Probabilistic Implicational Bases in FCA and Probabilistic Bases of GCIs in $\mathcal{EL}^\perp$

Francesco Kriegel

Institute for Theoretical Computer Science, TU Dresden, Germany
francesco.kriegel@tu-dresden.de
http://lat.inf.tu-dresden.de/~francesco

Abstract. A probabilistic formal context is a triadic context whose third dimension is a set of worlds equipped with a probability measure. After a formal definition of this notion, this document introduces probability of implications, and provides a construction for a base of implications whose probability satisfy a given lower threshold. A comparison between confidence and probability of implications is drawn, which yields the fact that both measures do not coincide, and cannot be compared. Furthermore, the results are extended towards the light-weight description logic $\mathcal{EL}^\perp$ with probabilistic interpretations, and a method for computing a base of general concept inclusions whose probability fulfill a certain lower bound is proposed.

Keywords: Formal Concept Analysis, Description Logics, Probabilistic Formal Context, Probabilistic Interpretation, Implication, General Concept Inclusion

1 Introduction

Most data-sets from real-world applications contain errors and noise. Hence, for mining them special techniques are necessary in order to circumvent the expression of the errors. This document focuses on rule mining, especially we attempt to extract rules that are approximately valid in data-sets, or families of data-sets, respectively. There are at least two measures for the approximate soundness of rules, namely *confidence* and *probability*. While confidence expresses the number of counterexamples in a single data-set, probability expresses somehow the number of data-sets in a data-set family that do not contain any counterexample. More specifically, we consider implications in the formal concept analysis setting [7], and general concept inclusions (GCIs) in the description logics setting [1] (in the light-weight description logic $\mathcal{EL}^\perp$).

Firstly, for axiomatizing rules from formal contexts possibly containing wrong incidences or having missing incidences the notion of a partial implication (also called association rule) and confidence has been defined by Luxenburger in [12]. Furthermore, Luxenburger introduced a method for the computation of a base of all partial implications holding in a formal context whose confidence is above a certain threshold. In [2] Borchmann has extended the results to the description logic $\mathcal{EL}^\perp$ by adjusting the notion of confidence to GCIs, and also gave a method for the construction of a base of confident GCIs for an interpretation.

Secondly, another perspective is a family of data-sets representing different views of the same domain, e.g., knowledge of different persons, or observations of an experiment that has been repeated several times, since some effects could not be observed in every case. In the field of formal concept analysis, Vityaev, Demin, and Ponomaryov, have introduced in probabilistic extensions of formal contexts and their formal concepts and implications, and furthermore gave some methods for their computation, cf. [4]. In [9] the author has shown some methods for the computation of a base of GCIs in probabilistic description logics where concept and role constructors are available to express probability directly in the concept descriptions. Here, we want to use another approach, and do not allow for probabilistic constructors, but define the notion of a probability of general concept inclusions in the light-weight description logic $\mathcal{EL}^{\perp}$. Furthermore, we provide a method for the computation of a base of GCIs satisfying a certain lower threshold for the probability. More specifically, we use the description logic $\mathcal{EL}^{\perp}$ with probabilistic interpretations that have been introduced by Lutz and Schröder in [11]. Beforehand, we only consider conjunctions in the language of formal concept analysis, and define the notion of a probabilistic formal context in a more general form than in [4], and provide a technique for the computation of base of implications satisfying a given lower probability threshold.

The document is structured as follows. In Section 2 some basic notions for probabilistic extensions of formal concept analysis are defined. Then, in Section 3 a method for the computation of a base for all implications satisfying a given lower probability threshold in a probabilistic formal context is developed, and its correctness is proven. The following sections extend the results to the description logic $\mathcal{EL}^{\perp}$. In particular, Section 4 introduces the basic notions for $\mathcal{EL}^{\perp}$, and defines probabilistic interpretations. Section 5 shows a technique for the construction of a base of GCIs holding in a probabilistic interpretation and fulfilling a lower probability threshold. Furthermore, a comparison of the notions of confidence and probability is drawn at the end of Section 3.

2 Probabilistic Formal Concept Analysis

A *probability measure* $\mathbb{P}$ on a countable set W is a mapping $\mathbb{P}\colon 2^W \to [0,1]$ such that $\mathbb{P}(\emptyset) = 0$ and $\mathbb{P}(W) = 1$ hold, and $\mathbb{P}$ is *σ-additive*, i.e., for all pairwise disjoint countable families $(U_n)_{n\in\mathbb{N}}$ with $U_n \subseteq W$ 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 $\mathbb{P}\{\, w \,\} > 0$ holds, 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.

Definition 1 (Probabilistic Formal Context). *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* $\cdot^{I_w}\colon 2^G \to 2^M$ *and* $\cdot^{I_w}\colon 2^M \to 2^G$ *where*

$$A^{I_w} := \{\, m \in M \mid \forall g \in A\colon (g, m, w) \in I \,\}$$
$$B^{I_w} := \{\, g \in G \mid \forall m \in B\colon (g, m, w) \in I \,\}$$

for object sets $A \subseteq G$ *and 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.

Definition 2 (Implication, Probability). *Let $\mathbb{K} = (G, M, W, I, \mathbb{P})$ be a probabilistic formal context. For attribute sets $X, Y \subseteq M$ we call $X \to Y$ an* implication *over M, and its* probability in $\mathbb{K}$ *is defined as the measure of the set of worlds it holds in, i.e.,*

$$\mathbb{P}(X \to Y) := \mathbb{P}\{\, w \in W \mid X^{I_w} \subseteq Y^{I_w} \,\}.$$

Furthermore, we define the following properties for implications $X \to Y$:

1. *$X \to Y$* holds *in world w of $\mathbb{K}$ if $X^{I_w} \subseteq Y^{I_w}$ is satisfied.*
2. *$X \to Y$* certainly holds *in $\mathbb{K}$ if it holds in all worlds of $\mathbb{K}$.*
3. *$X \to Y$* almost certainly holds *in $\mathbb{K}$ if it holds in all possible worlds of $\mathbb{K}$.*
4. *$X \to Y$* possibly holds *in $\mathbb{K}$ if it holds in a possible world of $\mathbb{K}$.*
5. *$X \to Y$ is* impossible *in $\mathbb{K}$ if it does not hold in any possible world of $\mathbb{K}$.*
6. *$X \to Y$ is* refuted *by $\mathbb{K}$ if does not hold in any world of $\mathbb{K}$.*

It is readily verified that $\mathbb{P}(X \to Y) = \mathbb{P}\{\, w \in W_\varepsilon \mid X^{I_w} \subseteq Y^{I_w} \,\} = \sum\{\, \mathbb{P}\{\, w \,\} \mid w \in W_\varepsilon \text{ and } X^{I_w} \subseteq Y^{I_w} \,\}$. An implication $X \to Y$ almost certainly holds if $\mathbb{P}(X \to Y) = 1$, possibly holds if $\mathbb{P}(X \to Y) > 0$, and is impossible if $\mathbb{P}(X \to Y) = 0$. If $X \to Y$ certainly holds, then it is almost certain, and if $X \to Y$ is refuted, then it is impossible.

3 Probabilistic Implicational Bases

At first we introduce the notion of a probabilistic implicational base. Then we will develop and prove a construction for such bases w.r.t. probabilistic formal contexts. If the underlying context is finite, then the base is computable. The reader should be aware of the standard notions of formal concept analysis in [7]. Recall that an implication follows from an implication set if, and only if, it can be syntactically deduced using the so-called *Armstrong rules* as follows: 1. From $X \supseteq Y$ infer $X \to Y$. 2. From $X \to Y$ and $Y \to Z$ infer $X \to Z$. 3. From $X_1 \to Y_1$ and $X_2 \to Y_2$ infer $X_1 \cup X_2 \to Y_1 \cup Y_2$.

Definition 3 (Probabilistic Implicational Base). *Let $\mathbb{K} = (G, M, W, I, \mathbb{P})$ be a probabilistic formal context, and $p \in [0, 1]$ a threshold. A* probabilistic implicational base *for $\mathbb{K}$ and p is an implication set $\mathcal{B}$ over M that satisfies the following properties:*

1. *$\mathcal{B}$ is* sound *for $\mathbb{K}$ and p, i.e., $\mathbb{P}(X \to Y) \geq p$ holds for all implications $X \to Y \in \mathcal{B}$, and*
2. *$\mathcal{B}$ is* complete *for $\mathbb{K}$ and p, i.e., if $\mathbb{P}(X \to Y) \geq p$, then $X \to Y$ follows from $\mathcal{B}$.*

A probabilistic implicational base is irredundant *if none of its implications follows from the others, and is* minimal *if it has minimal cardinality among all bases for $\mathbb{K}$ and p.*

It is readily verified that the above definition is a straight-forward generalization of implicational bases as defined in [7, Definition 37], in particular formal contexts coincide with probabilistic formal contexts having only one possible world, and implications holding in the formal context coincide with implications having probability 1.

We now define a transformation from probabilistic formal contexts to formal contexts. It allows to decide whether an implication (almost) certainly holds, and furthermore it can be utilized to construct an implicational base for the (almost) certain implications.

Definition 4 (Scaling). *Let $\mathbb{K}$ be a probabilistic formal context. The* certain scaling *of $\mathbb{K}$ is the formal context $\mathbb{K}^\times := (G \times W, M, I^\times)$ where $((g, w), m) \in I^\times$ iff $(g, m, w) \in I$, and the* almost certain scaling *of $\mathbb{K}$ is the subcontext $\mathbb{K}^\times_\varepsilon := (G \times W_\varepsilon, M, I^\times_\varepsilon)$ of $\mathbb{K}^\times$.*

Lemma 5. *Let $\mathbb{K} = (G, M, W, I, \mathbb{P})$ be a probabilistic formal context, and let $X \to Y$ be a formal implication. Then the following statements are satisfied:*

1. *$X \to Y$ certainly holds in $\mathbb{K}$ if, and only if, $X \to Y$ holds in $\mathbb{K}^{\times}$.*
2. *$X \to Y$ almost certainly holds in $\mathbb{K}$ if, and only if, $X \to Y$ holds in $\mathbb{K}^{\times}_{\varepsilon}$.*

Proof. It is readily verified that the following equivalences hold:

$$\begin{aligned}\mathbb{P}(X \to Y) = 1 &\Leftrightarrow \forall w \in W\colon X^{I_w} \subseteq Y^{I_w} \\ &\Leftrightarrow X^{I^{\times}} = \biguplus_{w \in W} X^{I_w} \times \{\, w \,\} \subseteq \biguplus_{w \in W} Y^{I_w} \times \{\, w \,\} = Y^{I^{\times}} \\ &\Leftrightarrow \mathbb{K}^{\times} \models X \to Y.\end{aligned}$$

The second statement can be proven analogously. □

Recall the notion of a pseudo-intent [6–8]: An attribute set $P \subseteq M$ of a formal context (G, M, I) is a *pseudo-intent* if $P \neq P^{II}$, and $Q^{II} \subseteq P$ holds for all pseudo-intents $Q \subsetneq P$. Furthermore, it is well-known that the *canonical implicational base* of a formal context (G, M, I) consists of all implications $P \to P^{II}$ where P is a pseudo-intent, cf. [6–8]. Consequently, the next corollary is an immediate consequence of Lemma 5.

Corollary 6. *Let $\mathbb{K}$ be a probabilistic formal context. Then the following statements hold:*

1. *An implicational base for $\mathbb{K}^{\times}$ is an implicational base for the certain implications of $\mathbb{K}$, in particular this holds for the following implication set:*

$$\mathcal{B}_{\mathbb{K}} := \{\, P \to P^{I^{\times} I^{\times}} \mid P \in \mathsf{PsInt}(\mathbb{K}^{\times}) \,\}.$$

2. *An implicational base for $\mathbb{K}^{\times}_{\varepsilon}$ w.r.t. the background knowledge $\mathcal{B}_{\mathbb{K}}$ is an implicational base for the almost certain implications of $\mathbb{K}$, in particular this holds for the following implication set:*

$$\mathcal{B}_{\mathbb{K},1} := \mathcal{B}_{\mathbb{K}} \cup \{\, P \to P^{I^{\times}_{\varepsilon} I^{\times}_{\varepsilon}} \mid P \in \mathsf{PsInt}(\mathbb{K}^{\times}_{\varepsilon}, \mathcal{B}_{\mathbb{K}}) \,\}.$$

Lemma 7. *Let $\mathbb{K} = (G, M, W, I, \mathbb{P})$ be a probabilistic formal context. Then the following statements are satisfied:*

1. *$Y \subseteq X$ implies that $X \to Y$ certainly holds in $\mathbb{K}$.*
2. *$X_1 \subseteq X_2$ and $Y_1 \supseteq Y_2$ imply $\mathbb{P}(X_1 \to Y_1) \leq \mathbb{P}(X_2 \to Y_2)$.*
3. *$X_0 \subseteq X_1 \subseteq \ldots \subseteq X_n$ implies $\mathbb{P}(X_0 \to X_n) \leq \bigwedge_{i=1}^{n} \mathbb{P}(X_{i-1} \to X_i)$.*

Proof. 1. If $Y \subseteq X$, then $X^{I_w} \subseteq Y^{I_w}$ follows for all worlds $w \in W$.

2. Assume $X_1 \subseteq X_2$ and $Y_2 \subseteq Y_1$. Then $X_1^{I_w} \supseteq X_2^{I_w}$ and $Y_2^{I_w} \supseteq Y_1^{I_w}$ follow for all worlds $w \in W$. Consider a world $w \in W$ where $X_1^{I_w} \subseteq Y_1^{I_w}$. Of course, we may conclude that $X_2^{I_w} \subseteq Y_2^{I_w}$. As a consequence we get $\mathbb{P}(X_1 \to Y_1) \leq \mathbb{P}(X_2 \to Y_2)$.
3. We prove the third claim by induction on n. For $n = 0$ there is nothing to show, and the case $n = 1$ is trivial. Hence, consider $n = 2$ for the induction base, and let $X_0 \subseteq X_1 \subseteq X_2$. Then we have that $X_0^{I_w} \supseteq X_1^{I_w} \supseteq X_2^{I_w}$ is satisfied in all worlds $w \in W$. Now consider a world $w \in W$ where $X_0^{I_w} \subseteq X_2^{I_w}$ is true.

Of course, it then follows that $X_0^{I_w} \subseteq X_1^{I_w} \subseteq X_2^{I_w}$. Consequently, we conclude $\mathbb{P}(X_0 \to X_2) \leq \mathbb{P}(X_0 \to X_1)$ and $\mathbb{P}(X_0 \to X_2) \leq \mathbb{P}(X_1 \to X_2)$.
For the induction step let $n > 2$. The induction hypothesis yields that

$$\mathbb{P}(X_0 \to X_{n-1}) \leq \bigwedge_{i=1}^{n-1} \mathbb{P}(X_{i-1} \to X_i).$$

Of course, it also holds that $X_0 \subseteq X_{n-1} \subseteq X_n$, and it follows by induction hypothesis and the previous inequality that

$$\mathbb{P}(X_0 \to X_n) \leq \mathbb{P}(X_0 \to X_{n-1}) \wedge \mathbb{P}(X_{n-1} \to X_n) \leq \bigwedge_{i=1}^{n} \mathbb{P}(X_{i-1} \to X_i). \qquad \square$$

Lemma 8. *Let $\mathbb{K} = (G, M, W, I, \mathbb{P})$ be a probabilistic formal context. Then for all implications $X \to Y$ the following equalities are valid:*

$$\mathbb{P}(X \to Y) = \mathbb{P}(X^{I^\times I^\times} \to Y^{I^\times I^\times}) = \mathbb{P}(X^{I_\varepsilon^\times I_\varepsilon^\times} \to Y^{I_\varepsilon^\times I_\varepsilon^\times}).$$

Proof. Let $X \to Y$ be an implication. Then for all worlds $w \in W$ it holds that

$$g \in X^{I_w} \Leftrightarrow \forall m \in X\colon (g, m, w) \in I \Leftrightarrow \forall m \in X\colon ((g, w), m) \in I^\times \Leftrightarrow (g, w) \in X^{I^\times},$$

and we conclude that $X^{I_w} = \pi_1(X^{I^\times} \cap (G \times \{ w \}))$. Furthermore, we then infer $X^{I_w} = X^{I^\times I^\times I_w}$, and thus the following equations hold:

$$\begin{aligned} \mathbb{P}(X \to Y) &= \mathbb{P}\{ w \in W \mid X^{I_w} \subseteq Y^{I_w} \} \\ &= \mathbb{P}\{ w \in W \mid X^{I^\times I^\times I_w} \subseteq Y^{I^\times I^\times I_w} \} = \mathbb{P}(X^{I^\times I^\times} \to Y^{I^\times I^\times}). \end{aligned}$$

In particular, for all possible worlds $w \in W_\varepsilon$ it holds that $g \in X^{I_w} \Leftrightarrow (g, w) \in X^{I_\varepsilon^\times}$, and thus $X^{I_w} = \pi_1(X^{I_\varepsilon^\times} \cap (G \times \{ w \}))$ and $X^{I_w} = X^{I_\varepsilon^\times I_\varepsilon^\times I_w}$ are satisfied. Consequently, it may be concluded that $\mathbb{P}(X \to Y) = \mathbb{P}(X^{I_\varepsilon^\times I_\varepsilon^\times} \to Y^{I_\varepsilon^\times I_\varepsilon^\times})$. $\square$

Lemma 9. *Let $\mathbb{K}$ be a probabilistic formal context. Then the following statements hold:*

1. *If $\mathcal{B}$ is an implicational base for the certain implications of $\mathbb{K}$, then the implication $X \to Y$ follows from $\mathcal{B} \cup \{ X^{I^\times I^\times} \to Y^{I^\times I^\times} \}$.*
2. *If $\mathcal{B}$ is an implicational base for the almost certain implications of $\mathbb{K}$, then the implication $X \to Y$ follows from $\mathcal{B} \cup \{ X^{I_\varepsilon^\times I_\varepsilon^\times} \to Y^{I_\varepsilon^\times I_\varepsilon^\times} \}$.*

Proof. Of course, the implication $X \to X^{I^\times I^\times}$ holds in $\mathbb{K}^\times$, i.e., certainly holds in $\mathbb{K}$ by Lemma 5, and hence follows from $\mathcal{B}$. Thus, the implication $X \to Y^{I^\times I^\times}$ is entailed by $\mathcal{B} \cup \{ X^{I^\times I^\times} \to Y^{I^\times I^\times} \}$, and because of $Y \subseteq Y^{I^\times I^\times}$ the claim follows.
The second statement follows analogously. $\square$

Lemma 10. *Let $\mathbb{K}$ be a probabilistic formal context. Then the following statements hold:*

1. $\mathbb{P}(X^{I_\varepsilon^\times I_\varepsilon^\times} \to Y^{I_\varepsilon^\times I_\varepsilon^\times}) = \mathbb{P}(X^{I_\varepsilon^\times I_\varepsilon^\times} \to (X \cup Y)^{I_\varepsilon^\times I_\varepsilon^\times})$,
2. $(X \cup Y)^{I_\varepsilon^\times I_\varepsilon^\times} \to Y^{I_\varepsilon^\times I_\varepsilon^\times}$ *certainly holds in $\mathbb{K}$, and*

3. $X^{I_\varepsilon^\times I_\varepsilon^\times} \to Y^{I_\varepsilon^\times I_\varepsilon^\times}$ *is entailed by* $\{\, X^{I_\varepsilon^\times I_\varepsilon^\times} \to (X \cup Y)^{I_\varepsilon^\times I_\varepsilon^\times}, (X \cup Y)^{I_\varepsilon^\times I_\varepsilon^\times} \to Y^{I_\varepsilon^\times I_\varepsilon^\times} \,\}$.

$$\begin{tikzcd} & (X \cup Y)^{I_\varepsilon^\times I_\varepsilon^\times} \arrow[dr] & \\ X^{I_\varepsilon^\times I_\varepsilon^\times} \arrow[ur, "p"] \arrow[rr, "p"'] & & Y^{I_\varepsilon^\times I_\varepsilon^\times} \end{tikzcd}$$

Proof. First note that $(X^{I_\varepsilon^\times I_\varepsilon^\times} \cup Y^{I_\varepsilon^\times I_\varepsilon^\times})^{I_\varepsilon^\times I_\varepsilon^\times} = (X \cup Y)^{I_\varepsilon^\times I_\varepsilon^\times}$. As $Y^{I_\varepsilon^\times I_\varepsilon^\times}$ is a subset of $(X^{I_\varepsilon^\times I_\varepsilon^\times} \cup Y^{I_\varepsilon^\times I_\varepsilon^\times})^{I_\varepsilon^\times I_\varepsilon^\times}$, the implication $(X \cup Y)^{I_\varepsilon^\times I_\varepsilon^\times} \to Y^{I_\varepsilon^\times I_\varepsilon^\times}$ certainly holds in $\mathbb{K}$, cf. Statement 1 in Lemma 7.

Furthermore, we have that $X^{I_w} \subseteq Y^{I_w}$ if, and only if, $X^{I_w} \subseteq X^{I_w} \cap Y^{I_w} = (X \cup Y)^{I_w}$. Hence, the implication $X \to Y$ has the same probability as $X \to X \cup Y$. Consequently, we may conclude by means of Lemma 7 that

$$\mathbb{P}(X^{I_\varepsilon^\times I_\varepsilon^\times} \to Y^{I_\varepsilon^\times I_\varepsilon^\times}) = \mathbb{P}(X \to Y) = \mathbb{P}(X \to X \cup Y) = \mathbb{P}(X^{I_\varepsilon^\times I_\varepsilon^\times} \to (X \cup Y)^{I_\varepsilon^\times I_\varepsilon^\times}).$$

Obviously, $\{\, X^{I_\varepsilon^\times I_\varepsilon^\times} \to (X \cup Y)^{I_\varepsilon^\times I_\varepsilon^\times}, (X \cup Y)^{I_\varepsilon^\times I_\varepsilon^\times} \to Y^{I_\varepsilon^\times I_\varepsilon^\times} \,\}$ entails $X^{I_\varepsilon^\times I_\varepsilon^\times} \to Y^{I_\varepsilon^\times I_\varepsilon^\times}$. □

Lemma 11. *Let $\mathbb{K}$ be a probabilistic formal context, and X, Y be intents of $\mathbb{K}_\varepsilon^\times$ such that $X \subseteq Y$ and $\mathbb{P}(X \to Y) \geq p$. Then the following statements are true:*

1. *There is a chain of neighboring intents $X = X_0 \prec X_1 \prec X_2 \prec \ldots \prec X_n = Y$ in $\mathbb{K}_\varepsilon^\times$,*
2. *$\mathbb{P}(X_{i-1} \to X_i) \geq p$ for all $i \in \{\, 1, \ldots, n \,\}$, and*
3. *$X \to Y$ is entailed by $\{\, X_{i-1} \to X_i \mid i \in \{\, 1, \ldots, n \,\} \,\}$.*

Proof. The existence of a chain $X = X_0 \prec X_1 \prec X_2 \prec \ldots \prec X_{n-1} \prec X_n = Y$ of neighboring intents between X and Y in $\mathbb{K}_\varepsilon^\times$ follows from $X \subseteq Y$.

From Statement 3 in Lemma 7 it follows that all implications $X_{i-1} \to X_i$ have a probability of at least p in $\mathbb{K}$. It is trivial that they entail $X \to Y$. □

Theorem 12 (Probabilistic Implicational Base). *Let $\mathbb{K}$ be a probabilistic formal context, and $p \in [0, 1)$ a probability threshold. Then the following implication set is a probabilistic implicational base for $\mathbb{K}$ and p:*

$$\mathcal{B}_{\mathbb{K},p} := \mathcal{B}_{\mathbb{K},1} \cup \{\, X \to Y \mid X, Y \in \mathsf{Int}(\mathbb{K}_\varepsilon^\times) \textit{ and } X \prec Y \textit{ and } \mathbb{P}(X \to Y) \geq p \,\}.$$

Proof. All implications in $\mathcal{B}_{\mathbb{K},1}$ hold almost certainly in $\mathbb{K}$, and thus have probability 1. By construction, all other implications $X \to Y$ in the second subset have a probability $\geq p$. Hence, Statement 1 in Definition 3 is satisfied.

Now consider an implication $X \to Y$ over M such that $\mathbb{P}(X \to Y) \geq p$. We have to prove Statement 2 of Definition 3, i.e., that $X \to Y$ is entailed by $\mathcal{B}_{\mathbb{K},p}$.

Lemma 8 yields that both implications $X \to Y$ and $X^{I_\varepsilon^\times I_\varepsilon^\times} \to Y^{I_\varepsilon^\times I_\varepsilon^\times}$ have the same probability. Lemma 9 states that $X \to Y$ follows from $\mathcal{B}_{\mathbb{K},1} \cup \{\, X^{I_\varepsilon^\times I_\varepsilon^\times} \to Y^{I_\varepsilon^\times I_\varepsilon^\times} \,\}$. According to Lemma 10, the implication $X^{I_\varepsilon^\times I_\varepsilon^\times} \to Y^{I_\varepsilon^\times I_\varepsilon^\times}$ follows from $\{\, X^{I_\varepsilon^\times I_\varepsilon^\times} \to (X \cup Y)^{I_\varepsilon^\times I_\varepsilon^\times}, (X \cup Y)^{I_\varepsilon^\times I_\varepsilon^\times} \to Y^{I_\varepsilon^\times I_\varepsilon^\times} \,\}$. Furthermore, it holds that

$$\mathbb{P}(X^{I_\varepsilon^\times I_\varepsilon^\times} \to (X \cup Y)^{I_\varepsilon^\times I_\varepsilon^\times}) = \mathbb{P}(X^{I_\varepsilon^\times I_\varepsilon^\times} \to Y^{I_\varepsilon^\times I_\varepsilon^\times}) = \mathbb{P}(X \to Y) \geq p,$$

and the second implication $(X \cup Y)^{I_\varepsilon^\times I_\varepsilon^\times} \to Y^{I_\varepsilon^\times I_\varepsilon^\times}$ certainly holds, i.e., follows from $\mathcal{B}_{\mathbb{K},1}$. Finally, Lemma 11 states that there is a chain of neighboring intents of $\mathbb{K}_\varepsilon^\times$ starting at $X^{I_\varepsilon^\times I_\varepsilon^\times}$ and ending at $(X \cup Y)^{I_\varepsilon^\times I_\varepsilon^\times}$, i.e.,

$$X^{I_\varepsilon^\times I_\varepsilon^\times} = X_0^{I_\varepsilon^\times I_\varepsilon^\times} \prec X_1^{I_\varepsilon^\times I_\varepsilon^\times} \prec X_2^{I_\varepsilon^\times I_\varepsilon^\times} \prec \ldots \prec X_n^{I_\varepsilon^\times I_\varepsilon^\times} = (X \cup Y)^{I_\varepsilon^\times I_\varepsilon^\times},$$

such that all implications $X_{i-1}^{I_\varepsilon^\times I_\varepsilon^\times} \to X_i^{I_\varepsilon^\times I_\varepsilon^\times}$ have a probability $\geq p$, and are thus contained in $\mathcal{B}_{\mathbb{K},p}$. Hence, $\mathcal{B}_{\mathbb{K},p}$ entails the implication $X \to Y$. $\square$

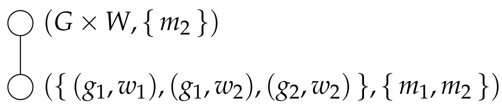
Corollary 13. *Let $\mathbb{K}$ be a probabilistic formal context. Then the following set is an implicational base for the possible implications of $\mathbb{K}$:*

$$\mathcal{B}_{\mathbb{K},\varepsilon} := \mathcal{B}_{\mathbb{K},1} \cup \{\, X \to Y \mid X, Y \in \mathsf{Int}(\mathbb{K}_\varepsilon^\times) \text{ and } X \prec Y \text{ and } \mathbb{P}(X \to Y) > 0 \,\}.$$

However, it is not possible to show irredundancy or minimality for the base of probabilistic implications given above in Theorem 12. Consider the probabilistic formal context $\mathbb{K} = (\{\, g_1, g_2 \,\}, \{\, m_1, m_2 \,\}, \{\, w_1, w_2 \,\}, I, \{\, \{\, w_1 \,\} \mapsto \frac{1}{2}, \{\, w_2 \,\} \mapsto \frac{1}{2} \,\})$ whose incidence relation I is defined as follows:

w_1	m_1	m_2
g_1	$\times$	$\times$
g_2		$\times$

w_2	m_1	m_2
g_1	$\times$	$\times$
g_2	$\times$	$\times$



The only pseudo-intent of $\mathbb{K}^\times$ is $\emptyset$, and the concept lattice of $\mathbb{K}^\times$ is shown above. Hence, we have the following probabilistic implicational base for $p = \frac{1}{2}$:

$$\mathcal{B}_{\mathbb{K},\frac{1}{2}} = \{\, \emptyset \to \{\, m_2 \,\}, \{\, m_2 \,\} \to \{\, m_1, m_2 \,\} \,\}.$$

However, the set $\mathcal{B} := \{\, \emptyset \to \{\, m_1, m_2 \,\} \,\}$ is also a probabilistic implicational base for $\mathbb{K}$ and $\frac{1}{2}$ with less elements.

In order to compute a minimal base for the implications holding in a probabilistic formal context with a probability $\geq p$, one can for example determine the above given probabilistic base, and minimize it by means of constructing the Duquenne-Guigues base of it. This either requires the transformation of the implication set into a formal context that has this implication set as an implicational base, or directly compute all pseudo-closures of the closure operator induced by the (probabilistic) implicational base.

Recall that the confidence of an implication $X \to Y$ in a formal context (G, M, I) is defined as $\mathsf{conf}(X \to Y) := \left|(X \cup Y)^I\right| / \left|X^I\right|$, cf. [12]. In general, there is no correspondence between the probability of an implication in $\mathbb{K}$ and its confidence in $\mathbb{K}^\times$ or $\mathbb{K}_\varepsilon^\times$. To prove this we will provide two counterexamples. As first counterexample we consider the context $\mathbb{K}$ above. It is readily verified that $\mathbb{P}(\{\, m_2 \,\} \to \{\, m_1 \,\}) = \frac{1}{2}$ and $\mathsf{conf}(\{\, m_2 \,\} \to \{\, m_1 \,\}) = \frac{3}{4}$, i.e., the confidence is greater than the probability. Furthermore, consider the following modification of $\mathbb{K}$ as second counterexample:

w_1	m_1	m_2
g_1	$\times$	$\times$
g_2		

w_2	m_1	m_2
g_1		$\times$
g_2		$\times$

Then we have that $\mathbb{P}(\{\, m_2 \,\} \to \{\, m_1 \,\}) = \frac{1}{2}$ and $\mathsf{conf}(\{\, m_2 \,\} \to \{\, m_1 \,\}) = \frac{1}{3}$, i.e., the confidence is smaller than the probability.

4 The Description Logic $\mathcal{EL}^\perp$ and Probabilistic Interpretations

This section gives a brief overview on the light-weight description logic $\mathcal{EL}^\perp$ [1]. First, assume that (N_C, N_R) is a signature, i.e., N_C is a set of *concept names*, and N_R is a set of *role names*, respectively. Then $\mathcal{EL}^\perp$*-concept descriptions* C over (N_C, N_R) may be constructed according to the following inductive rule (where $A \in N_C$ and $r \in N_R$):

$$C ::= \perp \mid \top \mid A \mid C \sqcap C \mid \exists\, r.\, C.$$

We shall denote the set of all $\mathcal{EL}^\perp$-concept descriptions over (N_C, N_R) by $\mathcal{EL}^\perp(N_C, N_R)$.
Second, the semantics of $\mathcal{EL}^\perp$-concept descriptions is defined by means of interpretations: An *interpretation* is a tuple $\mathcal{I} = (\Delta^\mathcal{I}, \cdot^\mathcal{I})$ that consists of a set $\Delta^\mathcal{I}$, called *domain*, and an *extension function* $\cdot^\mathcal{I}\colon N_C \cup N_R \to 2^{\Delta^\mathcal{I}} \cup 2^{\Delta^\mathcal{I} \times \Delta^\mathcal{I}}$ that maps concept names $A \in N_C$ to subsets $A^\mathcal{I} \subseteq \Delta^\mathcal{I}$ and role names $r \in N_R$ to binary relations $r^\mathcal{I} \subseteq \Delta^\mathcal{I} \times \Delta^\mathcal{I}$. The extension function is extended to all $\mathcal{EL}^\perp$-concept descriptions as follows:

$$\begin{aligned}
\perp^\mathcal{I} &:= \emptyset, \\
\top^\mathcal{I} &:= \Delta^\mathcal{I}, \\
(C \sqcap D)^\mathcal{I} &:= C^\mathcal{I} \cap D^\mathcal{I}, \\
(\exists\, r.\, C)^\mathcal{I} &:= \{\, d \in \Delta^\mathcal{I} \mid \exists e \in \Delta^\mathcal{I}\colon (d, e) \in r^\mathcal{I} \text{ and } e \in C^\mathcal{I} \,\}.
\end{aligned}$$

A *general concept inclusion (GCI)* in $\mathcal{EL}^\perp$ is of the form $C \sqsubseteq D$ where C and D are $\mathcal{EL}^\perp$-concept descriptions. It *holds* in an interpretation $\mathcal{I}$ if $C^\mathcal{I} \subseteq D^\mathcal{I}$ is satisfied, and we then also write $\mathcal{I} \models C \sqsubseteq D$, and say that $\mathcal{I}$ is a *model* of $C \sqsubseteq D$. Furthermore, C is *subsumed* by D if $C \sqsubseteq D$ holds in all interpretations, and we shall denote this by $C \sqsubseteq D$, too. A *TBox* is a set of GCIs, and a *model* of a TBox is a model of all its GCIs. A TBox $\mathcal{T}$ *entails* a GCI $C \sqsubseteq D$, denoted by $\mathcal{T} \models C \sqsubseteq D$, if every model of $\mathcal{T}$ is a model of $C \sqsubseteq D$.

To introduce probability into the description logic $\mathcal{EL}^\perp$, we now present the notion of a probabilistic interpretation from [11]. It is simply a family of interpretations over the same domain and the same signature, indexed by a set of worlds that is equipped with a probability measure.

Definition 14 (Probabilistic Interpretation, [11]). *Let (N_C, N_R) be a signature. A* probabilistic interpretation *$\mathcal{I}$ is a tuple $(\Delta^\mathcal{I}, (\cdot^{\mathcal{I}_w})_{w \in W}, W, \mathbb{P})$ consisting of a set $\Delta^\mathcal{I}$, called* domain, *a countable set W of* worlds, *a probability measure $\mathbb{P}$ on W, and an* extension function *$\cdot^{\mathcal{I}_w}$ for each world $w \in W$, i.e., $(\Delta^\mathcal{I}, \cdot^{\mathcal{I}_w})$ is an interpretation for each $w \in W$.*

For a general concept inclusion $C \sqsubseteq D$ its probability *in $\mathcal{I}$ is defined as follows:*

$$\mathbb{P}(C \sqsubseteq D) := \mathbb{P}\{\, w \in W \mid C^{\mathcal{I}_w} \subseteq D^{\mathcal{I}_w} \,\}.$$

Furthermore, for a GCI $C \sqsubseteq D$ we define the following properties (as for probabilistic formal contexts): 1. $C \sqsubseteq D$ holds *in world w if $C^{\mathcal{I}_w} \subseteq D^{\mathcal{I}_w}$. 2. $C \sqsubseteq D$* certainly holds *in $\mathcal{I}$ if it holds in all worlds. 3. $C \sqsubseteq D$* almost certainly holds *in $\mathcal{I}$ if it holds in all possible worlds. 4. $C \sqsubseteq D$* possibly holds *in $\mathcal{I}$ if it holds in a possible world. 5. $C \sqsubseteq D$ is* impossible *in $\mathcal{I}$ if it does not hold in any possible world. 6. $C \sqsubseteq D$ is* refuted *by $\mathcal{I}$ if it does not hold in any world.*

It is readily verified that $\mathbb{P}(C \sqsubseteq D) = \mathbb{P}\{\, w \in W_\varepsilon \mid C^{\mathcal{I}_w} \subseteq D^{\mathcal{I}_w} \,\} = \sum\{\, \mathbb{P}\{\, w \,\} \mid w \in W_\varepsilon \text{ and } C^{\mathcal{I}_w} \subseteq D^{\mathcal{I}_w} \,\}$ for all general concept inclusions $C \sqsubseteq D$.

5 Probabilistic Bases of GCIs

In the following we construct from a probabilistic interpretation $\mathcal{I}$ a base of GCIs that entails all GCIs with a probability greater than a given threshold p w.r.t. $\mathcal{I}$.

Definition 15 (Probabilistic Base). *Let $\mathcal{I}$ be a probabilistic interpretation, and $p \in [0,1]$ a threshold. A* probabilistic base of GCIs *for $\mathcal{I}$ and p is a TBox $\mathcal{B}$ that satisfies the following conditions:*

1. *$\mathcal{B}$ is* sound *for $\mathcal{I}$ and p, i.e., $\mathbb{P}(C \sqsubseteq D) \geq p$ for all GCIs $C \sqsubseteq D \in \mathcal{B}$, and*
2. *$\mathcal{B}$ is* complete *for $\mathcal{I}$ and p, i.e., if $\mathbb{P}(C \sqsubseteq D) \geq p$, then $\mathcal{B} \models C \sqsubseteq D$.*

A probabilistic base $\mathcal{B}$ is irredundant *if none of its GCIs follows from the others, and is* minimal *if it has minimal cardinality among all probabilistic bases for $\mathcal{I}$ and p.*

For a probabilistic interpretation $\mathcal{I}$ we define its *certain scaling* as the disjoint union of all interpretations $\mathcal{I}_w$ with $w \in W$, i.e., as the interpretation $\mathcal{I}^\times := (\Delta^{\mathcal{I}} \times W, \cdot^{\mathcal{I}^\times})$ whose extension mapping is given as follows:

$$\begin{aligned} A^{\mathcal{I}^\times} &:= \{\, (d,w) \mid d \in A^{\mathcal{I}_w} \,\} && (A \in N_C), \\ r^{\mathcal{I}^\times} &:= \{\, ((d,w),(e,w)) \mid (d,e) \in r^{\mathcal{I}_w} \,\} && (r \in N_R). \end{aligned}$$

Furthermore, the *almost certain scaling* $\mathcal{I}^\times_\varepsilon$ of $\mathcal{I}$ is the disjoint union of all interpretations $\mathcal{I}_w$ where $w \in W_\varepsilon$ is a possible world. Analogously to Lemma 5, a GCI $C \sqsubseteq D$ certainly holds in $\mathcal{I}$ iff it holds in $\mathcal{I}^\times$, and almost certainly holds in $\mathcal{I}$ iff it holds in $\mathcal{I}^\times_\varepsilon$.

In [5] the so-called *model-based most-specific concept descriptions (mmscs)* have been defined w.r.t. greatest fixpoint semantics as follows: Let $\mathcal{J}$ be an interpretation, and $X \subseteq \Delta^{\mathcal{J}}$. Then a concept description C is a *mmsc* of X in $\mathcal{J}$, if $X \subseteq C^{\mathcal{J}}$ is satisfied, and $C \sqsubseteq D$ for all concept descriptions D with $X \subseteq D^{\mathcal{J}}$. It is easy to see that all mmscs of X are unique up to equivalence, and hence we denote *the* mmsc of X in $\mathcal{J}$ by $X^{\mathcal{J}}$. Please note that there is also a role-depth bounded variant w.r.t. descriptive semantics given in [3].

Lemma 16. *Let $\mathcal{I}$ be a probabilistic interpretation. Then the following statements hold:*

1. *$C^{\mathcal{I}_w} \times \{\, w \,\} = C^{\mathcal{I}^\times} \cap (\Delta^{\mathcal{I}} \times \{\, w \,\})$ for all concept descriptions C and worlds $w \in W$.*
2. *$C^{\mathcal{I}_w} \times \{\, w \,\} = C^{\mathcal{I}^\times_\varepsilon} \cap (\Delta^{\mathcal{I}} \times \{\, w \,\})$ for all concept descriptions C and possible worlds $w \in W_\varepsilon$.*
3. *$\mathbb{P}(C \sqsubseteq D) = \mathbb{P}(C^{\mathcal{I}^\times \mathcal{I}^\times} \sqsubseteq D^{\mathcal{I}^\times \mathcal{I}^\times}) = \mathbb{P}(C^{\mathcal{I}^\times_\varepsilon \mathcal{I}^\times_\varepsilon} \sqsubseteq D^{\mathcal{I}^\times_\varepsilon \mathcal{I}^\times_\varepsilon})$ for all GCIs $C \sqsubseteq D$.*

Proof. 1. We prove the claim by structural induction on C. By definition, the statement holds for $\bot$, $\top$, and all concept names $A \in N_C$. Consider a conjunction $C \sqcap D$, then

$$\begin{aligned} (C \sqcap D)^{\mathcal{I}_w} \times \{\, w \,\} &= (C^{\mathcal{I}_w} \cap D^{\mathcal{I}_w}) \times \{\, w \,\} \\ &= C^{\mathcal{I}_w} \times \{\, w \,\} \cap D^{\mathcal{I}_w} \times \{\, w \,\} \\ &\overset{\text{I.H.}}{=} C^{\mathcal{I}^\times} \cap D^{\mathcal{I}^\times} \cap (\Delta^{\mathcal{I}} \times \{\, w \,\}) \\ &= (C \sqcap D)^{\mathcal{I}^\times} \cap (\Delta^{\mathcal{I}} \times \{\, w \,\}). \end{aligned}$$

For an existential restriction $\exists\, r.\, C$ the following equalities hold:

$$
\begin{aligned}
&(\exists\, r.\, C)^{\mathcal{I}_w} \times \{\, w \,\} \\
&= \{\, d \in \Delta^{\mathcal{I}} \mid \exists e \in \Delta^{\mathcal{I}} \colon (d,e) \in r^{\mathcal{I}_w} \text{ and } e \in C^{\mathcal{I}_w} \,\} \times \{\, w \,\} \\
&= \{\, (d,w) \mid \exists (e,w) \colon ((d,w),(e,w)) \in r^{\mathcal{I}^\times} \text{ and } (e,w) \in C^{\mathcal{I}_w} \times \{\, w \,\} \,\} \\
&\overset{\text{I.H.}}{=} \{\, (d,w) \mid \exists (e,w) \colon ((d,w),(e,w)) \in r^{\mathcal{I}^\times} \text{ and } (e,w) \in C^{\mathcal{I}^\times} \,\} \\
&= (\exists\, r.\, C)^{\mathcal{I}^\times} \cap (\Delta^{\mathcal{I}} \times \{\, w \,\}).
\end{aligned}
$$

2. analogously.
3. Using the first statement we may conclude that the following equalities hold:

$$
\begin{aligned}
&\mathbb{P}(C \sqsubseteq D) \\
&= \mathbb{P}\{\, w \in W \mid C^{\mathcal{I}_w} \times \{\, w \,\} \subseteq D^{\mathcal{I}_w} \times \{\, w \,\} \,\} \\
&= \mathbb{P}\{\, w \in W \mid C^{\mathcal{I}^\times} \cap (\Delta^{\mathcal{I}} \times \{\, w \,\}) \subseteq D^{\mathcal{I}^\times} \cap (\Delta^{\mathcal{I}} \times \{\, w \,\}) \,\} \\
&= \mathbb{P}\{\, w \in W \mid C^{\mathcal{I}^\times \mathcal{I}^\times \mathcal{I}^\times} \cap (\Delta^{\mathcal{I}} \times \{\, w \,\}) \subseteq D^{\mathcal{I}^\times \mathcal{I}^\times \mathcal{I}^\times} \cap (\Delta^{\mathcal{I}} \times \{\, w \,\}) \,\} \\
&= \mathbb{P}\{\, w \in W \mid C^{\mathcal{I}^\times \mathcal{I}^\times \mathcal{I}_w} \times \{\, w \,\} \subseteq D^{\mathcal{I}^\times \mathcal{I}^\times \mathcal{I}_w} \times \{\, w \,\} \,\} \\
&= \mathbb{P}(C^{\mathcal{I}^\times \mathcal{I}^\times} \sqsubseteq D^{\mathcal{I}^\times \mathcal{I}^\times}).
\end{aligned}
$$

The second equality follows analogously. □

For a probabilistic interpretation $\mathcal{I} = (\Delta^{\mathcal{I}}, \cdot^{\mathcal{I}}, W, \mathbb{P})$ and a set M of $\mathcal{EL}^{\perp}$-concept descriptions we define their *induced context* as the probabilistic formal context $\mathbb{K}_{\mathcal{I},M} := (\Delta^{\mathcal{I}}, M, W, I, \mathbb{P})$ where $(d, C, w) \in I$ iff $d \in C^{\mathcal{I}_w}$.

Lemma 17. *Let $\mathcal{I}$ be a probabilistic interpretation, M a set of $\mathcal{EL}^{\perp}$-concept descriptions, and $X, Y \subseteq M$. Then the probability of the implication $X \to Y$ in the induced context $\mathbb{K}_{\mathcal{I},M}$ equals the probability of the GCI $\bigsqcap X \sqsubseteq \bigsqcap Y$ in $\mathcal{I}$, i.e., it holds that $\mathbb{P}(X \to Y) = \mathbb{P}(\bigsqcap X \sqsubseteq \bigsqcap Y)$.*

Proof. The following equivalences are satisfied for all $Z \subseteq M$ and worlds $w \in W$:

$$d \in Z^{I_w} \Leftrightarrow \forall C \in Z \colon (d, C, w) \in I \Leftrightarrow \forall C \in Z \colon d \in C^{\mathcal{I}_w} \Leftrightarrow d \in (\bigsqcap Z)^{\mathcal{I}_w}.$$

Now consider two subsets $X, Y \subseteq M$, then it holds that

$$
\begin{aligned}
\mathbb{P}(X \to Y) &= \mathbb{P}\{\, w \in W \mid X^{I_w} \subseteq Y^{I_w} \,\} \\
&= \mathbb{P}\{\, w \in W \mid (\bigsqcap X)^{\mathcal{I}_w} \subseteq (\bigsqcap Y)^{\mathcal{I}_w} \,\} = \mathbb{P}(\bigsqcap X \sqsubseteq \bigsqcap Y).
\end{aligned}
$$

□

Analogously to [5], the context $\mathbb{K}_{\mathcal{I}}$ is defined as $\mathbb{K}_{\mathcal{I},M_{\mathcal{I}}}$ with the following attributes:

$$M_{\mathcal{I}} := \{\, \perp \,\} \cup N_C \cup \{\, \exists\, r.\, X^{\mathcal{I}_\varepsilon^\times} \mid \emptyset \neq X \subseteq \Delta^{\mathcal{I}} \times W_\varepsilon \,\}.$$

For an implication set $\mathcal{B}$ over a set M of $\mathcal{EL}^{\perp}$-concept descriptions we define its *induced TBox* by $\bigsqcap \mathcal{B} := \{\, \bigsqcap X \sqsubseteq \bigsqcap Y \mid X \to Y \in \mathcal{B} \,\}$.

Corollary 18. *If $\mathcal{B}$ contains an almost certain implicational base for $\mathbb{K}_{\mathcal{I}}$, then $\bigsqcap \mathcal{B}$ is complete for the almost certain GCIs of $\mathcal{I}$.*

Proof. We know that a GCI almost certainly holds in $\mathcal{I}$ if, and only if, it holds in $\mathcal{I}_{\varepsilon}^{\times}$. Let $\mathcal{B}' \subseteq \mathcal{B}$ be an almost certain implicational base for $\mathbb{K}_{\mathcal{I}}$, i.e., an implicational base for $(\mathbb{K}_{\mathcal{I}})_{\varepsilon}^{\times} = \mathbb{K}_{\mathcal{I}_{\varepsilon}^{\times}}$. Then according to Distel in [5, Theorem 5.12] it follows that the TBox $\bigsqcap \mathcal{B}'$ is a base of GCIs for $\mathcal{I}_{\varepsilon}^{\times}$, i.e., a base for the almost certain GCIs of $\mathcal{I}$. Consequently, $\bigsqcap \mathcal{B}$ is complete for the almost certain GCIs of $\mathcal{I}$. $\square$

Theorem 19. *Let $\mathcal{I}$ be a probabilistic interpretation, and $p \in [0,1]$ a threshold. If $\mathcal{B}$ is a probabilistic implicational base for $\mathbb{K}_{\mathcal{I}}$ and p that contains an almost certain implicational base for $\mathbb{K}_{\mathcal{I}}$, then $\bigsqcap \mathcal{B}$ is a probabilistic base of GCIs for $\mathcal{I}$ and p.*

Proof. Consider a GCI $\bigsqcap X \sqsubseteq \bigsqcap Y \in \bigsqcap \mathcal{B}$. Then Lemma 17 yields that the implication $X \to Y$ and the GCI $\bigsqcap X \sqsubseteq \bigsqcap Y$ have the same probability. Since $\mathcal{B}$ is a probabilistic implicational base for $\mathbb{K}_{\mathcal{I}}$ and p, we conclude that $\mathbb{P}(\bigsqcap X \sqsubseteq \bigsqcap Y) \geq p$ is satisfied.

Assume that $C \sqsubseteq D$ is an arbitrary GCI with probability $\geq p$. We have to show that $\bigsqcap \mathcal{B}$ entails $C \sqsubseteq D$. Let $\mathcal{J}$ be an arbitrary model of $\bigsqcap \mathcal{B}$. Consider an implication $X \to Y \in \mathcal{B}$, then $\bigsqcap X \sqsubseteq \bigsqcap Y \in \bigsqcap \mathcal{B}$ holds, and hence it follows that $(\bigsqcap X)^{\mathcal{J}} \subseteq (\bigsqcap Y)^{\mathcal{J}}$. Consequently, the implication $X \to Y$ holds in the induced context $\mathbb{K}_{\mathcal{J},M_{\mathcal{I}}}$. (We here mean the non-probabilistic formal context that is induced by a non-probabilistic interpretation, cf. [2, 3, 5].)

Furthermore, since all model-based most-specific concept descriptions of $\mathcal{I}_{\varepsilon}^{\times}$ are expressible in terms of $M_{\mathcal{I}}$, we have that $E \equiv \bigsqcap \pi_{M_{\mathcal{I}}}(E)$ holds for all mmscs E of $\mathcal{I}_{\varepsilon}^{\times}$, cf. [2, 3, 5]. Hence, we may conclude that

$$\begin{aligned}
\mathbb{P}(C \sqsubseteq D) &= \mathbb{P}(C^{\mathcal{I}_{\varepsilon}^{\times}\mathcal{I}_{\varepsilon}^{\times}} \sqsubseteq D^{\mathcal{I}_{\varepsilon}^{\times}\mathcal{I}_{\varepsilon}^{\times}}) \\
&= \mathbb{P}(\bigsqcap \pi_{M_{\mathcal{I}}}(C^{\mathcal{I}_{\varepsilon}^{\times}\mathcal{I}_{\varepsilon}^{\times}}) \sqsubseteq \bigsqcap \pi_{M_{\mathcal{I}}}(D^{\mathcal{I}_{\varepsilon}^{\times}\mathcal{I}_{\varepsilon}^{\times}})) \\
&= \mathbb{P}(\pi_{M_{\mathcal{I}}}(C^{\mathcal{I}_{\varepsilon}^{\times}\mathcal{I}_{\varepsilon}^{\times}}) \to \pi_{M_{\mathcal{I}}}(D^{\mathcal{I}_{\varepsilon}^{\times}\mathcal{I}_{\varepsilon}^{\times}})).
\end{aligned}$$

Consequently, $\mathcal{B}$ entails the implication $\pi_{M_{\mathcal{I}}}(C^{\mathcal{I}_{\varepsilon}^{\times}\mathcal{I}_{\varepsilon}^{\times}}) \to \pi_{M_{\mathcal{I}}}(D^{\mathcal{I}_{\varepsilon}^{\times}\mathcal{I}_{\varepsilon}^{\times}})$, hence it holds in $\mathbb{K}_{\mathcal{J},M_{\mathcal{I}}}$, and furthermore the GCI $C^{\mathcal{I}_{\varepsilon}^{\times}\mathcal{I}_{\varepsilon}^{\times}} \sqsubseteq D^{\mathcal{I}_{\varepsilon}^{\times}\mathcal{I}_{\varepsilon}^{\times}}$ holds in $\mathcal{J}$. As $\mathcal{J}$ is an arbitrary interpretation, $\bigsqcap \mathcal{B}$ entails $C^{\mathcal{I}_{\varepsilon}^{\times}\mathcal{I}_{\varepsilon}^{\times}} \sqsubseteq D^{\mathcal{I}_{\varepsilon}^{\times}\mathcal{I}_{\varepsilon}^{\times}}$.

Corollary 18 yields that $\bigsqcap \mathcal{B}$ is complete for the almost certain GCIs of $\mathcal{I}$. In particular, the GCI $C \sqsubseteq C^{\mathcal{I}_{\varepsilon}^{\times}\mathcal{I}_{\varepsilon}^{\times}}$ almost certainly holds in $\mathcal{I}$, and hence follows from $\bigsqcap \mathcal{B}$. We conclude that $\bigsqcap \mathcal{B} \models C \sqsubseteq D^{\mathcal{I}_{\varepsilon}^{\times}\mathcal{I}_{\varepsilon}^{\times}}$. Of course, the GCI $D^{\mathcal{I}_{\varepsilon}^{\times}\mathcal{I}_{\varepsilon}^{\times}} \sqsubseteq D$ holds in all interpretations. Finally, we conclude that $\bigsqcap \mathcal{B}$ entails $C \sqsubseteq D$. $\square$

Corollary 20. *Let $\mathcal{I}$ be a probabilistic interpretation, and $p \in [0,1]$ a threshold. Then $\bigsqcap \mathcal{B}_{\mathbb{K}_{\mathcal{I}},p}$ is a probabilistic base of GCIs for $\mathcal{I}$ and p where $\mathcal{B}_{\mathbb{K}_{\mathcal{I}},p}$ is defined as in Theorem 12.*

6 Conclusion

We have introduced the notion of a probabilistic formal context as a triadic context whose third dimension is a set of worlds equipped with a probability measure. Then the probability of implications in such probabilistic formal contexts was defined, and a construction of a base of implications whose probability exceeds a given threshold has been proposed, and its correctness has been verified. Furthermore, the results have been applied to the light-weight description logic $\mathcal{EL}^{\perp}$ with probabilistic interpretations, and so we formulated a method for the computation of a base of general concept inclusions whose probability satisfies a given lower threshold.

For finite input data-sets all of the provided constructions are computable. In particular, [3, 5] provide methods for the computation of model-based most-specific concept descriptions, and the algorithms in [6, 10] can be utilized to compute concept lattices and canonical implicational bases (or bases of GCIs, respectively).

The author thanks Sebastian Rudolph for proof reading and a fruitful discussion, and the anonymous reviewers for their constructive comments.

References

[1] Franz Baader et al., eds. *The Description Logic Handbook: Theory, Implementation, and Applications*. New York, NY, USA: Cambridge University Press, 2003.

[2] Daniel Borchmann. "Learning Terminological Knowledge with High Confidence from Erroneous Data". PhD thesis. TU Dresden, Germany, 2014.

[3] Daniel Borchmann, Felix Distel, and Francesco Kriegel. *Axiomatization of General Concept Inclusions from Finite Interpretations*. LTCS-Report 15-13. Chair for Automata Theory, Institute for Theoretical Computer Science, TU Dresden, Germany, 2015.

[4] Alexander V. Demin, Denis K. Ponomaryov, and Evgenii Vityaev. "Probabilistic Concepts in Formal Contexts". In: *Perspectives of Systems Informatics - 8th International Andrei Ershov Memorial Conference, PSI 2011, Novosibirsk, Russia, June 27-July 1, 2011, Revised Selected Papers*. Ed. by Edmund M. Clarke, Irina Virbitskaite, and Andrei Voronkov. Vol. 7162. Lecture Notes in Computer Science. Springer, 2011, pp. 394–410.

[5] Felix Distel. "Learning Description Logic Knowledge Bases from Data using Methods from Formal Concept Analysis". PhD thesis. TU Dresden, Germany, 2011.

[6] 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 Baris Sertkaya. Vol. 5986. Lecture Notes in Computer Science. Springer, 2010, pp. 312–340.

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

[8] 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.

[9] Francesco Kriegel. "Axiomatization of General Concept Inclusions in Probabilistic Description Logics". In: *Proceedings of the 38th German Conference on Artificial Intelligence, KI 2015, Dresden, Germany, September 21-25, 2015*. Vol. 9324. Lecture Notes in Artificial Intelligence. Springer Verlag, 2015.

[10] Francesco Kriegel. *NextClosures – Parallel Exploration of Constrained Closure Operators*. LTCS-Report 15-01. Chair for Automata Theory, Institute for Theoretical Computer Science, TU Dresden, Germany, 2015.

[11] Carsten Lutz and Lutz Schröder. "Probabilistic Description Logics for Subjective Uncertainty". In: *Principles of Knowledge Representation and Reasoning: Proceedings of the Twelfth International Conference, KR 2010, Toronto, Ontario, Canada, May 9-13, 2010*. Ed. by Fangzhen Lin, Ulrike Sattler, and Miroslaw Truszczynski. AAAI Press, 2010.

[12] Michael Luxenburger. "Implikationen, Abhängigkeiten und Galois Abbildungen". PhD thesis. TH Darmstadt, Germany, 1993.