_k\},$$

for some $k \in \mathbb{N}$.

First of all notice that it is not generally true that $DC^p_\Gamma([\Delta, v]) \in \mathbb{R}$. This is, as claimed earlier, a consequence of our notion of satisfiability and, more particularly, the requirement that a conditioning event in a conditional statement be assigned probability strictly greater than zero. To show this more clearly with an example, consider the subset $[\Delta, v]$ given by the following set of rules in Φ_{CB}:

$$\{\langle D36|S157, 0.3\rangle, \langle D81|S157, 0.15\rangle, \langle D81|D36, 1\rangle\}.$$

It is clear that $[\Delta, v]$ is unsatisfiable. Let us consider the sequence of probability functions $\{\omega_n\}_{n\in\mathbb{N}}$ on $\{D36, D81, S157\}$ characterized by the following assignments:

- $\omega_n(S157) = \frac{1}{n}$,
- $\omega_n(D36) = \frac{0.3}{n} + \left(1 - \frac{1}{n}\right)$,
- $\omega_n(D81) = \frac{0.15}{n} + \left(1 - \frac{1}{n}\right)$,
- $\omega_n(S157 \wedge D36) = \frac{0.3}{n}$,
- $\omega_n(S157 \wedge D81) = \frac{0.15}{n}$,
- $\omega_n(D36 \wedge D81) = \frac{0.15}{n} + \left(1 - \frac{1}{n}\right)$.

Let $\lim_{n\to\infty} \omega_n = \omega$. It can be easily shown that such assignments fully determine ω_n and ω as probability functions on the language $\{D36, D81, S157\}$, for all $n \in \mathbb{N}$.

We have that, for all $n \in \mathbb{N}$, $\omega_n(D36|S157) = 0.3$, $\omega_n(D81|S157) = 0.15$ and $\omega_n(D81|D36) = \frac{n-0.85}{n-0.7}$. It is clear that for any $\delta > 0$ we can find $N \in \mathbb{N}$ such that, for all $n > N$, we have that $\omega_n(D81|D36) > 1 - \delta$. It is also clear that ω does not satisfy $[\Delta, v]$ (in fact, $[\Delta, v]$ is unsatisfiable). Thus we have that $DC^p([\Delta, v]) = \epsilon$ for all $p \in [1, \infty]$.

For the next result let us consider $[\Delta_1, v_1], [\Delta_2, v_2] \in \mathcal{PK}$ be such that $v_1(\phi|\theta) = v_2(\phi|\theta)$, for all $\phi|\theta \in \Delta_1 \cap \Delta_2$, and $\Gamma_1 \subset \Delta_1$, $\Gamma_2 \subset \Delta_2$.

Proposition 11. *For all $p \in [1, \infty]$ we have that*

$$DC^p_{\Gamma_1}([\Delta_1, v_1]) \leq DC^p_{\Gamma_1\cup\Gamma_2}([\Delta_1 \cup \Delta_2, v_1 + v_2]).$$

Proof. It follows directly from the definition of DC^p. $\square$

For the next proposition let us assume that $F^p_\Gamma([\Delta, v]) \neq \emptyset$.

Lemma 12. *$st(DC^p_\Gamma([\Delta, v]))$ is continuous and decreasing on p, for $p \in [1, \infty)$.*

Proof. The result follows from the well-known fact that, for general $\vec{x} \in \mathbb{R}^m$ and $m \in \mathbb{N}$, $\|\vec{x}\|_p$ is continuous and decreasing on p. $\square$

For the next results let $p_1, p_2 \in [1, \infty]$.

Proposition 13. *If $p_1 < p_2$ then $DC^{p_1}_\Gamma([\Delta, v]) \geq DC^{p_2}_\Gamma([\Delta, v])$.*

Proof. Let us assume that $p_1 < p_2$. If $DC^{p_2}_\Gamma([\Delta, v]) \in \mathbb{R}$ then the result follows from Lemma 12. Let us assume then that $DC^{p_2}_\Gamma([\Delta, v]) = t + \epsilon \notin \mathbb{R}$ and that $DC^{p_1}_\Gamma([\Delta, v]) = t$, for some $t \in \mathbb{R}$. Let $\vec{x} \in R_\Gamma([\Delta, v])$ be a p_1-optimal repair of $[\Delta, v]$ (i.e., $\|\vec{x}\|_{p_1} = t$) but notice that, since $\|\vec{x}\|_p$ is decreasing on the order p, we have that $\|\vec{x}\|_{p_1} \geq \|\vec{x}\|_{p_2}$ and thus $DC^{p_1}_\Gamma([\Delta, v]) \geq DC^{p_2}_\Gamma([\Delta, v])$. $\square$

Proposition 14. *$DC^p_\Gamma([\Delta, v]) = 0$ if and only if $[\Delta, v]$ is satisfiable.*

Proof. This result follows directly from the definition of DC^p. $\square$

Proposition 15. *Let $p \in [1, \infty)$. If $F^p_\Gamma([\Delta, v]) \neq \emptyset$ then $DC^p_\Gamma([\Delta, v])$ is at most $k^{\frac{1}{p}}$.*

Proof. Consider $[\Delta, v] = \{v(q_1|q_1) = 0, \ldots, v(q_k|q_k) = 0\}$, for $q_1, \ldots, q_k \in L$, and $\Gamma = \emptyset$. First of all notice that $q_i|q_i$, $q_j|q_j$, for any $i, j \in \{1, \ldots, k\}$ with $i \neq j$, are not *equivalent* in the sense that we can define a probability function ω on L such that $\omega(q_i|q_i) = 1$ but that assigns no probability to $q_j|q_j$ by setting $\omega(q_j) = 0$. It is clear that, for all $p \in [1, \infty)$, we will have that $DC^p_\Gamma([\Delta, v]) = k^{\frac{1}{p}}$. It is also clear that this is the biggest p-distance for a probabilistic database in $\mathcal{PK}$ given the premises. □

Corollary 16. *If $F^\infty_\Gamma([\Delta, v]) \neq \emptyset$ then $DC^\infty_\Gamma([\Delta, v])$ is at most 1.*

Next we want to prove that DC^{p_1} and DC^{p_2} (for $p_1 \neq p_2$) are *essentially* distinct, by which we mean that the difference between DC^{p_1} and DC^{p_2} does not reduce only to the (possible) difference in magnitude of the values they assign to a certain knowledge base in $\mathcal{PK}$ but also to the ordering they induce on $\mathcal{PK}$.

Proposition 17. *DC^{p_1} and DC^{p_2} induce distinct orderings on $\mathcal{PK}$.*

Proof. Let us assume that $p_1 < p_2$ and that $p_1, p_2 \in [1, \infty)$. Let us consider the set $[\Delta_1, v_1] = \{v(q|q) = 0\}$, for $q \in L$. We will have that $DC^{p_1}([\Delta_1, v_1]) = DC^{p_2}([\Delta_1, v_1]) = 1$. Consider now the set

$$[\Delta_2, v_2] = \{v(q_1|q_1) = 1 - \lambda, v(q_2|q_2) = 1 - \lambda\},$$

for $q_1, q_2 \in L$. We will have that $DC^{p_1}([\Delta_2, v_2]) = \lambda 2^{\frac{1}{p_1}}$ and that $DC^{p_2}([\Delta_2, v_2]) = \lambda 2^{\frac{1}{p_2}}$. First notice that, for all $\lambda \in (0, 1]$, it is the case that

$$\lambda 2^{\frac{1}{p_1}} > \lambda 2^{\frac{1}{p_2}}.$$

Notice also that, since $p_1 \in [1, \infty)$, there needs to exist $\lambda \in [0, 1]$ such that

$$\lambda 2^{\frac{1}{p_1}} > 1 > \lambda 2^{\frac{1}{p_2}}$$

and thus such that

$$DC^{p_1}([\Delta_1, v_1]) < DC^{p_1}([\Delta_2, v_2])$$

and

$$DC^{p_2}([\Delta_1, v_1]) > DC^{p_2}([\Delta_2, v_2]). \qquad \square$$

In simple words, what Proposition 17 tells us is that, for any two probabilistic knowledge bases $[\Delta_1, v_1], [\Delta_2, v_2] \in \mathcal{PK}$, we can have that $[\Delta_1, v_1]$ is *more* inconsistent than $[\Delta_2, v_2]$ according to DC^{p_1}, for some $p_1 \in [1, \infty]$, but that $[\Delta_1, v_1]$ is *less* inconsistent than $[\Delta_2, v_2]$ according to DC^{p_2}, with $p_1 \neq p_2$.

4.2. Normalized p-distance

In this subsection we introduce the notion of *normalized* p-distance. In order to do so let us consider $[\Delta, v] \in \mathcal{PK}$, with $\Gamma \subset \Delta$ and

$$\Delta - \Gamma = \{\phi_1|\theta_1, \ldots, \phi_k|\theta_k\},$$

for some $k \in \mathbb{N}$.

Definition 18. *Let $p \in [1, \infty)$. The normalized p-distance (to consistency) of $[\Delta, v]$ with respect to Γ –denoted $\overline{DC}^p_\Gamma([\Delta, v])$–is defined from $DC^p_\Gamma([\Delta, v])$ as follows:*

- *If $DC^p_\Gamma([\Delta, v])$ is finite (i.e., if $DC^p_\Gamma([\Delta, v]) \neq \frac{1}{\epsilon}$) then we set*

$$\overline{DC}^p_\Gamma([\Delta, v]) = \frac{st(DC^p_\Gamma([\Delta, v]))}{k^{\frac{1}{p}}} + DC^p_\Gamma([\Delta, v]) - st(DC^p_\Gamma([\Delta, v])).$$

- *If $DC^p_\Gamma([\Delta, v]) = \frac{1}{\epsilon}$ then we set $\overline{DC}^p_\Gamma([\Delta, v]) = DC^p_\Gamma([\Delta, v])$.*

The normalized ∞-distance of $[\Delta, v]$ coincides with its ∞-distance:

$$\overline{DC}^{\infty}_{\Gamma}([\Delta, v]) = DC^{\infty}_{\Gamma}([\Delta, v]).$$

The normalized p-distance gives us the *ratio* between the value $DC^p_{\Gamma}([\Delta, v])$ and the maximum value that $DC^p_{\Gamma}([\Delta, v])$ could attain given its cardinality and configuration (i.e., $k^{\frac{1}{p}}$, see Proposition 15). Thus, clearly, $\overline{DC}^p_{\Gamma}([\Delta, v])$ is at most 1. The normalized p-distance brings into play the size of the knowledge base and gives a better ground to *compare* the amount of inconsistency of distinct knowledge bases.

For the next results let us assume that $F^p_{\Gamma}([\Delta, v]) \neq \emptyset$.

Lemma 19. *$st(\overline{DC}^p_{\Gamma}([\Delta, v]))$ is continuous and increasing on p, for $p \in [1, \infty)$.*

Proof. The result follows from the fact that, for general $\vec{x} \in \mathbb{R}^m$ and $m \in \mathbb{N}$, $\frac{\|\vec{x}\|_p}{k^{\frac{1}{p}}}$ is continuous and increasing on p. $\square$

For the next proposition let $p_1, p_2 \in [1, \infty]$.

Proposition 20. *If $p_1 < p_2$ then $\overline{DC}^{p_1}_{\Gamma}([\Delta, v]) \leq \overline{DC}^{p_2}_{\Gamma}([\Delta, v])$.*

Proof. Let us assume that $p_1 < p_2$. If $\overline{DC}^{p_1}_{\Gamma}([\Delta, v]) \in \mathbb{R}$ then the result follows from Lemma 19. Let us assume then that $\overline{DC}^{p_1}_{\Gamma}([\Delta, v]) = t + \epsilon \notin \mathbb{R}$ and that $\overline{DC}^{p_2}_{\Gamma}([\Delta, v]) = t$, for some $t \in \mathbb{R}$. Let $\vec{x} \in R_{\Gamma}([\Delta, v])$ be a p_2-optimal repair of $[\Delta, v]$ (i.e., $\|\vec{x}\|_{p_2} = t$) but notice that, since $\frac{\|\vec{x}\|_p}{k^{\frac{1}{p}}}$ is increasing on p, we have that

$$\frac{\|\vec{x}\|_{p_1}}{k^{\frac{1}{p_1}}} \leq \frac{\|\vec{x}\|_{p_2}}{k^{\frac{1}{p_2}}}$$

and thus $\overline{DC}^{p_1}_{\Gamma}([\Delta, v]) \leq \overline{DC}^{p_2}_{\Gamma}([\Delta, v])$. $\square$

Most results proved about DC^p in the previous subsection also hold for $\overline{DC}^p$. In particular, for $p_1, p_2 \in [1, \infty]$ and $p_1 \neq p_2$, we will also have that $\overline{DC}^{p_1}$ and $\overline{DC}^{p_2}$ induce distinct orderings on $\mathcal{PK}$ and thus that $\overline{DC}^{p_1}$ and $\overline{DC}^{p_2}$ are *essentially* distinct measures in the sense indicated above.

4.3. Examples and some remarks

In this subsection we consider some examples (the first one extracted from Φ_{CB}) in order to illustrate some differences in DC^p and the p-optimal repairs for distinct values of p. The values we will focus on are the most common when dealing with p-norms (i.e., 1, 2 and ∞).

Let us consider for our first example the set $[\Delta, v] \subset \Phi_{CB}$ given by the rules

$$\langle D10|S668, 0.8\rangle, \langle D25|S668, 0.1\rangle, \langle D70|S668, 0.8\rangle,$$

$$\langle D25|D10, 1\rangle, \langle D25|D70, 1\rangle$$

in Φ_{CB}, with $S668 \in S$ and $D10, D25, D70 \in D$. Let us assume that the rules $\langle D25|D10, 1\rangle$ and $\langle D25|D70, 1\rangle$ are believed to be accurate by the modeler and thus set

$$\Gamma = \{D25|D10, D25|D70\}.$$

We will have that $DC^1_{\Gamma}([\Delta, v]) = 0.7$, $DC^2_{\Gamma}([\Delta, v]) \simeq 0.57$[6] and also that $DC^{\infty}_{\Gamma}([\Delta, v]) = 0.35$. The normalized distances will be $\overline{DC}^1_{\Gamma}([\Delta, v]) \simeq 0.23$, $\overline{DC}^2_{\Gamma}([\Delta, v]) \simeq 0.33$ and $\overline{DC}^{\infty}_{\Gamma}([\Delta, v]) = 0.35$.

There is a unique 1-optimal repair in $R_{\Gamma}([\Delta, v])$ and it yields a unique modification in the assignment of the conditional statement $D25|S668$ by a magnitude of 0.7 (i.e., such optimal repair would consist of the replacement of $\langle D25|S668, 0.1\rangle$ for $\langle D25|S668, [0.1, 0.8]\rangle$). Notice that, in such repair, the whole weight of the adjustment rests on a single statement in Φ_{CB} and that, in the lack of knowledge of how reliable or accurate the assignment v is on the distinct statements in $\Delta - \Gamma$ and in the lack of any other rationality criterion that allows us to discriminate among possible repairs, might not be the most reasonable repair strategy. Instead, the unique ∞-optimal repair in $R_{\Gamma}([\Delta, v])$ offers in that sense a more balanced approach by inducing an adjustment of magnitude 0.35 in the assignment on each one of the statements in $\Delta - \Gamma$. The unique 2-optimal repair in $R_{\Gamma}([\Delta, v])$ involves an adjustment of magnitude $\frac{7}{15}$ in the assignment on the statement $D25|S668$ and an adjustment of magnitude $\frac{7}{30}$ on both $D10|S668$ and $D70|S668$. This simple example favours the view that, in the lack

[6] '$\simeq$' stands here and throughout for 'approximately'. Like in this case, we sometimes round decimal numbers to two significant figures in order to make comparisons more easily visualised.

of knowledge about the reliability of the statements in our knowledge base and in the lack of any other applicable rationality criterion, large values of p seem to be more adequate when it comes to considering p-optimal repairs for our knowledge base.

For our second example consider

$$[\Delta, v]_k = \{v(q_1|q_1) = 1, \dots, v(q_{k-1}|q_{k-1}) = 1, v(q_k|q_k) = 0\},$$

with $q_1, \dots, q_k \in L$ distinct propositional variables, for $k \in \mathbb{N}$. Here we will have that

$$DC^1([\Delta, v]_k) = DC^2([\Delta, v]_k) = DC^\infty([\Delta, v]_k) = 1$$

and that $\overline{DC}^1([\Delta, v]_k) = \frac{1}{k}$, $\overline{DC}^2([\Delta, v]_k) = \frac{1}{\sqrt{k}}$ and $\overline{DC}^\infty([\Delta, v]_k) = 1$. $\overline{DC}^\infty$ assigns to $[\Delta, v]_k$ the highest possible amount of inconsistency in a knowledge base regardless of the cardinality of $[\Delta, v]_k$. Examples like this one in which the repair to restore consistency rests on a small number of statements or on a severe adjustment in a small number of them brings $\overline{DC}^p$, for big values of p, to assign high amounts of inconsistency regardless of the cardinality of the knowledge base. Although arguable, it seems reasonable to consider $[\Delta, v]_{k_1}$ *more* inconsistent than $[\Delta, v]_{k_2}$ whenever $k_1 < k_2$ based on the simple fact that the former is *bigger*. On that assumption and in the view of examples like these in which the adjustment of possible optimal repairs of the knowledge base concentrates on a small number of statements it seems more reasonable to measure the p-distance to consistency by considering small values of p.

5. Computation of DC^p and R

In this section we deal with some issues regarding the computation of DC^p (and thus R) for a general knowledge base $[\Delta, v] \in \mathcal{PK}$. In particular, we show the connection between its computation and the solution to certain constrained optimization problems.

Let us consider $[\Delta, v]$, with $\Delta' = \Delta - \Gamma$, $\Gamma \subset \Delta$,

$$[\Gamma, v_{|\Gamma}] = \{v(\psi_1|\chi_1) = [\underline{\lambda}_1, \overline{\lambda}_1], \dots, v(\psi_j|\chi_j) = [\underline{\lambda}_j, \overline{\lambda}_j]\}$$

and

$$[\Delta', v_{|\Delta'}] = \{v(\phi_1|\theta_1) = [\underline{\eta}_1, \overline{\eta}_1], \dots, v(\phi_k|\theta_k) = [\underline{\eta}_k, \overline{\eta}_k]\},$$

for some $j \in \{0\} \cup \mathbb{N}$ and $k \in \mathbb{N}$.

Consider the following constrained optimization problem with *optimization variable* the vector $\vec{x} \in \mathbb{R}^{2k}$:

$$\text{minimize} \quad \|\vec{x}\|_p \tag{1}$$

subject to the following constraints:

- $(\underline{\eta}_i - x_i)\vec{r}_{\theta_i} \cdot \vec{y} \le \vec{r}_{\theta_i \wedge \phi_i} \cdot \vec{y} \le (\overline{\eta}_i + x_{i+k})\vec{r}_{\theta_i} \cdot \vec{y}$ for each $i \in \{1, \dots, k\}$,
- $\underline{\lambda}_i\vec{r}_{\chi_i} \cdot \vec{y} \le \vec{r}_{\psi_i \wedge \chi_i} \cdot \vec{y} \le \overline{\lambda}_i\vec{r}_{\chi_i} \cdot \vec{y}$ for each $i \in \{1, \dots, j\}$,
- $\sum_{i=1}^{2^l} y_i = 1$,
- $y_i \ge 0$ for each $i \in \{1, \dots, 2^l\}$,
- $\vec{r}_\theta \cdot \vec{y} \ge \delta$ for each $\phi|\theta \in \Delta$ for fixed $\delta \in (0, 1]$,
- $x_i \ge 0$ for each $i \in \{1, \dots, 2k\}$,
- $0 \le \underline{\eta}_i - x_i \le \overline{\eta}_i + x_{i+k} \le 1$ for each $i \in \{1, \dots, k\}$.

The value attained by the optimization variable $\vec{x}$ in (1) represents a possible repair of $[\Delta, v]$ (i.e., an element in $R_\Gamma([\Delta, v])$) and the vector $\vec{y}$ a probability function on L. The value δ is introduced as a lower bound probability threshold for the conditioning events in the statements in Δ.

We define the constrained optimization problem (2) from (1) by replacing the constraint

$$\vec{r}_\theta \cdot \vec{y} \ge \delta \text{ for each } \phi|\theta \in \Delta \text{ for fixed } \delta \in (0, 1],$$

for the following strict inequality:

$$\vec{r}_\theta \cdot \vec{y} > 0 \text{ for each } \phi|\theta \in \Delta.$$

Let us denote the constrained optimization problems (1) and (2) by $\mathcal{C}_\Gamma^{p,\delta}([\Delta, v])$ and $\mathcal{C}_\Gamma^{p}([\Delta, v])$ respectively. We define $\mathcal{SC}_\Gamma^{p,\delta}([\Delta, v])$, the solution to (1), as follows:

$$\mathcal{SC}_\Gamma^{p,\delta}([\Delta, v]) = \inf_{\vec{y} \in \mathbb{D}_{2^l}} \{\|\vec{x}\|_p \mid (\vec{x}, \vec{y}) \in \mathbb{R}^{2k+2^l} \text{ is } \textit{feasible}\}.$$

By $(\vec{x}, \vec{y})$ being feasible we mean that $(\vec{x}, \vec{y})$ satisfies the constraints in (1). The collection of all such vectors is called the *feasible set* of (1) –see, for example, [5] for more on these concepts and, in general, on the terminology and basic definitions for constrained optimization problems–. Notice that $\mathcal{SC}_\Gamma^{p,\delta}([\Delta, v])$ may not exist.

We define $\mathcal{SC}_\Gamma^{p}([\Delta, v])$, with respect to the feasible set of (2), in an analogous way.

It is clear from the definition of $F_\Gamma^p([\Delta, v])$ that, if $F_\Gamma^p([\Delta, v]) \neq \emptyset$,

$$\mathcal{SC}_\Gamma^{p}([\Delta, v]) = \inf F_\Gamma^p([\Delta, v])$$

and that an *optimal point* $(\vec{x}, \vec{y}) \in \mathbb{R}^{2k+2^l}$ may not exist (i.e., a point $(\vec{x}, \vec{y})$ at which $\mathcal{SC}_\Gamma^{p}([\Delta, v])$, provided it exists, is attained). As is clear, $\vec{x}$ in an optimal point $(\vec{x}, \vec{y})$ yields a p-optimal repair.

From now we will focus on $\mathcal{C}_\Gamma^{p,\delta}([\Delta, v])$ and its computation.

For what follows let us set $C = \{\theta \in SL \mid \phi|\theta \in \Delta\}$ and define $\Pi_{[\Gamma, v_{|\Gamma}]}$ to be the set of probability functions on L that satisfy $[\Gamma, v_{|\Gamma}]$ (if $\Gamma = \emptyset$ then $\Pi_{[\Gamma, v_{|\Gamma}]}$ or just Π coincides with the set of probability functions on L).

Definition 21. *We say that C is η-consistent with respect to $[\Gamma, v_{|\Gamma}]$ if there exists a probability function $\omega \in \Pi_{[\Gamma, v_{|\Gamma}]}$ such that $\omega(\theta) \geq \eta$ for all $\theta \in C$.*

Definition 22. *We say that C is maximally η-consistent with respect to $[\Gamma, v_{|\Gamma}]$ if these two conditions hold:*

- *For all $\lambda < \eta$ we have that C is λ-consistent with respect to $[\Gamma, v_{|\Gamma}]$.*
- *For all $\mu > \eta$ we have that C is not μ-consistent with respect to $[\Gamma, v_{|\Gamma}]$.*

These two definitions are generalizations of the notions of *η-consistency* and *maximal η-consistency* introduced in [13] as *measures* of consistency of classical propositional knowledge bases.

We will denote the maximal η-consistency of C with respect to our set $[\Gamma, v_{|\Gamma}]$ by $mc_{[\Gamma, v_{|\Gamma}]}(C)$ –if $\Gamma = \emptyset$ then we will just write $mc(C)$– and will use the abbreviation $\omega(C) \geq \eta$ to express the fact that $\omega(\theta) \geq \eta$ for all $\theta \in C$. If $\Pi_{[\Gamma, v_{|\Gamma}]} = \emptyset$ then $mc_{[\Gamma, v_{|\Gamma}]}$ is not defined.

Let us assume now that C is maximally η-consistent with respect to $[\Gamma, v_{|\Gamma}]$. Notice that it is not generally the case that there exists a probability function $\omega \in \Pi_{[\Gamma, v_{|\Gamma}]}$ such that $\omega(C) \geq \eta$. To see this consider the example where $[\Gamma, v_{|\Gamma}] = \{\langle p|q, \frac{1}{2}\rangle\}$ and $C = \{\neg q\}$, for $p, q \in L$. It is clear that $mc_{[\Gamma, v_{|\Gamma}]}(C) = 1$ and that there is no probability function $\omega \in \Pi_{[\Gamma, v_{|\Gamma}]}$ such that $\omega(C) = 1$. However, if $\Gamma = \emptyset$ then it is proved in [13] that $mc(C)$ is a rational number and attained by some probability function ω on L.

Consider the following constrained optimization problem with *optimization variable* $t \in \mathbb{R}$:

$$\text{maximize} \quad t$$

subject to the following constraints:

- $\vec{r}_\phi \cdot \vec{y} \geq t$ for each $\phi \in C$,
- $\sum_{i=1}^{2^l} y_i = 1$,
- $y_i \geq 0$ for each $i \in \{1, \ldots, 2^l\}$,
- $\underline{\lambda}_i \vec{r}_{\chi_i} \cdot \vec{y} \leq \vec{r}_{\psi_i \wedge \chi_i} \cdot \vec{y} \leq \overline{\lambda}_i \vec{r}_{\chi_i} \cdot \vec{y}$ for each $i \in \{1, \ldots, j\}$,
- $\vec{r}_\chi \cdot \vec{y} > 0$ for each $\psi|\chi \in \Gamma$.

Notice that if $\Gamma = \emptyset$ then the problem above is a *linear program*.

The solution to this problem, if it exists, yields $mc_{[\Gamma, v_{|\Gamma}]}(C)$:

$$mc_{[\Gamma, v_{|\Gamma}]}(C) = \sup_{\vec{y} \in \mathbb{D}_{2^l}} \{t \mid (t, \vec{y}) \in \mathbb{R}^{t+2^l} \text{ is feasible}\}.$$

Proposition 23. $\mathcal{C}_\Gamma^{p,\delta}([\Delta, v])$ *has an optimal point for all $\delta \in (0, mc_{[\Gamma, v_{|\Gamma}]}(C))$, provided that $mc_{[\Gamma, v_{|\Gamma}]}(C) > 0$.*

Proof. First of all notice that, since $mc_{[\Gamma, v_{|\Gamma}]}(C) > 0$, the feasible set of $\mathcal{C}_\Gamma^{p,\delta}([\Delta, v])$ will not be empty. The existence of $\mathcal{SC}_\Gamma^{p,\delta}([\Delta, v])$ follows from the fact that the feasible set is a closed set. □

Notice that the vector $\vec{x}$ in an optimal point $(\vec{x}, \vec{y}) \in \mathbb{R}^{2k+2^l}$ of $\mathcal{C}_\Gamma^{p,\delta}([\Delta, v])$ is a repair of $[\Delta, v]$ (i.e., $\vec{x} \in R_\Gamma([\Delta, v])$) although, of course, not necessarily p-optimal.

The next propositions state the relation between $DC_\Gamma^p([\Delta, v])$ and $\mathcal{SC}_\Gamma^{p,\delta}([\Delta, v])$.

Proposition 24. *$DC^p_\Gamma([\Delta, v]) = \frac{1}{\epsilon}$ if and only if, for all $\delta \in (0, 1]$, $\mathcal{SC}^{p,\delta}_\Gamma([\Delta, v])$ does not exist.*

Proof. It follows trivially from the definition of DC^p. □

For the next proposition let us assume that $mc_{[\Gamma, v_{|\Gamma}]}(C) > 0$. Consider $\{\delta_n\}_{n\in\mathbb{N}}$ a decreasing sequence that converges to 0, with $\delta_n \in (0, mc_{[\Gamma, v_{|\Gamma}]}(C))$ for all $n \in \mathbb{N}$, and

$$\ell = \lim_{n\to\infty} \mathcal{SC}^{p,\delta_n}_\Gamma([\Delta, v]).$$

Proposition 25. *If $|\mathcal{SC}^{p,\delta_n}_\Gamma([\Delta, v]) - \ell| > 0$ for all $n \in \mathbb{N}$ then $DC^p_\Gamma([\Delta, v]) = \ell + \epsilon$, otherwise $DC^p_\Gamma([\Delta, v]) = \ell$.*

Proof. It follows trivially from the definition of the map DC^p. □

Corollary 26. *If $DC^p_\Gamma([\Delta, v]) = \ell \in \mathbb{R}$ then there exists $N \in \mathbb{N}$ such that, for all $n, m \geq N$, $|\mathcal{SC}^{p,\delta_n}_\Gamma([\Delta, v]) - \mathcal{SC}^{p,\delta_m}_\Gamma([\Delta, v])| = 0$.*

Notice that the *constraint functions* of the form

$$(\underline{\eta}_i - x_i)\vec{r}_{\theta_i} \cdot \vec{y}, \qquad (\overline{\eta}_i + x_{i+k})\vec{r}_{\theta_i} \cdot \vec{y}$$

in (1) for each $i \in \{1, \ldots, k\}$ are not linear. That places our problem $\mathcal{C}^{p,\delta}_\Gamma([\Delta, v])$ within *non-convex* optimization grounds and, unfortunately, there seem to be no *effective* methods to obtain $\mathcal{SC}^{p,\delta}_\Gamma([\Delta, v])$ for the general case (see for example [4] or [5] for more on these issues). Nevertheless, there exist methods to deal with at least small instances of these problems. Such methods normally take some compromise that could consist, for example, in a simplification of the problem by means of a modification in the constraints and/or the obtention of an approximation to the solution at best.

6. Measuring inconsistency in Φ_{CB}

In the companion paper [12] it was shown that Φ_{CB} is unsatisfiable. A study of the sources of inconsistency was also carried out and we found that, for a slightly relaxed version of Φ_{CB} given by replacing every rule of the form $\langle p|q, \eta\rangle$ in it with $\eta \in (0, 1)$ for $\langle p|q, [\eta - 0.01, \eta + 0.01]\rangle$[7] the number of minimal unsatisfiable subsets is 695. These subsets happen to be also minimal unsatisfiable subsets of Φ_{CB} itself (i.e., under the natural, point-valued interpretation for the rules). We actually know the number of minimal unsatisfiable subsets of Φ_{CB} itself is much larger yet not all of them have been computed so far (we refer to [12] for more on these issues).

On a straightforward reading, one could claim that such a large number of minimal unsatisfiable subsets makes Φ_{CB} *highly* inconsistent as each inconsistent subset is an inconsistency. Our notion of p-distance constitutes a *graded* approach to measuring inconsistency in probabilistic knowledge bases and gives a very different reading as to how inconsistent Φ_{CB} is.

6.1. Measuring $DC^p(\Phi_{CB})$

The computation of DC^p for a general knowledge base $[\Delta, v] \in \mathcal{PK}$ is, as we have seen in the previous section, not a trivial matter. In particular, in order to compute $DC^p(\Phi_{CB})$, we will use some heuristics and will take advantage of its *simple* structure, as we will shortly show.

Let $[\Delta, v] \in \mathcal{PK}$ be such that, for all $\phi|\theta \in \Delta$, we have that $\phi, \theta \in L$. We can regard $[\Delta, v]$ as a directed graph where the edges and nodes are given by the conditional statements in Δ (i.e., the conditional statement $\phi|\theta$ yields the edge directed from the node θ to the node ϕ).

For what follows let us consider the collection of real intervals $\mathfrak{I}^r \subset \mathfrak{I}$ that differs from $\mathfrak{I}$ in that the end points of its members are rational (i.e., for $[\eta_1, \eta_2] \in \mathfrak{I}^r$ we will have that $\eta_1, \eta_2 \in \mathbb{Q}$).

For the next lemma let $[\Delta, v] \in \mathcal{PK}^r$, where

$$\mathcal{PK}^r = \{[\Delta, v] \in \mathcal{PK} \mid v(\phi|\theta) \in \mathfrak{I}^r \text{ for all } \phi|\theta \in \Delta\}.$$

Lemma 27. *$[\Delta, v]$ is satisfiable if and only if there exists a model $\mathcal{I}$ of $[\Delta, v]$.*

Proof. That there exists a model $\mathcal{I}$ of $[\Delta, v]$ if and only if $[\Delta, v]$ is satisfied by a rational probability function is clear. We need to prove that if a probability function satisfies $[\Delta, v]$ then there exists a rational probability function that satisfies $[\Delta, v]$.

[7] All the degrees of confirmation in the rules of Φ_{CB} contain at most two decimal digits and so these intervals are well defined.

Let us consider the following linear system with variables the 2^l coordinates of the vector $\vec{y}$ given by the collection of inequalities of the form

$$\underline{\eta}\vec{r}_\theta \cdot \vec{y} \le \vec{r}_{\phi\wedge\theta} \cdot \vec{y} \le \overline{\eta}\vec{r}_\theta \cdot \vec{y},$$

for all $\phi|\theta \in \Delta$, with $v(\phi|\theta) = [\underline{\eta}, \overline{\eta}] \in \mathfrak{I}^r$.

Along with these inequalities we also consider the constraints $\vec{y} \in \mathbb{D}_{2^l}$ (i.e., $\vec{y}$ is a probability function on L) and $\vec{r}_\theta \cdot \vec{y} > 0$ for each $\phi|\theta \in \Delta$.

We assume that this linear system has a solution (i.e., that there exists a probability function that satisfies $[\Delta, v]$) and thus its solution set is non-empty. Notice that the set of solutions of the system needs to contain rational solutions given the nature of its linear constraints and the form of the intervals in $\mathfrak{I}^r$. Thus we can conclude that if a probability function satisfies $[\Delta, v]$ then there has to exist a rational probability function that satisfies $[\Delta, v]$. $\square$

For the next proposition let $[\Delta, v] \in \mathcal{PK}$ be as follows:

- $\phi, \theta \in L$ for all $\phi|\theta \in \Delta$.
- $[\Delta, v]$ is, as a directed graph, acyclic.
- $v(\phi|\theta) \in (0, 1) \cap \mathbb{Q}$, for all $\phi|\theta \in \Delta$.

Proposition 28. $[\Delta, v]$ *is satisfiable.*

Proof. By Lemma 27 the pair $[\Delta, v]$ will be satisfiable if and only if it has a model. We aim at constructing a model of $[\Delta, v]$.

Let $C = \{v(\phi|\theta) \mid (\phi|\theta) \in \Delta\}$. By assumption on $[\Delta, v]$, C will be a set of rational values. Let Q be the least commom multiple of the denominators of all the values in C.

For the construction of $\mathcal{I}$ we consider first the set of *nodes* in $[\Delta, v]$ that have no *parents*, when dealing with $[\Delta, v]$ as a graph. Let p be a node in $[\Delta, v]$ with no parents. We set a collection of Q elements for p, labelled $D_p = \{a_1^p, \ldots, a_Q^p\}$. At the n^{th} step in the construction of $\mathcal{I}$ we select a node q in $[\Delta, v]$, with parents $q_1, \ldots, q_k$ (for some $k \in \mathbb{N}$) and assume, without loss of generality, that the sets $D_{q_1}, \ldots, D_{q_k}$ have been defined at a previous step. We set a collection of $Q(\sum_{i=1}^k |D_{q_i}|)$ elements, labelled $D_q = \{a_1^q, \ldots, a^q_{Q(\sum_{i=1}^k |D_{q_i}|)}\}$, for q.

We impose some restrictions on the sets of the form D_p, for $p \in L$. We construct $\mathcal{I}$ in a way that, for $p, q \in L$, $D_p \cap D_q \neq \emptyset$ if and only if one of $p|q$, $q|p$ is in Δ or there exists $r \in L$ such that $p|r, q|r \in \Delta$.[8] That this can be achieved is clear given the restrictions on the cardinality of such sets and the fact that $v(\phi|\theta)$ is at most $\frac{Q-1}{Q}$ for all $\phi|\theta \in \Delta$.

For $p \in L$, with parents $q_1, \ldots, q_k$ (for some $k \in \mathbb{N}$) and sets $D_{q_1}, \ldots, D_{q_k}$ already defined, we will set D_p to have exactly $|D_{q_i}|v(p|q_i)$ elements of D_{q_i}, for each $i \in \{1, \ldots, k\}$. The elements of D_{q_i} chosen for D_p could also be elements of D_{q_j}, for $j \in \{1, \ldots, i-1, i+1, \ldots, k\}$, (depending on whether q_i and q_j are related in Δ by a conditional statement or not) but of no other sets thus far defined. That a suitable definition of D_p can be achieved (considering that the sets $D_{q_1}, \ldots, D_{q_k}$ might not be pairwise disjoint) is clear given the restrictions on the cardinality of such sets and the fact that $\frac{1}{Q} \le v(\phi|\theta) \le \frac{Q-1}{Q}$ for all $\phi|\theta \in \Delta$.

We define

$$D^{\mathcal{I}} = \bigcup_{p \in L} D_p$$

and $v^{\mathcal{I}}$ on $L \times D^{\mathcal{I}}$ as expected. For $p \in L$ and $a \in D^{\mathcal{I}}$,

$$v^{\mathcal{I}}(p, a) = \begin{cases} 1 & \text{if } a \in D_p, \\ 0, & \text{otherwise.} \end{cases} \quad \square$$

Proposition 29. $DC^p(\Phi_{CB}) = \epsilon$ *for all* $p \in [1, \infty]$.

Proof. We know from results in [12] that Φ_{CB} is inconsistent and thus that $DC^p(\Phi_{CB}) > 0$ for all $p \in [1, \infty]$ (i.e., at least ϵ). On the other hand Φ_{CB} is, when regarded as a graph, acyclic and with nodes in L. Replacing 0 and 1 by any values in $(0, 1) \cap \mathbb{Q}$ as close as desired to 0 and 1 in rules of the form $\langle q|p, 0\rangle$, $\langle q|p, 1\rangle \in \Phi_{CB}$ respectively guarantees satisfiability, by Proposition 28. Therefore we can conclude that $DC^p(\Phi_{CB}) = \epsilon$ for all $p \in [1, \infty]$. $\square$

Proposition 29 tells us that the amount of inconsistency in Φ_{CB}, as measured by DC^p, is *infinitesimal* and thus Φ_{CB} is very close to consistency (i.e., *almost* consistent).

[8] In general, we could impose the condition that distinct sets of the form D_p, for $p \in L$, should have as little domain elements in common as possible.

By considering the notion of satisfiability taken in [17] (which differs from that based on Definition 3 in that a probability function ω on L with $\omega(\theta) = 0$ is assumed to *satisfy* any statement of the form $\langle \phi|\theta, \Omega\rangle$, for $\phi|\theta \in \mathcal{K}$ and $\Omega \in \Im$) we would have that Φ_{CB} is satisfiable (for example, one can consider a probability function ω on L such that $\omega(\neg p_1 \wedge \cdots \wedge \neg p_l) = 1$, which satisfies Φ_{CB}). Thus the amount of inconsistency in Φ_{CB} according to the metric presented in [17] would be zero. Certainly, in terms of magnitude, there is not much of a difference between the metric presented in [17] and that given by DC^p when it comes to quantifying the amount of inconsistency of Φ_{CB} (in magnitude, such difference is just *infinitesimal*).

In order to see in more detail the difference between these two approaches (i.e., that in [17] and ours) let us consider a small fragment of Φ_{CB}, given by the following subset of rules:

$$\Phi = \{\langle D21|S1022, 1\rangle, \langle D20|S1022, 0.1\rangle, \langle D20|D21, 1\rangle\},$$

with $D20, D21 \in D$ and $S1022 \in S$.

By considering the notion of satisfiability derived from Definition 3 Φ is unsatisfiable (actually, Φ is a minimal unsatisfiable subset) and $DC^p(\Phi) = \epsilon$ whereas according to the notion of satisfiability in [17] Φ is satisfiable (for example, consider the probability distribution ω on $\{D21, S1022, D20\}$ such that $\omega(\neg D21 \wedge \neg D20 \wedge \neg S1022) = 1$, which would correspond, in set theoretical or first order terms, to the *empty model*). Beyond magnitudes or degrees of inconsistency, it is important to appreciate the *qualitative* difference between these two approaches for our particular example. DC^p acknowledges the unsatisfiability of Φ (and, we believe, modelers would agree in that Φ should be regarded as unsatisfiable and thus as non-desirable) whereas the approach in [17] does not.

Let us assume now that the conditional $D20|D21$ is regarded as well evaluated (i.e., that $\langle D20|D21, 1\rangle$ is regarded by the modeler as a reliable piece of information). We will have in this case that $DC^1_\Gamma(\Phi) = 0.9$ (with normalized measure $\overline{DC}^1_\Gamma(\Phi) = 0.3$), where $\Gamma = \{D20|D21\}$, and thus that there is a considerable difference between our approach and that of [17] also in terms of magnitude in this particular example.

7. Conclusion

We have presented and analyzed a family of measures (p-measures) aimed at helping the modeler in evaluating inconsistency in probabilistic knowledge bases. Unlike most approaches to measuring inconsistency in the literature, ours is graded in the sense that we take into account *minimal* adjustments in the degrees of certainty (i.e., probabilities) that make the knowledge base consistent. We have also seen that the computation of the measures here presented yield possible repairs to bring or restore consistency in such knowledge bases.

There is still much left to be done in relation to the family of measures presented here. In particular, a deeper understanding of the differences among distinct p-measures (and, by extension, among the possible repairs they yield), both on theoretical and practical considerations, would be desirable. This task is left for future work and, hopefully, will constitute the core of a future paper.

References

[1] K. Adlassnig, Fuzzy set theory in medical diagnosis, IEEE Transactions on Systems, Man and Cybernetics 16 (2) (1986) 260–265.
[2] K. Adlassnig, G. Kolarz, W. Effenberger, H. Grabner, CADIAG: approaches to computer assisted medical diagnosis, Computers in Biology and Medicine 15 (1985) 315–335.
[3] K. Adlassnig, G. Kolarz, W. Scheithauer, H. Grabner, Approach to a hospital-based application of a medical expert system, Medical Informatics 11 (3) (1986) 205–223.
[4] D. Bertsekas, Nonlinear Programming, Athena Scientific, 2009.
[5] S. Boyd, L. Vandenberghe, Convex Optimization, Cambridge University Press, 2004.
[6] J.S. Chun Lam, D.H. Sleeman, J.Z. Pan, W. Weber Vasconcelos, A fine-grained approach to resolving unsatisfiable ontologies, Journal of Data Semantics 10 (2010) 62–95.
[7] A. Ciabattoni, T. Vetterlein, On the fuzzy (logical) content of CADIAG-2, Fuzzy Sets and Systems 161 (14) (2010) 1941–1958.
[8] R. Goldblatt, Lectures on the hyperreals, .An Introduction to Nonstandard Analyses, Graduate Texts in Mathematics, vol. 188, Springer, 1998
[9] A. Hunter, S. Konieczny, Measuring inconsistency through minimal inconsistent sets, in: G. Brewka, J. Lang (Eds.), Eleventh International Conference on the Principles of Knowledge Representation and Reasoning, AAAI Press, 2008, pp. 358–366.
[10] A. Hunter, S. Konieczny, Approaches to measuring inconsistent information. Inconsistency tolerance, Lecture Notes in Computer Science, vol. 3300, Springer, 2005, pp. 189–234.
[11] A. Kalyanpur, B. Parsia, E. Sirin, B.C. Grau, Repairing unsatisfiable concepts in OWL ontologies, in: Y. Sure, J. Domingue (Eds.), Third European Semantic Web Conference. The Semantic Web: Research and Applications. Lecture Notes in Computer Science, vol. 4011, Springer, 2006, pp. 170–184.
[12] P. Klinov, B. Parsia, D. Picado Muiño, The consistency of the CADIAG-2 knowledge base: a probabilistic approach, in: C.G. Fermüller, A. Voronkov (Eds.), Logic for Programming, Artificial Intelligence and Reasoning – 17th International Conference, Lecture Notes in Computer Science, vol. 6397, Springer, 2010, pp. 432–446.
[13] K.M. Knight, Measuring inconsistency, Journal of Philosophical Logic 31 (2002) 77–98.
[14] H. Leitich, K. Adlassnig, G. Kolarz, Evaluation of two different models of semiautomatic knowledge acquisition for the medical consultant system CADIAG-2/RHEUMA, Artificial Intelligence in Medicine 25 (2002) 215–225.
[15] D. Picado Muiño, The (probabilistic) logical content of CADIAG-2, in: J. Filipe, A. Fred, B. Sharp (Eds.), Second International Conference on Agents and Artificial Intelligence, Artificial Intelligence, vol. 1, Valencia, Spain, 2010, pp. 28–35.
[16] R. Reiter, A theory of diagnosis from first principles, Artificial Intelligence 32 (1987) 57–95.
[17] M. Thimm, Measuring inconsistency in probabilistic knowledge bases, Twenty-Fifth Conference on Uncertainty in Artificial Intelligence, AUAI Press, Corvallis, Oregon, 2009, pp. 530–537.
[18] L. Zadeh, Fuzzy logic and approximate reasoning, Synthese 30 (1975) 407–428.