Implications over Probabilistic Attributes

Francesco Kriegel

Institute of Theoretical Computer Science,
Technische Universität Dresden, Dresden, Germany
francesco.kriegel@tu-dresden.de
orcid.org/0000-0003-0219-0330

Abstract. We consider the task of acquisition of terminological knowledge from given assertional data. However, when evaluating data of real-world applications we often encounter situations where it is impractical to deduce only crisp knowledge, due to the presence of exceptions or errors. It is rather appropriate to allow for degrees of uncertainty within the derived knowledge. Consequently, suitable methods for knowledge acquisition in a probabilistic framework should be developed.
In particular, we consider data which is given as a probabilistic formal context, i.e., as a triadic incidence relation between objects, attributes, and worlds, which is furthermore equipped with a probability measure on the set of worlds. We define the notion of a probabilistic attribute as a probabilistically quantified set of attributes, and define the notion of validity of implications over probabilistic attributes in a probabilistic formal context. Finally, a technique for the axiomatization of such implications from probabilistic formal contexts is developed. This is done is a sound and complete manner, i.e., all derived implications are valid, and all valid implications are deducible from the derived implications. In case of finiteness of the input data to be analyzed, the constructed axiomatization is finite, too, and can be computed in finite time.

Keywords: Knowledge Acquisition · Probabilistic Formal Context · Probabilistic Attribute · Probabilistic Implication · Knowledge Base

1 Introduction

We consider data which is given as a probabilistic formal context, i.e., as a triadic incidence relation between objects, attributes, and worlds, which is furthermore equipped with a probability measure on the set of worlds. We define the notion of a probabilistic attribute as a probabilistically quantified set of attributes, and define the notion of validity of implications over probabilistic attributes in a probabilistic formal context. Finally, a technique for the axiomatization of such implications from probabilistic formal contexts is developed. This is done is a sound and complete manner, i.e., all derived implications are valid, and all valid implications are deducible from the derived implications. In case of finiteness of the input data to be analyzed, the constructed axiomatization is finite, too, and can be computed in finite time.

This document is structured as follows. A brief introduction on *Formal Concept Analysis* is given in Section 2, and then the subsequent Section 3 presents basics on *Probabilistic Formal Concept Analysis*. Then, in Section 4 we define the notion of implications over probabilistic attributes, and infer some characterizing statements. The most important part of this document is the Section 5, in which we constructively develop a method for the axiomatization of probabilistic implications from probabilistic formal contexts; the section closes with a proof of soundness and completeness of the proposed knowledge base. Eventually, in Section 6 some closing remarks as well as future steps for extending and applying the results are given. In order to explain and motivate the definitions and the theoretical results, Sections 3 – 5 contain a running example.

2 Formal Concept Analysis

This section briefly introduces the standard notions of *Formal Concept Analysis* (abbr. *FCA*) [5]. A *formal context* $\mathbb{K} := (G, M, I)$ consists of a set G of *objects*, a set M of *attributes*, and an *incidence relation* $I \subseteq G \times M$. For a pair $(g, m) \in I$, we say that g *has* m. The *derivation operators* of $\mathbb{K}$ are the mappings $\cdot^I\colon \wp(G) \to \wp(M)$ and $\cdot^I\colon \wp(M) \to \wp(G)$ that are defined by

$$
\begin{aligned}
& A^I := \{\, m \in M \mid \forall\, g \in A\colon (g,m) \in I \,\} \quad \text{for object sets } A \subseteq G, \\
\text{and}\quad & B^I := \{\, g \in G \mid \forall\, m \in B\colon (g,m) \in I \,\} \quad \text{for attribute sets } B \subseteq M.
\end{aligned}
$$

It is well-known [5] that both derivation operators constitute a so-called *Galois connection* between the powersets $\wp(G)$ and $\wp(M)$, i.e., the following statements hold true for all subsets $A, A_1, A_2 \subseteq G$ and $B, B_1, B_2 \subseteq M$:

1. $A \subseteq B^I \Leftrightarrow B \subseteq A^I \Leftrightarrow A \times B \subseteq I$
2. $A \subseteq A^{II}$
3. $A^I = A^{III}$
4. $A_1 \subseteq A_2 \Rightarrow A_2^I \subseteq A_1^I$
5. $B \subseteq B^{II}$
6. $B^I = B^{III}$
7. $B_1 \subseteq B_2 \Rightarrow B_2^I \subseteq B_1^I$

For obvious reasons, formal contexts can be represented as binary tables the rows of which are labeled with the objects, the columns of which are labeled with the attributes, and the occurrence of a cross $\times$ in the cell at row g and column m indicates that the object g has the attribute m.

An *intent* of $\mathbb{K}$ is an attribute set $B \subseteq M$ with $B = B^{II}$. The set of all intents of $\mathbb{K}$ is denoted by $\mathsf{Int}(\mathbb{K})$. An *implication* over M is an expression $X \to Y$ where $X, Y \subseteq M$. It is *valid* in $\mathbb{K}$, denoted as $\mathbb{K} \models X \to Y$, if $X^I \subseteq Y^I$, i.e., if each object of $\mathbb{K}$ that possesses all attributes in X also has all attributes in Y. An implication set $\mathcal{L}$ is *valid* in $\mathbb{K}$, denoted as $\mathbb{K} \models \mathcal{L}$, if all implications in $\mathcal{L}$ are valid in $\mathbb{K}$. Furthermore, the relation $\models$ is lifted to implication sets as follows: an implication set $\mathcal{L}$ *entails* an implication $X \to Y$, symbolized as $\mathcal{L} \models X \to Y$, if $X \to Y$ is valid in all formal contexts in which $\mathcal{L}$ is valid. More specifically, $\models$ is called the *semantic entailment relation*.

It was shown that entailment can also be decided *syntactically* by applying *deduction rules* to the implication set $\mathcal{L}$ without the requirement to consider all formal contexts in which $\mathcal{L}$ is valid. Recall that an implication $X \to Y$ is *syntactically entailed* by an implication set $\mathcal{L}$, denoted as $\mathcal{L} \vdash X \to Y$, if $X \to Y$ can be constructed from $\mathcal{L}$ by the application of *inference axioms*, cf. [12, Page 47], that are described as follows:

(F1)	*Reflexivity:*	$\emptyset \vdash X \to X$
(F2)	*Augmentation:*	$\{X \to Y\} \vdash X \cup Z \to Y$
(F3)	*Additivity:*	$\{X \to Y, X \to Z\} \vdash X \to Y \cup Z$
(F4)	*Projectivity:*	$\{X \to Y \cup Z\} \vdash X \to Y$
(F5)	*Transitivity:*	$\{X \to Y, Y \to Z\} \vdash X \to Z$
(F6)	*Pseudotransitivity:*	$\{X \to Y, Y \cup Z \to W\} \vdash X \cup Z \to W$

In the inference axioms above the symbols X, Y, Z, and W, denote arbitrary subsets of the considered set M of attributes. Formally, we define $\mathcal{L} \vdash X \to Y$ if there is a sequence of implications $X_0 \to Y_0, \ldots, X_n \to Y_n$ such that the following conditions hold:

1. For each $i \in \{0, ..., n\}$, there is a subset $\mathcal{L}_i \subseteq \mathcal{L} \cup \{X_0 \to Y_0, \ldots, X_{i-1} \to Y_{i-1}\}$ such that $\mathcal{L}_i \vdash X_i \to Y_i$ matches one of the Axioms F1 − F6.
2. $X_n \to Y_n = X \to Y$.

Often, the Axioms F1, F2, and F6, are referred to as *Armstrong's axioms*. These three axioms constitute a *complete* and *independent* set of inference axioms for implicational entailment, i.e., from it the other Axioms F3 − F5 can be derived, and none of them is derivable from the others.

The semantic entailment and the syntactic entailment coincide, i.e., an implication $X \to Y$ is semantically entailed by an implication set $\mathcal{L}$ if, and only if, $\mathcal{L}$ syntactically entails $X \to Y$, cf. [12, Theorem 4.1 on Page 50] as well as [5, Proposition 21 on Page 81]. Consequently, we do not have distinguish between both entailment relations $\models$ and $\vdash$, when it is up to decide whether an implication follows from a set of implications.

A *model* of an implication set $\mathcal{L}$ is an attribute set $Z \subseteq M$ such that $X \subseteq Z$ implies $Y \subseteq Z$ for all $X \to Y \in \mathcal{L}$. By $X^{\mathcal{L}}$ we denote the smallest superset of X that is a model of $\mathcal{L}$.

The data encoded in a formal context can be visualized as a *line diagram* of the corresponding *concept lattice*, which we shall shortly describe. A *formal concept* of a formal context $\mathbb{K} := (G, M, I)$ is a pair (A, B) consisting of a set $A \subseteq G$ of objects as well as a set $B \subseteq M$ of attributes such that $A^I = B$ and $B^I = A$. We then also refer to A as the *extent*, and to B as the *intent*, respectively, of (A, B). In the denotation of $\mathbb{K}$ as a cross table, those formal concepts are the maximal rectangles full of crosses (modulo reordering of rows and columns). Then, the set of all formal concepts of $\mathbb{K}$ is denoted as $\mathfrak{B}(\mathbb{K})$, and it is ordered by defining $(A, B) \leq (C, D)$ if, and only if, $A \subseteq C$. It was shown that this order always induces a complete lattice $\boldsymbol{\mathfrak{B}}(\mathbb{K}) := (\mathfrak{B}(\mathbb{K}), \leq, \wedge, \vee, \top, \bot)$, called the *concept lattice* of $\mathbb{K}$, cf. [5, 14], in which the infimum and the supremum operation satisfy the equations

$$\begin{aligned} &\bigwedge_{t\in T}(A_t,B_t) = \big(\bigcap_{t\in T} A_t, (\bigcup_{t\in T} B_t)^{II}\big), \\ \text{and} \quad &\bigvee_{t\in T}(A_t,B_t) = \big((\bigcup_{t\in T} A_t)^{II}, \bigcap_{t\in T} B_t\big), \end{aligned}$$

and where $\top = (\emptyset^I, \emptyset^{II})$ is the greatest element, and where $\bot = (\emptyset^{II}, \emptyset^I)$ is the smallest element, respectively. Furthermore, the concept lattice of $\mathbb{K}$ can be nicely represented as a *line diagram* as follows: Each formal concept is depicted as a vertex. Furthermore, there is an upward directed edge from each formal concept to its upper neighbors, i.e., to all those formal concepts which are greater with respect to $\leq$, but for which there is no other formal concept in between. The nodes are labeled as follows: an attribute $m \in M$ is an upper label of the *attribute concept* $(\{m\}^I, \{m\}^{II})$, and an object $g \in G$ is a lower label of the *object concept* $(\{g\}^{II}, \{g\}^I)$. Then, the extent of the formal concept represented by a vertex consists of all objects which label vertices reachable by a downward directed path, and dually the intent is obtained by gathering all attribute labels of vertices reachable by an upward directed path.

Let $\mathbb{K} \models \mathcal{L}$. A *pseudo-intent* of a formal context $\mathbb{K}$ relative to an implication set $\mathcal{L}$ is an attribute set $P \subseteq M$ which is no intent of $\mathbb{K}$, but is a model of $\mathcal{L}$, and satisfies $Q^{II} \subseteq P$ for all pseudo-intents $Q \subsetneq P$. The set of all those pseudo-intents is symbolized by $\mathsf{PsInt}(\mathbb{K}, \mathcal{L})$. Then the implication set

$$\mathsf{Can}(\mathbb{K}, \mathcal{L}) := \{\, P \to P^{II} \mid P \in \mathsf{PsInt}(\mathbb{K}, \mathcal{L}) \,\}$$

constitutes an *implicational base* of $\mathbb{K}$ relative to $\mathcal{L}$, i.e., for each implication $X \to Y$ over M, the following equivalence is satisfied:

$$\mathbb{K} \models X \to Y \Leftrightarrow \mathsf{Can}(\mathbb{K}, \mathcal{L}) \cup \mathcal{L} \models X \to Y.$$

$\mathsf{Can}(\mathbb{K}, \mathcal{L})$ is called the *canonical base* of $\mathbb{K}$ relative to $\mathcal{L}$. It can be shown that it is a *minimal* implicational base of $\mathbb{K}$ relative to $\mathcal{L}$, i.e., there is no implicational base of $\mathbb{K}$ relative to $\mathcal{L}$ with smaller cardinality. Further information is given by [3, 4, 6, 13]. The most prominent algorithm for computing the canonical base is certainly *NextClosure* developed by Bernhard Ganter [3, 4]. A parallel algorithm called *NextClosures* is also available [8, 11], and an implementation is provided in *Concept Explorer FX* [7]; its advantage is that its processing time scales almost inverse linear with respect to the number of available CPU cores.

Eventually, in case a given formal context is not complete in the sense that it does not contain enough objects to refute invalid implications, i.e., only contains some observed objects in the domain of interest, but one aims at exploring all valid implications over the given attribute set, a technique called *Attribute Exploration* can be utilized, which guides the user through the process of axiomatizing an implicational base for the underlying domain in a way the number of questions posed to the user is minimal. For a sophisticated introduction as well as for theoretical and technical details, the interested reader is rather referred to [2–4, 9, 13]. A parallel variant of the *Attribute Exploration* also exists, cf. [8, 9], which is implemented in *Concept Explorer FX* [7].

3 Probabilistic Formal Concept Analysis

This section presents probabilistic extensions of the common notions of *Formal Concept Analysis*, which were first introduced in [10]. A *probability measure* $\mathbb{P}$ on a countable set W is a mapping $\mathbb{P}\colon \wp(W) \to [0,1]$ such that $\mathbb{P}(\emptyset) = 0$, $\mathbb{P}(W) = 1$, and $\mathbb{P}$ is σ-*additive*, i.e., for all countable families $(U_n)_{n\in\mathbb{N}}$ of pairwise disjoint sets $U_n \subseteq W$ it holds that $\mathbb{P}(\bigcup_{n\in\mathbb{N}} U_n) = \sum_{n\in\mathbb{N}} \mathbb{P}(U_n)$. A world $w \in W$ is *possible* if $\mathbb{P}\{w\} > 0$, and *impossible* otherwise. The set of all possible worlds is denoted by W_ε, and the set of all impossible worlds is denoted by W_0. Obviously, $W_\varepsilon \uplus W_0$ is a partition of W. Of course, such a probability measure can be completely characterized by the definition of the probabilities of the singleton subsets of W, since it holds true that $\mathbb{P}(U) = \mathbb{P}(\bigcup_{w\in U}\{w\}) = \sum_{w\in U} \mathbb{P}(\{w\})$.

Definition 1. *A* probabilistic formal context $\mathbb{K}$ *is a tuple* $(G, M, W, I, \mathbb{P})$ *that consists of a set* G *of* objects, *a set* M *of* attributes, *a countable set* W *of* worlds, *an* incidence relation $I \subseteq G \times M \times W$, *and a probability measure* $\mathbb{P}$ *on* W. *For a triple* $(g, m, w) \in I$ *we say that* object g has attribute m in world w. *Furthermore, we define the* derivations in world w *as operators* $\cdot^{I(w)}\colon \wp(G) \to \wp(M)$ *and* $\cdot^{I(w)}\colon \wp(M) \to \wp(G)$ *where*

$$A^{I(w)} := \{\, m \in M \mid \forall\, g \in A\colon (g, m, w) \in I \,\} \quad \textit{for object sets } A \subseteq G,$$
$$\textit{and} \quad B^{I(w)} := \{\, g \in G \mid \forall\, m \in B\colon (g, m, w) \in I \,\} \quad \textit{for attribute sets } B \subseteq M,$$

i.e., $A^{I(w)}$ *is the set of all common attributes of all objects in* A *in the world* w, *and* $B^{I(w)}$ *is the set of all objects that have all attributes in* B *in* w. *The formal context induced by a world* $w \in W$ *is defined as* $\mathbb{K}(w) := (G, M, I(w))$.

$\mathbb{K}_{\mathsf{ex}}(w_1)$	m_1	m_2	m_3
g_1	×	·	×
g_2	·	×	×
g_3	·	·	×

$\mathbb{P}_{\mathsf{ex}}(w_1) := \frac{1}{2}$

$\mathbb{K}_{\mathsf{ex}}(w_2)$	m_1	m_2	m_3
g_1	×	·	×
g_2	·	×	·
g_3	×	·	×

$\mathbb{P}_{\mathsf{ex}}(w_2) := \frac{1}{3}$

$\mathbb{K}_{\mathsf{ex}}(w_3)$	m_1	m_2	m_3
g_1	×	·	×
g_2	·	×	×
g_3	·	×	·

$\mathbb{P}_{\mathsf{ex}}(w_3) := \frac{1}{6}$

Fig. 1. An exemplary probabilistic formal context $\mathbb{K}_{\mathsf{ex}}$

As a running example for the current and the up-coming sections, we consider the probabilistic formal context $\mathbb{K}_{\mathsf{ex}}$ presented in Figure 1. It consists of three objects g_1, g_2, g_3, three attributes m_1, m_2, m_3, and three worlds w_1, w_2, w_3. In $\mathbb{K}_{\mathsf{ex}}$ it holds true that the object g_1 has the attribute m_1 in all three worlds, and the object g_3 has the attribute m_3 in all worlds except in w_3.

Definition 2. *Let $\mathbb{K}$ be a probabilistic formal context. The* almost certain scaling *of $\mathbb{K}$ is the formal context $\mathbb{K}_\varepsilon^\times := (G \times W_\varepsilon, M, I_\varepsilon^\times)$ where $((g,w),m) \in I^\times$ if $(g,m,w) \in I$.*

$(\mathbb{K}_{\mathsf{ex}})_\varepsilon^\times$	m_1	m_2	m_3
(g_1, w_1)	×	·	×
(g_2, w_1)	·	×	×
(g_3, w_1)	·	·	×
(g_1, w_2)	×	·	×
(g_2, w_2)	·	×	·
(g_3, w_2)	×	·	×
(g_1, w_3)	×	·	×
(g_2, w_3)	·	×	×
(g_3, w_3)	·	×	·

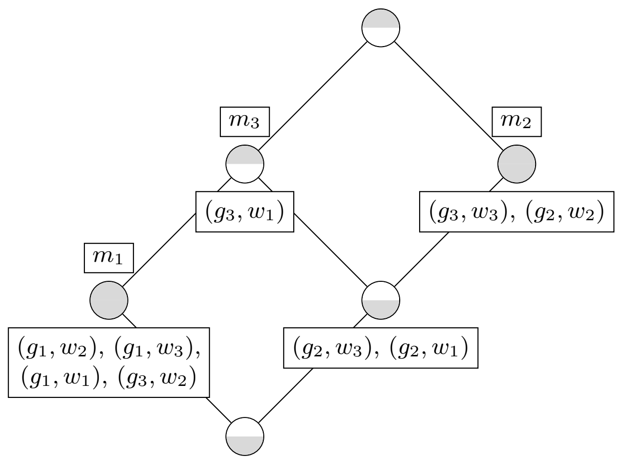

Fig. 2. The certain scaling of $\mathbb{K}_{\mathsf{ex}}$ from Figure 1 and its concept lattice

For our running example $\mathbb{K}_{\mathsf{ex}}$, the almost certain scaling is displayed in Figure 2. As it can be read off it essentially consists of the subposition of the three formal contexts $\mathbb{K}_{\mathsf{ex}}(w_1)$, $\mathbb{K}_{\mathsf{ex}}(w_2)$, and $\mathbb{K}_{\mathsf{ex}}(w_3)$.

Lemma 3. *Let $\mathbb{K} := (G, M, W, I, \mathbb{P})$ be a probabilistic formal context. Then for all subsets $B \subseteq M$ and for all possible worlds $w \in W_\varepsilon$, it holds true that $B^{I(w)} = B^{I_\varepsilon^\times I_\varepsilon^\times I(w)}$.*

Proof. Let $X \subseteq M$. Then for all possible worlds $w \in W_\varepsilon$ it holds that

$$\begin{aligned} g \in X^{I(w)} &\Leftrightarrow \forall\, m \in X\colon (g,m,w) \in I \\ &\Leftrightarrow \forall\, m \in X\colon ((g,w),m) \in I_\varepsilon^\times \Leftrightarrow (g,w) \in X^{I_\varepsilon^\times}, \end{aligned}$$

and we conclude that $X^{I(w)} = \pi_1(X^{I_\varepsilon^\times} \cap (G \times \{w\}))$. Furthermore, we then infer $X^{I(w)} = X^{I_\varepsilon^\times I_\varepsilon^\times I(w)}$. $\square$

4 Implications over Probabilistic Attributes

In [10], the notion of probability of an implication in a probabilistic formal context has been defined. However, it was not possible to express implications between probabilistically quantified attributes, e.g., we could not state that having attribute m with probability $\frac{1}{3}$ implies having attribute n with probability $\frac{2}{3}$. In this section we will resolve this issue by defining the notion of probabilistic attributes, and considering implications over probabilistic attributes. Furthermore, a technique for the construction of bases of such probabilistic implications is proposed in the next Section 5.

Definition 4. *Let $\mathbb{K} := (G, M, W, I, \mathbb{P})$ be a probabilistic formal context. For object sets $A \subseteq G$ and attribute sets $B \subseteq M$, the* incidence probability *is defined as*

$$\mathbb{P}(A, B) := \mathbb{P}\{\, w \in W \mid A \times B \times \{w\} \subseteq I \,\}.$$

A probabilistic attribute *over M is an expression $\mathsf{d} \geq p.\, B$ where $B \subseteq M$, and $p \in [0, 1]$. The set of all probabilistic attributes over M is denoted as $\mathsf{d}(M)$. For a subset $\mathbf{X} \subseteq \mathsf{d}(M)$, its* extension *in $\mathbb{K}$ is given by*

$$\mathbf{X}^I := \{\, g \in G \mid \forall\ \mathsf{d} \geq p.\, B \in \mathbf{X} \colon \mathbb{P}(\{g\}, B) \geq p \,\}.$$

Considering our exemplary probabilistic formal context $\mathbb{K}_{\mathsf{ex}}$ from Figure 1, the set $\{\mathsf{d} \geq \frac{1}{2}.\{m_1\}\}$ has the following extension in $\mathbb{K}_{\mathsf{ex}}$:

$$\{\mathsf{d} \geq \tfrac{1}{2}.\{m_1\}\}^I = \{\, g \in G \mid \mathbb{P}(\{g\}, \{m_1\}) \geq \tfrac{1}{2} \,\} = \{g_1\},$$

i.e., only the object g_1 has the attribute m_1 with a probability of at least $\frac{1}{2}$.

Lemma 5. *Let $\mathbb{K}$ be a probabilistic formal context, and $A \subseteq G$ as well as $B \subseteq M$. Then it holds true that $\mathbb{P}(A, B) = \mathbb{P}(A, B^{I_\varepsilon^\times I_\varepsilon^\times})$.*

Proof. In Lemma 3 we have shown that $B^{I(w)} = B^{I_\varepsilon^\times I_\varepsilon^\times I(w)}$ for all $B \subseteq M$ and all possible worlds $w \in W_\varepsilon$. Hence, we may conclude that

$$A \times B \times \{w\} \subseteq I \Leftrightarrow A \subseteq B^{I(w)} \Leftrightarrow A \subseteq B^{I_\varepsilon^\times I_\varepsilon^\times I(w)} \Leftrightarrow A \times B^{I_\varepsilon^\times I_\varepsilon^\times} \times \{w\} \subseteq I,$$

and so $\mathbb{P}(A, B) = \mathbb{P}(A, B^{I_\varepsilon^\times I_\varepsilon^\times})$. □

Definition 6. *Let $\mathbb{K} := (G, M, W, I, \mathbb{P})$ be a probabilistic formal context.*

1. *A* probabilistic implication *over M is an implication over $\mathsf{d}(M)$, and the set of all probabilistic implications over M is denoted as $\mathsf{dImp}(M)$. A probabilistic implication $\mathbf{X} \to \mathbf{Y}$ is* valid *in $\mathbb{K}$ if $\mathbf{X}^I \subseteq \mathbf{Y}^I$ is satisfied, and we shall denote this by $\mathbb{K} \models \mathbf{X} \to \mathbf{Y}$.*
2. *An implication $P \to Q$ over $[0, 1]$ is* valid *in $\mathbb{K}$, denoted as $\mathbb{K} \models P \to Q$, if for all objects $g \in G$ and all attribute sets $B \subseteq M$, the following condition is satisfied:*

$$\mathbb{P}(\{g\}, B) \in P \text{ implies } \mathbb{P}(\{g\}, B) \in Q.$$

3. *An implication* $X \to Y$ *over* M *is* valid *in* $\mathbb{K}$, *and we symbolize this by* $\mathbb{K} \models X \to Y$, *if for all objects* $g \in G$ *and all probability values* $p \in [0,1]$, *the following condition is satisfied:*

$$\mathbb{P}(\{g\}, X) \geq p \textit{ implies } \mathbb{P}(\{g\}, Y) \geq p.$$

An example for a probabilistic implication within the domain of the probabilistic formal context presented in Figure 1 is

$$\{\mathsf{d} \geq \tfrac{1}{2}.\{m_1\}\} \to \{\mathsf{d} \geq \tfrac{2}{3}.\{m_2\}, \mathsf{d} \geq \tfrac{4}{5}.\{m_3\}\}.$$

However, it is not valid in $\mathbb{K}$ since the premise's extension $\{g_1\}$ of $\{\mathsf{d} \geq \frac{1}{2}.\{m_1\}\}$ is not a subset of the conclusion's extension

$$\begin{aligned}\{\mathsf{d} \geq \tfrac{2}{3}.\{m_2\}, \mathsf{d} \geq \tfrac{4}{5}.\{m_3\}\}^I &= \{\mathsf{d} \geq \tfrac{2}{3}.\{m_2\}\}^I \cap \{\mathsf{d} \geq \tfrac{4}{5}.\{m_3\}\}^I \\ &= \{g_2\} \cap \{g_1, g_3\} = \emptyset.\end{aligned}$$

It can be easily verified that the probabilistic implication $\{\mathsf{d} \geq \frac{1}{2}.\{m_1\}\} \to \{\mathsf{d} \geq \frac{4}{5}.\{m_3\}\}$ is valid in $\mathbb{K}$.

Lemma 7. *Let* $\mathbb{K} \coloneqq (G, M, W, I, \mathbb{P})$ *be a probabilistic formal context in which the implication* $(p, q] \to \{q\}$ *over* $[0,1]$ *is valid. Then, for each attribute set* $X \subseteq M$ *and each probability value* $r \in (p, q]$, *the probabilistic implication* $\{\mathsf{d} \geq r.\,X\} \to \{\mathsf{d} \geq q.\,X\}$ *is valid in* $\mathbb{K}$, *too.*

Proof. Assume that $\mathbb{K} \models (p, q] \to \{q\}$. It easily follows from Definition 6 that for each object $g \in G$, $\mathbb{P}(\{g\}, X) \in (p, q]$ implies $\mathbb{P}(\{g\}, X) = q$, and furthermore it is trivial that $\mathbb{P}(\{g\}, X) > q$ implies $\mathbb{P}(\{g\}, X) \geq q$. Hence, the probabilistic implication $\{\mathsf{d} \geq r.\,X\} \to \{\mathsf{d} \geq q.\,X\}$ is valid in $\mathbb{K}$ for each value $r \in (p, q]$. □

Lemma 8. *Let* $\mathbb{K}$ *be a probabilistic formal context, and assume that the implication* $X \to Y \in \mathsf{Imp}(M)$ *is valid in the almost certain scaling* $\mathbb{K}_\varepsilon^\times$. *Then* $X \to Y$ *is valid in* $\mathbb{K}$, *too, and in particular for each probability value* $p \in [0,1]$, *it holds true that* $\mathbb{K} \models \{\mathsf{d} \geq p.\,X\} \to \{\mathsf{d} \geq p.\,Y\}$.

Proof. Since $X \to Y$ is valid in $\mathbb{K}_\varepsilon^\times$, we conclude that $X^{I_\varepsilon^\times} \subseteq Y^{I_\varepsilon^\times}$, or equivalently $Y \subseteq X^{I_\varepsilon^\times I_\varepsilon^\times}$. By Lemma 5 we have that $\mathbb{P}(\{g\}, X) = \mathbb{P}(\{g\}, X^{I_\varepsilon^\times I_\varepsilon^\times})$. As a consequence we get that $\mathbb{P}(\{g\}, X) = \mathbb{P}(\{g\}, X^{I_\varepsilon^\times I_\varepsilon^\times}) \leq \mathbb{P}(\{g\}, Y)$, and thus $\mathbb{P}(\{g\}, X) \geq p$ implies $\mathbb{P}(\{g\}, Y) \geq p$. As an immediate corollary it then follows that the probabilistic implication $\{\mathsf{d} \geq p.\,X\} \to \{\mathsf{d} \geq p.\,Y\}$ is valid in $\mathbb{K}$. □

5 Probabilistic Implicational Knowledge Bases

In this main section, we define the notion of a probabilistic implicational knowledge base, and develop a method for the construction of a sound and complete knowledge base for a given probabilistic formal context. Furthermore, in case of finiteness of the input probabilistic formal context, the construction will yield a knowledge base which is finite as well.

Definition 9. *A* probabilistic implicational knowledge base *over an attribute set M is a triple $(\mathcal{V}, \mathcal{L}, \mathcal{P})$ where $\mathcal{V} \subseteq \mathsf{Imp}([0,1])$, $\mathcal{L} \subseteq \mathsf{Imp}(M)$, and $\mathcal{P} \subseteq \mathsf{dImp}(M)$. It is* valid *in a probabilistic formal context $\mathbb{K} := (G, M, W, I, \mathbb{P})$ if each implication in $\mathcal{V} \cup \mathcal{L} \cup \mathcal{P}$ is valid in $\mathbb{K}$. Then we shall denote this as $\mathbb{K} \models (\mathcal{V}, \mathcal{L}, \mathcal{P})$. The relation $\models$ is lifted as usual, i.e., for two probabilistic implicational knowledge bases $\mathcal{K}_1$ and $\mathcal{K}_2$, we say that $\mathcal{K}_1$* entails *$\mathcal{K}_2$, symbolized as $\mathcal{K}_1 \models \mathcal{K}_2$, if $\mathcal{K}_2$ is valid in all probabilistic formal contexts in which $\mathcal{K}_1$ is valid, i.e., if for all probabilistic formal contexts $\mathbb{K}$, it holds true that $\mathbb{K} \models \mathcal{K}_1$ implies $\mathbb{K} \models \mathcal{K}_2$.*

Furthermore, a probabilistic implicational knowledge base for *a probabilistic formal context $\mathbb{K}$ is a probabilistic implicational knowledge base $\mathcal{K}$ that satisfies the following condition for all probabilistic implications $\mathbf{X} \to \mathbf{Y}$ over M:*

$$\mathbb{K} \models \mathbf{X} \to \mathbf{Y} \textit{ if, and only if, } \mathcal{K} \models \mathbf{X} \to \mathbf{Y}.$$

5.1 Trivial Background Knowledge

Lemma 10. *Let $\mathbb{K} := (G, M, W, I, \mathbb{P})$ be a probabilistic formal context. Then the following probabilistic implications are valid in $\mathbb{K}$.*

1. $\{\mathsf{d} \geq p.\, X\} \to \{\mathsf{d} \geq q.\, Y\}$ *if $X \supseteq Y$ and $p \geq q$.*
2. $\{\mathsf{d} \geq p.\, X, \mathsf{d} \geq q.\, Y\} \to \{\mathsf{d} \geq p+q-1.\, X \cup Y\}$ *if $p+q-1 > 0$.*

Proof. Clearly, $\mathbb{P}(\{g\}, X) \geq p$ implies that $\mathbb{P}(\{g\}, X) \geq q$. Furthermore, since $X \supseteq Y$ yields that $\{g\} \times X \times \{w\} \supseteq \{g\} \times Y \times \{w\}$, we conclude that $\mathbb{P}(\{g\}, Y) \geq q$. For the second implication, observe that in case $p+q-1 > 0$ the intersection of $\{\, w \in W \mid \forall\, x \in X \colon (g, x, w) \in I \,\}$ and $\{\, w \in W \mid \forall\, y \in Y \colon (g, y, w) \in I \,\}$ must be non-empty, and in particular must have a $\mathbb{P}$-measure of at least $p+q-1$. $\square$

Define the background knowledge $\mathcal{T}(M)$ to consist of all trivial probabilistic implications over M, i.e.,

$$\mathcal{T}(M) := \{\, \mathbf{X} \to \mathbf{Y} \mid \mathbf{X} \to \mathbf{Y} \in \mathsf{dImp}(M) \text{ and } \emptyset \models \mathbf{X} \to \mathbf{Y} \,\}.$$

In particular, $\mathcal{T}(M)$ contains all implications from Lemma 10. Since our aim is to construct a knowledge base for a given probabilistic formal context, we do not have to explicitly compute these trivial implications, and we will hence utilize them as background knowledge, which is already present and known.

Lemma 11. *Let $\mathcal{P} \cup \{\mathbf{X} \to \mathbf{Y}\} \subseteq \mathsf{dImp}(M)$ be a set of probabilistic implications over M. If $\mathcal{P} \cup \mathcal{T}(M) \models \mathbf{X} \to \mathbf{Y}$ with respect to non-probabilistic entailment, then $\mathcal{P} \models \mathbf{X} \to \mathbf{Y}$ with respect to probabilistic entailment.*

Proof. Consider a probabilistic formal context $\mathbb{K}$ with $\mathbb{K} \models \mathcal{P}$. Then Lemma 14 implies that $\mathsf{d}(\mathbb{K}) \models \mathcal{P}$. Furthermore, it is trivial that $\mathsf{d}(\mathbb{K}) \models \mathcal{T}(M)$. Consequently, the probabilistic implication $\mathbf{X} \to \mathbf{Y}$ is valid in $\mathsf{d}(\mathbb{K})$, and another application of Lemma 14 yields that $\mathbb{K} \models \mathbf{X} \to \mathbf{Y}$. $\square$

5.2 Approximations of Probabilities

Assume that $\mathbb{K} \coloneqq (G, M, W, I, \mathbb{P})$ is a probabilistic context. Then the set

$$V(\mathbb{K}) \coloneqq \{\, \mathbb{P}(\{g\}, B) \mid g \in G \text{ and } B \subseteq M \,\}$$

contains all values that can occur when evaluating the validity of implications over probabilistic attributes in $\mathbb{K}$. Note that according to Lemma 5 it holds true that

$$V(\mathbb{K}) = \{\, \mathbb{P}(\{g\}, B) \mid g \in G \text{ and } B \in \mathsf{Int}(\mathbb{K}_{\varepsilon}^{\times}) \,\}.$$

Furthermore, we define an *upper approximation* of probability values as follows: For an arbitrary $p \in [0, 1]$, let $\lceil p \rceil_{\mathbb{K}}$ be the smallest value above p which can occur when evaluating a probabilistic attribute in $\mathbb{K}$, i.e., we define

$$\lceil p \rceil_{\mathbb{K}} \coloneqq 1 \wedge \bigwedge \{\, q \mid p \leq q \text{ and } q \in V(\mathbb{K}) \,\}.$$

Dually, we define the *lower approximation* of $p \in [0, 1]$ in $\mathbb{K}$ as

$$\lfloor p \rfloor_{\mathbb{K}} \coloneqq 0 \vee \bigvee \{\, q \mid q \leq p \text{ and } q \in V(\mathbb{K}) \,\}.$$

It is easy to verify that for all attribute sets $B \subseteq M$ and all probability values $p \in [0, 1]$, the following entailment is valid:

$$\mathbb{K} \models \{\{\mathsf{d} \geq p.\, B\} \to \{\mathsf{d} \geq \lceil p \rceil_{\mathbb{K}}.\, B\}, \{\mathsf{d} \geq \lceil p \rceil_{\mathbb{K}}.\, B\} \to \{\mathsf{d} \geq p.\, B\}\}.$$

The second implication is in fact valid in arbitrary probabilistic contexts, since we can apply Lemma 10 with $p \leq \lceil p \rceil_{\mathbb{K}}$. For the first implication, observe that for all objects $g \in G$ and all attribute set $B \subseteq M$, it holds true that

$$\lfloor \mathbb{P}(\{g\}, B) \rfloor_{\mathbb{K}} = \mathbb{P}(\{g\}, B) = \lceil \mathbb{P}(\{g\}, B) \rceil_{\mathbb{K}},$$

and thus in particular $\mathbb{P}(\{g\}, B) \geq p$ if, and only if, $\mathbb{P}(\{g\}, B) \geq \lceil p \rceil_{\mathbb{K}}$. Analogously, $\mathbb{P}(\{g\}, B) \leq p$ is equivalent to $\mathbb{P}(\{g\}, B) \leq \lfloor p \rfloor_{\mathbb{K}}$.

Lemma 12. *Let $\mathbb{K}$ be a probabilistic formal context, and assume that $p \in (0, 1)$ is a probability value. Then the implication $(\lfloor p \rfloor_{\mathbb{K}}, \lceil p \rceil_{\mathbb{K}}] \to \{\lceil p \rceil_{\mathbb{K}}\}$ is valid in $\mathbb{K}$. Furthermore, the implications $[0, \lceil 0 \rceil_{\mathbb{K}}] \to \{\lceil 0 \rceil_{\mathbb{K}}\}$ and $(\lfloor 1 \rfloor_{\mathbb{K}}, 1] \to \{1\}$ are valid in $\mathbb{K}$.*

Proof. Consider an arbitrary object $g \in G$ as well as an arbitrary attribute set $B \subseteq M$, and assume that $\lfloor p \rfloor_{\mathbb{K}} < \mathbb{P}(\{g\}, B) \leq \lceil p \rceil_{\mathbb{K}}$. Of course, it holds true that $(\lfloor p \rfloor_{\mathbb{K}}, \lceil p \rceil_{\mathbb{K}}] \cap V(\mathbb{K}) = \{\lceil p \rceil_{\mathbb{K}}\}$, and consequently $\mathbb{P}(\{g\}, B) = \lceil p \rceil_{\mathbb{K}}$.

Now consider the implication $[0, \lceil 0 \rceil_{\mathbb{K}}] \to \{\lceil 0 \rceil_{\mathbb{K}}\}$. In case $0 \in V(\mathbb{K})$ we have that $\lceil 0 \rceil_{\mathbb{K}} = 0$, and then the implication is trivial. Otherwise, it follows that $[0, \lceil 0 \rceil_{\mathbb{K}}] \cap V(\mathbb{K}) = \{\lceil 0 \rceil_{\mathbb{K}}\}$, and consequently $0 \leq \mathbb{P}(\{g\}, B) \leq \lceil 0 \rceil_{\mathbb{K}}$ implies $\mathbb{P}(\{g\}, B) = \lceil 0 \rceil_{\mathbb{K}}$.

Eventually, we prove the validity of the implication $(\lfloor 1 \rfloor_{\mathbb{K}}, 1] \to \{1\}$. If $1 \in V(\mathbb{K})$, then $\lfloor 1 \rfloor_{\mathbb{K}} = 1$, and hence the premise interval $(1, 1]$ is empty, i.e., the implication trivially holds in $\mathbb{K}$. Otherwise, $(\lfloor 1 \rfloor_{\mathbb{K}}, 1] \cap V(\mathbb{K}) = \emptyset$, and the implication is again trivial. $\square$

5.3 The Probabilistic Scaling

Definition 13. *Let* $\mathbb{K} := (G, M, W, I, \mathbb{P})$ *be a probabilistic formal context. The* probabilistic scaling *of* $\mathbb{K}$ *is defined as the formal context* $\mathsf{d}(\mathbb{K}) := (G, \mathsf{d}(M), I)$ *the incidence relation* I *of which is defined by*

$$(g, \mathsf{d} \geq p.\, B) \in I \text{ if } \mathbb{P}(\{g\}, B) \geq p,$$

and by $\mathsf{d}^*(\mathbb{K})$ *we denote the subcontext of* $\mathsf{d}(\mathbb{K})$ *with the attribute set*

$$\mathsf{d}^*(M) := \{\, \mathsf{d} \geq \lceil p \rceil_{\mathbb{K}}.\, B^{I_\varepsilon^\times I_\varepsilon^\times} \mid p \in [0,1],\ \lceil p \rceil_{\mathbb{K}} \neq 0,\ \text{and } B \subseteq M,\ B^{I_\varepsilon^\times I_\varepsilon^\times} \neq \emptyset \,\}.$$

Figure 3 shows the probabilistic scaling $\mathsf{d}^*(\mathbb{K}_{\mathsf{ex}})$ the attribute set of which is given by

$$\mathsf{d}^*(\{m_1, m_2, m_3\}) = \{\, \mathsf{d} \geq p.\, B \mid B \in \mathsf{Int}((\mathbb{K}_{\mathsf{ex}})_\varepsilon^\times) \setminus \{\emptyset\} \text{ and } p \in V(\mathbb{K}_{\mathsf{ex}}) \setminus \{0\} \,\}$$
$$= \left\{ \begin{array}{l} \mathsf{d} \geq p.\{m_2\}, \mathsf{d} \geq p.\{m_3\}, \mathsf{d} \geq p.\{m_1, m_3\}, \\ \mathsf{d} \geq p.\{m_2, m_3\}, \mathsf{d} \geq p.\{m_1, m_2, m_3\} \end{array} \,\middle|\, p \in \{\tfrac{1}{6}, \tfrac{1}{3}, \tfrac{2}{3}, \tfrac{5}{6}, 1\} \right\}.$$

Please note that the dual formal context is displayed, i.e., the incidence table is transposed. More formally, for a formal context $\mathbb{K} := (G, M, I)$ its *dual context* is $\mathbb{K}^\partial := (M, G, \{\, (m, g) \mid (g, m) \in I \,\})$.

Lemma 14. *Let* $\mathbb{K}$ *be a probabilistic formal context. Then for all probabilistic implications* $\{\, \mathsf{d} \geq p_t.\, X_t \mid t \in T \,\} \to \{\mathsf{d} \geq q.\, Y\}$*, the following statements are equivalent:*

1. $\mathbb{K} \models \{\, \mathsf{d} \geq p_t.\, X_t \mid t \in T \,\} \to \{\mathsf{d} \geq q.\, Y\}$
2. $\mathsf{d}(\mathbb{K}) \models \{\, \mathsf{d} \geq p_t.\, X_t \mid t \in T \,\} \to \{\mathsf{d} \geq q.\, Y\}$
3. $\mathsf{d}^*(\mathbb{K}) \models \{\, \mathsf{d} \geq \lceil p_t \rceil_{\mathbb{K}}.\, X_t^{I_\varepsilon^\times I_\varepsilon^\times} \mid t \in T \,\} \to \{\mathsf{d} \geq \lceil q \rceil_{\mathbb{K}}.\, Y^{I_\varepsilon^\times I_\varepsilon^\times}\}$

Proof. Statements 1 and 2 are equivalent by Definition 4. Furthermore, from Lemma 5 we conclude that $(\mathsf{d} \geq p.\, X)^I = (\mathsf{d} \geq p.\, X^{I_\varepsilon^\times I_\varepsilon^\times})^I$ for all probabilistic attributes $\mathsf{d} \geq p.\, X$. From this and the fact that $\mathbb{P}(\{g\}, B) \geq p$ is equivalent to $\mathbb{P}(\{g\}, B) \geq \lceil p \rceil_{\mathbb{K}}$, the equivalence of Statements 2 and 3 then follows easily. □

5.4 Construction of the Probabilistic Implicational Knowledge Base

Theorem 15. *Let* $\mathbb{K} := (G, M, W, I, \mathbb{P})$ *be a probabilistic context. Then* $\mathcal{K}(\mathbb{K}) := (\mathcal{V}(\mathbb{K}), \mathsf{Can}(\mathbb{K}_\varepsilon^\times), \mathsf{Can}(\mathsf{d}^*(\mathbb{K}), \mathcal{T}(M)))$ *is a probabilistic implicational base for* $\mathbb{K}$*, where*

$$\mathcal{V}(\mathbb{K}) := \{\, [0, \lceil 0 \rceil_{\mathbb{K}}] \to \{\lceil 0 \rceil_{\mathbb{K}}\} \mid 0 \neq \lceil 0 \rceil_{\mathbb{K}} \,\} \cup \{\, (\lfloor 1 \rfloor_{\mathbb{K}}, 1] \to \{1\} \mid \lfloor 1 \rfloor_{\mathbb{K}} \neq 1 \,\}$$
$$\cup \{\, (\lfloor p \rfloor_{\mathbb{K}}, \lceil p \rceil_{\mathbb{K}}] \to \{\lceil p \rceil_{\mathbb{K}}\} \mid p \in (\lceil 0 \rceil_{\mathbb{K}}, \lfloor 1 \rfloor_{\mathbb{K}}) \setminus V(\mathbb{K}) \,\}.$$

If $\mathbb{K}$ *is finite, then* $\mathcal{K}(\mathbb{K})$ *is finite, too. Furthermore, finiteness of* $\mathcal{K}(\mathbb{K})$ *is also ensured in case of finiteness of both* M *and* $V(\mathbb{K})$*.*

$(\mathsf{d}^*(\mathbb{K}_{\mathsf{ex}}))^\partial$	g_1	g_2	g_3
$\mathsf{d} \geq \frac{1}{6}.\{m_2\}$	·	×	×
$\mathsf{d} \geq \frac{1}{6}.\{m_3\}$	×	×	×
$\mathsf{d} \geq \frac{1}{6}.\{m_1, m_3\}$	×	·	×
$\mathsf{d} \geq \frac{1}{6}.\{m_2, m_3\}$	·	×	·
$\mathsf{d} \geq \frac{1}{6}.\{m_1, m_2, m_3\}$	·	·	·
$\mathsf{d} \geq \frac{1}{3}.\{m_2\}$	·	×	·
$\mathsf{d} \geq \frac{1}{3}.\{m_3\}$	×	×	×
$\mathsf{d} \geq \frac{1}{3}.\{m_1, m_3\}$	×	·	×
$\mathsf{d} \geq \frac{1}{3}.\{m_2, m_3\}$	·	×	·
$\mathsf{d} \geq \frac{1}{3}.\{m_1, m_2, m_3\}$	·	·	·
$\mathsf{d} \geq \frac{2}{3}.\{m_2\}$	·	×	·
$\mathsf{d} \geq \frac{2}{3}.\{m_3\}$	×	×	×
$\mathsf{d} \geq \frac{2}{3}.\{m_1, m_3\}$	×	·	·
$\mathsf{d} \geq \frac{2}{3}.\{m_2, m_3\}$	·	×	·
$\mathsf{d} \geq \frac{2}{3}.\{m_1, m_2, m_3\}$	·	·	·
$\mathsf{d} \geq \frac{5}{6}.\{m_2\}$	·	×	·
$\mathsf{d} \geq \frac{5}{6}.\{m_3\}$	×	·	×
$\mathsf{d} \geq \frac{5}{6}.\{m_1, m_3\}$	×	·	·
$\mathsf{d} \geq \frac{5}{6}.\{m_2, m_3\}$	·	·	·
$\mathsf{d} \geq \frac{5}{6}.\{m_1, m_2, m_3\}$	·	·	·
$\mathsf{d} \geq 1.\{m_2\}$	·	×	·
$\mathsf{d} \geq 1.\{m_3\}$	×	·	·
$\mathsf{d} \geq 1.\{m_1, m_3\}$	×	·	·
$\mathsf{d} \geq 1.\{m_2, m_3\}$	·	·	·
$\mathsf{d} \geq 1.\{m_1, m_2, m_3\}$	·	·	·

Fig. 3. The probabilistic scaling of the probabilistic formal context given in Figure 1

Proof. We first prove soundness. Lemma 12 yields that $\mathcal{V}(\mathbb{K})$ is valid in $\mathbb{K}$, Lemma 8 proves that $\mathsf{Can}(\mathbb{K}_\varepsilon^\times)$ is valid in $\mathbb{K}$, and Lemma 14 shows the validity of $\mathsf{Can}(\mathsf{d}^*(\mathbb{K}), \mathcal{T}(M))$ in $\mathbb{K}$.

We proceed with proving completeness. Assume that

$$\mathbb{K} \models \{\, \mathsf{d} \geq p_t.\, X_t \mid t \in T \,\} \to \{\mathsf{d} \geq q.\, Y\}.$$

Then by Lemma 14 the implication

$$\{\, \mathsf{d} \geq \lceil p_t \rceil_{\mathbb{K}}.\, X_t^{I_\varepsilon^\times I_\varepsilon^\times} \mid t \in T \,\} \to \{\mathsf{d} \geq \lceil q \rceil_{\mathbb{K}}.\, Y^{I_\varepsilon^\times I_\varepsilon^\times}\}$$

is valid in $\mathsf{d}^*(\mathbb{K})$, and must hence be entailed by $\mathsf{Can}(\mathsf{d}^*(\mathbb{K}), \mathcal{T}(M)) \cup \mathcal{T}(M)$ with respect to non-probabilistic entailment. Lemma 11 then yields that it must also be entailed by $\mathsf{Can}(\mathsf{d}^*(\mathbb{K}), \mathcal{T}(M))$ with respect to probabilistic entailment. Furthermore, the implications

$$(\lfloor p \rfloor_{\mathbb{K}}, \lceil p \rceil_{\mathbb{K}}] \to \{\lceil p \rceil_{\mathbb{K}}\}$$

for $p \in (\{\, p_t \mid t \in T \,\} \cup \{q\}) \setminus V(\mathbb{K})$ are contained in $\mathcal{V}(\mathbb{K})$, and the implications $X \to X^{I_\varepsilon^\times I_\varepsilon^\times}$ for $X \in \{\, X_t \mid t \in T \,\} \cup \{Y\}$ are entailed by $\mathsf{Can}(\mathbb{K}_\varepsilon^\times)$. Utilizing Lemmas 7 and 8 and summing up, it holds true that

$$\begin{aligned} \mathcal{K}(\mathbb{K}) \models \{\, &\{\mathsf{d} \geq p_t.\, X_t\} \to \{\mathsf{d} \geq \lceil p_t \rceil_{\mathbb{K}}.\, X_t\} \mid t \in T \,\} \\ &\cup \{\, \{\mathsf{d} \geq \lceil p_t \rceil_{\mathbb{K}}.\, X_t\} \to \{\mathsf{d} \geq \lceil p_t \rceil_{\mathbb{K}}.\, X_t^{I_\varepsilon^\times I_\varepsilon^\times}\} \mid t \in T \,\} \\ &\cup \{\{\, \mathsf{d} \geq \lceil p_t \rceil_{\mathbb{K}}.\, X_t^{I_\varepsilon^\times I_\varepsilon^\times} \mid t \in T \,\} \to \{\mathsf{d} \geq \lceil q \rceil_{\mathbb{K}}.\, Y^{I_\varepsilon^\times I_\varepsilon^\times}\}\}, \end{aligned}$$

and so $\mathcal{K}(\mathbb{K})$ also entails $\{\, \mathsf{d} \geq p_t.\, X_t \mid t \in T \,\} \to \{\mathsf{d} \geq \lceil q \rceil_{\mathbb{K}}.\, Y^{I_\varepsilon^\times I_\varepsilon^\times}\}$. Of course, the implications

$$\{\mathsf{d} \geq \lceil q \rceil_{\mathbb{K}}.\, Y^{I_\varepsilon^\times I_\varepsilon^\times}\} \to \{\mathsf{d} \geq q.\, Y^{I_\varepsilon^\times I_\varepsilon^\times}\} \text{ and } \{\mathsf{d} \geq q.\, Y^{I_\varepsilon^\times I_\varepsilon^\times}\} \to \{\mathsf{d} \geq q.\, Y\}$$

are trivial, i.e., are valid in all probabilistic formal contexts. Eventually, we have thus just shown that the considered probabilistic implication $\{\, \mathsf{d} \geq p_t.\, X_t \mid t \in T \,\} \to \{\mathsf{d} \geq q.\, Y\}$ follows from $\mathcal{K}(\mathbb{K})$.

As last step, we prove the claim on finiteness of $\mathcal{K}(\mathbb{K})$ if both M and $V(\mathbb{K})$ are finite. Clearly, then $\mathcal{V}(\mathbb{K})$ is finite. The almost certain scaling $\mathbb{K}_\varepsilon^\times$ then possess a finite attribute set, and thus its canonical base must be finite. It also follows that the attribute set $\mathsf{d}^*(M)$ of the probabilistic scaling $\mathsf{d}^*(\mathbb{K})$ is finite, and hence the canonical base of $\mathsf{d}^*(\mathbb{K})$ is finite as well. □

Returning back to our running example $\mathbb{K}_{\mathsf{ex}}$ from Figure 1, we now construct its probabilistic implicational knowledge base $\mathcal{K}(\mathbb{K}_{\mathsf{ex}})$. The first component is the following set of implications between probability values:

$$\{(0, \tfrac{1}{6}] \to \{\tfrac{1}{6}\},\ (\tfrac{1}{6}, \tfrac{1}{3}] \to \{\tfrac{1}{3}\},\ (\tfrac{1}{3}, \tfrac{2}{3}] \to \{\tfrac{2}{3}\},\ (\tfrac{2}{3}, \tfrac{5}{6}] \to \{\tfrac{5}{6}\},\ (\tfrac{5}{6}, 1] \to \{1\}\}.$$

Note that we left out the trivial implication $\{0\} \to \{0\}$. The canonical base of the certain scaling $(\mathbb{K}_{\mathsf{ex}})_\varepsilon^\times$ was computed as $\{\{m_1\} \to \{m_3\}\}$, which consequently is the second component $\mathcal{K}(\mathbb{K}_{\mathsf{ex}})$. For the computation of the third

$$\mathsf{Can}(\mathsf{d}^*(\mathbb{K}_{\mathsf{ex}}), \mathcal{T}'(\{m_1, m_2, m_3\})) = \left\{ \begin{array}{l}
\quad \emptyset \\
\to \{\mathsf{d} \geq \frac{2}{3}.\{m_3\}\}, \\
\quad \{\mathsf{d} \geq 1.\{m_1, m_3\}, \mathsf{d} \geq \frac{1}{6}.\{m_2\}\} \\
\to \{\mathsf{d} \geq 1.\{m_1, m_2, m_3\}\}, \\
\quad \{\mathsf{d} \geq \frac{2}{3}.\{m_2, m_3\}, \mathsf{d} \geq \frac{5}{6}.\{m_3\}, \mathsf{d} \geq 1.\{m_2\}, \mathsf{d} \geq \frac{1}{3}.\{m_1, m_3\}\} \\
\to \{\mathsf{d} \geq 1.\{m_1, m_2, m_3\}\}, \\
\quad \{\mathsf{d} \geq 1.\{m_3\}, \mathsf{d} \geq \frac{1}{3}.\{m_1, m_3\}\} \\
\to \{\mathsf{d} \geq 1.\{m_1, m_3\}\}, \\
\quad \{\mathsf{d} \geq \frac{2}{3}.\{m_1, m_3\}, \mathsf{d} \geq \frac{5}{6}.\{m_3\}\} \\
\to \{\mathsf{d} \geq 1.\{m_1, m_3\}\}, \\
\quad \{\mathsf{d} \geq \frac{2}{3}.\{m_3\}, \mathsf{d} \geq \frac{1}{6}.\{m_2, m_3\}\} \\
\to \{\mathsf{d} \geq \frac{2}{3}.\{m_2, m_3\}, \mathsf{d} \geq 1.\{m_2\}\}, \\
\quad \{\mathsf{d} \geq \frac{2}{3}.\{m_3\}, \mathsf{d} \geq \frac{1}{3}.\{m_2\}\} \\
\to \{\mathsf{d} \geq \frac{2}{3}.\{m_2, m_3\}, \mathsf{d} \geq 1.\{m_2\}\}, \\
\quad \{\mathsf{d} \geq \frac{5}{6}.\{m_3\}\} \\
\to \{\mathsf{d} \geq \frac{1}{3}.\{m_1, m_3\}\}, \\
\quad \{\mathsf{d} \geq \frac{2}{3}.\{m_3\}, \mathsf{d} \geq \frac{1}{6}.\{m_1, m_3\}\} \\
\to \{\mathsf{d} \geq \frac{5}{6}.\{m_3\}, \mathsf{d} \geq \frac{1}{3}.\{m_1, m_3\}\}
\end{array} \right\}$$

Fig. 4. The implicational base of $\mathsf{d}^*(\mathbb{K}_{\mathsf{ex}})$ with respect to the background implications that are described in Lemma 10.

component of $\mathcal{K}(\mathbb{K}_{\mathsf{ex}})$, we consider the probabilistic scaling of $\mathbb{K}_{\mathsf{ex}}$. In order to avoid the axiomatization of trivial implications, we construct the implicational base of $\mathsf{d}^*(\mathbb{K}_{\mathsf{ex}})$ relative to the implication set containing all those probabilistic implications which are described in Lemma 10. This set $\mathcal{T}'(\{m_1, m_2, m_3\})$ of background knowledge contains, among others, the following implications:

$$\begin{aligned}
\{\mathsf{d} \geq \tfrac{5}{6}.\{m_1, m_3\}\} &\to \{\mathsf{d} \geq \tfrac{1}{3}.\{m_1\}\} \\
\{\mathsf{d} \geq \tfrac{5}{6}.\{m_1, m_3\}\} &\to \{\mathsf{d} \geq \tfrac{2}{3}.\{m_3\}\} \\
&\vdots \\
\{\mathsf{d} \geq \tfrac{2}{3}.\{m_2\}, \mathsf{d} \geq \tfrac{2}{3}.\{m_3\}\} &\to \{\mathsf{d} \geq \tfrac{1}{3}.\{m_2, m_3\}\} \\
\{\mathsf{d} \geq \tfrac{5}{6}.\{m_1, m_3\}, \mathsf{d} \geq \tfrac{5}{6}.\{m_2, m_3\}\} &\to \{\mathsf{d} \geq \tfrac{2}{3}.\{m_1, m_2, m_3\}\} \\
&\vdots
\end{aligned}$$

The resulting probabilistic implication set is presented in Figure 4. Please note that we possibly did not include all trivial probabilistic implications in the background knowledge, as we do not yet know an appropriate reasoning procedure.

6 Conclusion

In this document we have investigated a method for the axiomatization of rules that are valid in a given probabilistic data set, which is represented as a multi-world view over the same set of entities and vocabulary to describe the entities, and which is furthermore equipped with a probability measure on the set of worlds. We have developed such a method in the field of Formal Concept Analysis, where it is possible to assign properties to single objects. We have achieved the description of a technique for a sound and complete axiomatization of terminological knowledge which is valid in the input data set and expressible in the chosen description language. It is only natural to extend the results to a probabilistic version of the light-weight description logic $\mathcal{EL}^{\perp}$, which not only allows for assigning properties (*concept names*) to entities, but also allows for connecting pairs of entities by binary relations (*role names*). This will be subject of an upcoming publication.

It remains to apply the proposed method to concrete real-world data sets, e.g., in the medical domain, where worlds are represented by patients, or in natural sciences, where worlds are represented by repetitions of the same experiment, etc. From a theoretical perspective, it is interesting to investigate whether the constructed probabilistic implicational knowledge base is of a minimal size – as it holds true for the canonical base in the non-probabilistic case. Furthermore, so far we have only considered a model-theoretic semantics, and the induced semantic entailment between implication sets. For practical purposes, a syntactic entailment in the favor of the *Armstrong rules* in FCA is currently missing.

Acknowledgements. The author gratefully thanks the anonymous reviewers for their constructive hints and helpful remarks.

References

[1] Sadok Ben Yahia and Jan Konecny, eds. *Proceedings of the Twelfth International Conference on Concept Lattices and Their Applications, Clermont-Ferrand, France, October 13-16, 2015*. Vol. 1466. CEUR Workshop Proceedings. CEUR-WS.org, 2015.

[2] Bernhard Ganter. "Attribute Exploration with Background Knowledge". In: *Theoretical Computer Science* 217.2 (1999), pp. 215–233.

[3] Bernhard Ganter. *Two Basic Algorithms in Concept Analysis*. FB4-Preprint 831. Darmstadt, Germany: Technische Hochschule Darmstadt, 1984.

[4] Bernhard Ganter. "Two Basic Algorithms in Concept Analysis". In: *Formal Concept Analysis, 8th International Conference, ICFCA 2010, Agadir, Morocco, March 15-18, 2010. Proceedings*. Ed. by Léonard Kwuida and Barış Sertkaya. Vol. 5986. Lecture Notes in Computer Science. Springer, 2010, pp. 312–340.

[5] Bernhard Ganter and Rudolf Wille. *Formal Concept Analysis - Mathematical Foundations*. Springer, 1999.

[6] Jean-Luc Guigues and Vincent Duquenne. “Famille minimale d’implications informatives résultant d’un tableau de données binaires”. In: *Mathématiques et Sciences Humaines* 95 (1986), pp. 5–18.

[7] Francesco Kriegel. *Concept Explorer FX*. Software for Formal Concept Analysis with Description Logic Extensions. 2010–2017. URL: `https://github.com/francesco-kriegel/conexp-fx`.

[8] Francesco Kriegel. *NextClosures – Parallel Exploration of Constrained Closure Operators*. LTCS-Report 15-01. Chair for Automata Theory, Technische Universität Dresden, 2015.

[9] Francesco Kriegel. “Parallel Attribute Exploration”. In: *Graph-Based Representation and Reasoning - 22nd International Conference on Conceptual Structures, ICCS 2016, Annecy, France, July 5-7, 2016, Proceedings*. Ed. by Ollivier Haemmerlé, Gem Stapleton, and Catherine Faron-Zucker. Vol. 9717. Lecture Notes in Computer Science. Springer, 2016, pp. 91–106.

[10] Francesco Kriegel. “Probabilistic Implicational Bases in FCA and Probabilistic Bases of GCIs in $\mathcal{EL}^{\perp}$”. In: *Proceedings of the Twelfth International Conference on Concept Lattices and Their Applications, Clermont-Ferrand, France, October 13-16, 2015*. Ed. by Sadok Ben Yahia and Jan Konecny. Vol. 1466. CEUR Workshop Proceedings. CEUR-WS.org, 2015, pp. 193–204.

[11] Francesco Kriegel and Daniel Borchmann. “NextClosures: Parallel Computation of the Canonical Base”. In: *Proceedings of the Twelfth International Conference on Concept Lattices and Their Applications, Clermont-Ferrand, France, October 13-16, 2015*. Ed. by Sadok Ben Yahia and Jan Konecny. Vol. 1466. CEUR Workshop Proceedings. CEUR-WS.org, 2015, pp. 181–192.

[12] David Maier. *The Theory of Relational Databases*. Computer Science Press, 1983.

[13] Gerd Stumme. “Attribute Exploration with Background Implications and Exceptions”. English. In: *Data Analysis and Information Systems: Statistical and Conceptual Approaches, Proceedings of the 19th Annual Conference of the Gesellschaft für Klassifikation e.V., University of Basel, March 8–10, 1995*. Ed. by Hans-Hermann Bock and Wolfgang Polasek. Studies in Classification, Data Analysis, and Knowledge Organization. Berlin, Heidelberg: Springer, 1996, pp. 457–469.

[14] Rudolf Wille. “Restructuring Lattice Theory: An Approach Based on Hierarchies of Concepts”. In: *Ordered Sets: Proceedings of the NATO Advanced Study Institute held at Banff, Canada, August 28 to September 12, 1981*. Ed. by Ivan Rival. Dordrecht: Springer Netherlands, 1982, pp. 445–470.