TECHNISCHE UNIVERSITÄT DRESDEN

Faculty of Computer Science Institute of Theoretical Computer Science, Chair of Automata Theory

LEARNING DESCRIPTION LOGIC AXIOMS FROM DISCRETE PROBABILITY DISTRIBUTIONS OVER DESCRIPTION GRAPHS (EXTENDED VERSION)

LTCS-REPORT 18-12

Francesco Kriegel

francesco.kriegel@tu-dresden.de

27th February 2019

Learning Description Logic Axioms from Discrete Probability Distributions over Description Graphs

Francesco Kriegel[0000−0003−0219−0330]

Institute of Theoretical Computer Science,
Technische Universität Dresden, Dresden, Germany
francesco.kriegel@tu-dresden.de

Abstract. Description logics in their standard setting only allow for representing and reasoning with crisp knowledge without any degree of uncertainty. Of course, this is a serious shortcoming for use cases where it is impossible to perfectly determine the truth of a statement. For resolving this expressivity restriction, probabilistic variants of description logics have been introduced. Their model-theoretic semantics is built upon so-called probabilistic interpretations, that is, families of directed graphs the vertices and edges of which are labeled and for which there exists a probability measure on this graph family. Results of scientific experiments, e.g., in medicine, psychology, or biology, that are repeated several times can induce probabilistic interpretations in a natural way. In this document, we shall develop a suitable axiomatization technique for deducing terminological knowledge from the assertional data given in such probabilistic interpretations. More specifically, we consider a probabilistic variant of the description logic $\mathcal{EL}^{\bot}$, and provide a method for constructing a set of rules, so-called concept inclusions, from probabilistic interpretations in a sound and complete manner.

Keywords: Data mining · Knowledge acquisition · Probabilistic description logic · Knowledge base · Probabilistic interpretation · Concept inclusion

1 Introduction

Description Logics (abbrv. DLs) [2] are frequently used knowledge representation and reasoning formalisms with a strong logical foundation. In particular, these provide their users with automated inference services that can derive implicit knowledge from the explicitly represented knowledge. Decidability and computational complexity of common reasoning tasks have been widely explored for most DLs. Besides being used in various application domains, their most notable success is the fact that DLs constitute the logical underpinning of the *Web Ontology Language* (abbrv. OWL) and its profiles.

Logics in their standard form only allow for representing and reasoning with *crisp* knowledge without any degree of *uncertainty*. Of course, this is a serious shortcoming for use cases where it is impossible to perfectly determine the truth of a statement or where there exist degrees of truth. For resolving this expressivity restriction, probabilistic variants of logics have been introduced. A thorough article on extending first-order logics with means for representing and reasoning with probabilistic knowledge was published by Halpern [12]. In particular, Halpern explains why it is important to distinguish between two contrary types of probabilities: *statistical information* (type 1) and *degrees* *of belief* (type 2). The crucial difference between both types is that type-1 probabilities represent information about one particular world, the *real* world, and assume that there is a probability distribution on the objects, while type-2 probabilities represent information about a multi-world view such that there is a probability distribution on the set of possible worlds. Following his arguments and citing two of his examples, the first following statement can only be expressed in type-1 probabilistic logics and the second one is only expressible in type-2 probabilistic logics.

1. *"The probability that a randomly chosen bird will fly is greater than 0.9."*
2. *"The probability that Tweety (a particular bird) flies is greater than 0.9."*

Bacchus has published a further early work on probabilistic logics [3]. In particular, he defined the probabilistic first-order logic **Lp**, which allows to express various kinds of probabilistic/statistical knowledge: relative, interval, functional, conditional, independence. It is of type 1, since its semantics is based on probability measures over the domain of discourse (the objects). However, it also supports the deduction of degrees of belief (type 2) from given knowledge by means of an inference mechanism that is called *belief formation* and is based on an *inductive assumption of randomization*.

In [13], Heinsohn introduced the probabilistic description logic $\mathcal{ALCP}$ as an extension of $\mathcal{ALC}$. An $\mathcal{ALCP}$ ontology is a union of some acyclic $\mathcal{ALC}$ TBox and a finite set of so-called *p-conditionings*, which are expressions of the form $C \overset{[p,q]}{\to} D$ where C and D are Boolean combinations of concept names and where p and q are real numbers from the unit interval $[0, 1]$. $\mathcal{ALCP}$ allows for expressing type-1 probabilities only, since a p-conditioning $C \overset{[p,q]}{\to} D$ is defined to be valid in an interpretation $\mathcal{I}$ if it holds true that $p \leq {}^{|C^{\mathcal{I}} \cap D^{\mathcal{I}}|}/_{|C^{\mathcal{I}}|} \leq q$, that is, a uniform distribution on the domain of $\mathcal{I}$ is assumed and it is measured which percentage of the objects satisfying the premise C also satisfies the conclusion D. In particular, this means that only finite models are considered, which is a major restriction. Heinsohn shows how important reasoning problems (consistency and determining minimal p-conditionings) can be translated into problems of linear algebra. Please note that there is a strong correspondence with the notion of *confidence* of a concept inclusion as utilized by Borchmann in [4], cf. Section 2.

Another probabilistic extension of $\mathcal{ALC}$ was devised by Jaeger [14]: the description logic $\mathcal{PALC}$. Probabilities can be assigned to both terminological information and assertional information, rendering it a mixture of means for expressing type-1 and type-2 probabilities. A $\mathcal{PALC}$ ontology is a union of an acyclic $\mathcal{ALC}$ TBox, a finite set of *probabilistic terminological axioms* of the form $P(C \mid D) = p$, and a finite set of *probabilistic assertions* of the form $P(a \in C) = p$. The model-theoretic semantics are defined by extending the usual notion of a model with a probability measures: one measure μ dedicated to the probabilistic terminological axioms, and one measure ν_a dedicated to the probabilistic assertions for each individual a. Furthermore, these probability measures are defined on some finite subalgebra of the Lindenbaum-Tarski algebra of $\mathcal{ALC}$ concept descriptions that is generated by the concept descriptions occurring in the ontology, and it is further required that each ABox measure ν_a has minimal *cross entropy* to the TBox measure μ.

Lukasiewicz introduced in [23] the description logics P-*DL-Lite*, P-$\mathcal{SHIF}(\mathbf{D})$, and P-$\mathcal{SHOIN}(\mathbf{D})$ that are probabilistic extensions of *DL-Lite* and of the DLs underlying OWL Lite and OWL DL, respectively. We shall now briefly explain P-$\mathcal{SHOIN}(\mathbf{D})$, the others are analogous. It allows for expressing *conditional constraints* of the form $(\phi|\psi)[l, u]$ where ϕ and ψ are elements from some fixed, finite set $\mathcal{C}$ of $\mathcal{SHOIN}(\mathbf{D})$

concept descriptions, so-called *basic classification concepts*, and where l and u are real numbers from the unit interval $[0,1]$. Similar to $\mathcal{PALC}$, P-$\mathcal{SHOIN}(\mathbf{D})$ ontologies are unions of some $\mathcal{SHOIN}(\mathbf{D})$ ontology, a finite set of conditional constraints ($PTBox$) as probabilistic terminological knowledge, and a finite set of conditional constraints ($PABox$) as probabilistic assertional knowledge for each probabilistic individual. The semantics are then defined using interpretations that are additionally equipped with a discrete probability measure on the Lindenbaum-Tarski algebra generated by $\mathcal{C}$. Note that, in contrast to $\mathcal{PALC}$, there is only one probability measure available in each interpretation. While the terminological knowledge is, just like for $\mathcal{PALC}$, the default knowledge from which we only differ for a particular individual if the corresponding knowledge requires us to do so, the inference process is different, i.e., cross entropy is not utilized in any way. In order to allow for drawing inferences from a P-$\mathcal{SHOIN}(\mathbf{D})$ ontology, *lexicographic entailment* is defined for deciding whether a conditional constraint follows from the terminological part or for a certain individual. A thorough complexity analysis shows that the decision problems in these three logics are **NP**-complete, **EXP**-complete, and **NEXP**-complete, respectively.

Gutiérrez-Basulto, Jung, Lutz, and Schröder consider in [11] the probabilistic description logics Prob-$\mathcal{ALC}$ and Prob-$\mathcal{EL}$ where probabilities are always interpreted as degrees of belief (type 2). Among other language constructs, a new concept constructor is introduced that allows to probabilistically quantify a concept description. The semantics are based on multi-world interpretations where a discrete probability measure on the set of worlds is defined. Consistency and entailment is then defined just as usual, but using such probabilistic interpretations. A thorough investigation of computational complexity for various probabilistic extensions of DLs is provided: for instance, the common reasoning problems in Prob-$\mathcal{EL}$ and in Prob-$\mathcal{ALC}$ are **EXP**-complete, that is, not more expensive than the same problems in $\mathcal{ALC}$.

One should never mix up probabilistic and fuzzy variants of (description) logics. Although at first sight one could get the impression that both are suitable for any use cases where imprecise knowledge is to be represented and reasoned with, this is definitely not the case. A very simple argument against this is that in fuzzy logics we can easily evaluate conjunctions by means of the underlying fixed triangular norm (abbrv. t-norm), while it is not (always) possible to deduce the probability of a conjunction given the probabilities of the conjuncts. For instance, consider statements α and β. If both have fuzzy truth degree $1/2$ and the t-norm is Gödel's minimum, then $\alpha \wedge \beta$ has the fuzzy truth degree $1/2$ as well. In contrast, if both have probabilistic truth degree $1/2$, then the probability of $\alpha \wedge \beta$ might be any value in the interval $[0, 1/2]$, but without additional information we cannot bound it further or even determine it exactly.

Within this document, we make use of the syntax and semantics of [11]. It is easy to see that the probabilistic multi-world interpretations can be represented as families of directed graphs the vertices and edges of which are labeled and for which there exists a probability measure on this graph family. More specifically, we shall develop a suitable axiomatization technique for deducing terminological knowledge from the assertional data given in such probabilistic interpretations. In order to prevent the generated ontology from *overfitting*, a description logic that is not closed under Boolean operations is chosen. Since conjunction is essential, this implies that we leave out disjunction and negation. We consider a probabilistic variant $\mathsf{Prob}^{>}\mathcal{EL}^{\perp}$ of the description logic $\mathcal{EL}^{\perp}$, show that reasoning in $\mathsf{Prob}^{>}\mathcal{EL}^{\perp}$ is **EXP**-complete, and provide a method for constructing a set of rules, so-called concept inclusions, from probabilistic interpretations in a sound and complete manner. Within this document, the usage of probability restrictions is only allowed for lower probability bounds. This choice shall ease readability; it is not hard to verify that similar results can be obtained when additionally allowing for upper probability bounds.

Results of scientific experiments, e.g., in medicine, psychology, biology, finance, or economy, that are repeated several times can induce probabilistic interpretations in a natural way. Each repetition corresponds to a world, and the results of a particular repetition are encoded in the graph structure of that world. For instance, a researcher could collect data on consumption of the drugs ethanol and nicotine as well as on occurrence of serious health effects, e.g., cancer, psychological disorders, pneumonia, etc., such that a world corresponds to a single person and all worlds are equally likely. Then, the resulting probabilistic interpretation could be analyzed with the procedure described in the sequel of this document, which produces a *sound and complete axiomatization* of it. In particular, the outcome would then be a *logical-statistical evaluation* of the input data, and could include concept inclusions like the following.[1]

$$\begin{aligned}&\exists\,\mathsf{drinks}.\,(\mathsf{Alcohol}\sqcap\exists\,\mathsf{frequency}.\,\mathsf{TwiceAWeek})\\ \sqsubseteq\;&\mathsf{d}\geq {}^{1}\!/_{10}.\,\exists\,\mathsf{suffersFrom}.\,\mathsf{Cancer}\sqcap\mathsf{d}\geq {}^{1}\!/_{5}.\,\exists\,\mathsf{develops}.\,\mathsf{PsychologicalDisorder}\end{aligned}$$

$$\begin{aligned}&\exists\,\mathsf{smokes}.\,\mathsf{Tobacco}\\ \sqsubseteq\;&\mathsf{d}\geq {}^{1}\!/_{4}.\,\exists\,\mathsf{suffersFrom}.\,\mathsf{Cancer}\sqcap\mathsf{d}\geq {}^{1}\!/_{3}.\,\exists\,\mathsf{suffersFrom}.\,\mathsf{Pneumonia}\end{aligned}$$

The first one states that any person who drinks alcohol twice a week suffers from cancer with a probability of at least 10 % and develops some psychological disorder with a probability of at least 20 %; the second one expresses that each person smoking tobacco suffers from cancer with a probability of at least 25 % and suffers from pneumonia with a probability of at least $33{}^{1}\!/_{3}$ %.

However, one should be cautious when interpreting the results, since the procedure, like any other existing statistical evaluation techniques, cannot distinguish between *causality* and *correlation*. It might as well be the case that an application of our procedure yields concept inclusions of the following type.

$$\begin{aligned}&\mathsf{d}\geq {}^{1}\!/_{2}.\,\exists\,\mathsf{develops}.\,\mathsf{PsychologicalDisorder}\\ \sqsubseteq\;&\mathsf{d}\geq {}^{1}\!/_{3}.\,\exists\,\mathsf{drinks}.\,(\mathsf{Alcohol}\sqcap\exists\,\mathsf{frequency}.\,\mathsf{Daily})\end{aligned}$$

The above concept inclusion reads as follows: any person who develops a psychological disorder with a probability of at least 50 % drinks alcohol on a daily basis with a probability of at least $33{}^{1}\!/_{3}$ %.

It should further be mentioned that for evaluating observations by means of the proposed technique no hypotheses are necessary. Instead, the procedure simply provides a sound and complete axiomatization of the observations, and the output is, on the one hand, not too hard to be understood by humans (at least if, the probability depth is not set too high) and, on the other hand, well-suited to be further processed by a computer.

This document also resolves an issue found by Franz Baader with the techniques described by the author in [15, Sections 5 and 6]. In particular, the concept inclusion

[1] Please note that, although similar statements with adjusted probability bounds do hold true in real world, the mentioned statements are not taken from any publications in the medical or psychological domain. The author has simply read according Wikipedia articles and then wrote down the statements.

base proposed therein in Proposition 2 is only complete with respect to those probabilistic interpretations that are also quasi-uniform with a probability ε of each world. Herein, we describe a more sophisticated axiomatization technique of not necessarily quasi-uniform probabilistic interpretations that ensures completeness of the constructed concept inclusion base with respect to *all* probabilistic interpretations, but which, however, only allows for bounded nesting of probability restrictions. It is not hard to generalize the following results to a more expressive probabilistic description logic, for example to a probabilistic variant $\mathsf{Prob}^{>}\mathcal{M}$ of the description logic $\mathcal{M}$, for which an axiomatization technique is available [17]. That way, we can regain the same, or even a greater, expressivity as the author has tried to tackle in [15], but without the possibility to nest probability restrictions arbitrarily deep. A first step for resolving this issue has already been made in [20] where a nesting of probability restrictions is not supported. As a follow-up, we now expand on these results in [20] with the goal to allow for nesting of probabilistically quantified concept descriptions.

Due to space constraints, no proofs could be included here, but have rather been moved to a corresponding technical report [21].

2 Related Work

So far, several approaches for axiomatizing concept inclusions (abbrv. CIs) in different description logics have been developed, and many of these utilize sophisticated techniques from Formal Concept Analysis [8,9]: on the one hand, there is the so-called *canonical base*, cf. Guigues and Duquenne in [10], that provides a concise representation of the implicative theory of a formal context in a sound and complete manner and, on the other hand, the interactive algorithm *attribute exploration* exists, which guides an expert through the process of axiomatizing the theory of implications that are valid in a domain of interest, cf. Ganter in [7]. In particular, attribute exploration is an interactive variant of an algorithm for computing canonical bases [7], and it works as follows: the input is a formal context that only partially describes the domain of interest (that is, there may be implications that are not valid, but for which this partial description does not provide a counterexample), and during the run of the exploration process a minimal number of questions is enumerated and posed to the expert (such a question is an implication for which no counterexample has been explored, and the expert can either confirm its validity or provide a suitable counterexample). On termination, a minimal sound and complete representation of the theory of implications that are valid in the considered domain has been generated.

A first pioneering work on axiomatizing CIs in the description logic $\mathcal{FLE}$ has been developed by Rudolph [24], which allows for the exploration of a CI base for a given interpretation in a multi-step approach such that each step increases the role depth of concept descriptions occurring in the CIs. Later, a refined approach has been developed by Baader and Distel [1,6] for axiomatizing CI bases in the description logic $\mathcal{EL}^{\bot}$. They found techniques for computing and for exploring such bases that contain a *minimal* number of CIs and that are both sound and complete not only for those valid CIs up to certain role depth but instead for *all* valid ones. However, due to possible presence of cycles in the input interpretation they need to apply greatest fixed-point semantics; luckily, there is a finite *closure ordinal* for any finitely representable interpretation, that is, there is a certain role depth up to which the concept descriptions in the base can be unraveled to obtain a base for *all* valid CIs with respect to the standard semantics. Borchmann, Distel, and the author devised a variant of these techniques in [5] that circumvents the use of greatest fixed-point semantics, but which can only compute *minimal* CI bases that are sound and complete for all concept inclusions up to a set role depth—of course, if one chooses the closure ordinal as role-depth bound, then also these bases are sound and complete for *all* valid CIs w.r.t. standard semantics. Further variants that allow for the incorporation of background knowledge or allow for a more expressive description logic can be found in [16,17,22].

However, all of the mentioned approaches have in common that they heavily rely on the assumption that the given input interpretation to be axiomatized does not contain errors—otherwise these errors would be reflected in the constructed CI base. A reasonable solution avoiding this assumption has been proposed by Borchmann in [4]. He defined the notion of *confidence* as a statistical measure of validity of a CI in a given interpretation, and developed means for the computation and exploration of CI bases in $\mathcal{EL}^{\bot}$ that are sound and complete for those CIs the confidence of which exceeds a pre-defined threshold. Furthermore, in [19] the author defined the notion of *probability* of a CI in a probabilistic interpretation, and showed how corresponding bases of CIs exceeding a probability threshold can be constructed in a sound and complete manner. Both works have in common that they only allow for a statistical or probabilistic quantification of CIs, that is, it is only possible to assign a degree of truth to whole CIs, and not to concept descriptions occurring in these. For instance, one can express that $A \sqsubseteq \exists\, r.\, B$ has a confidence or probability of $2/3$, but one cannot write that every object which satisfies A with a probability of $5/6$ also satisfies $\exists\, r.\, B$ with a probability of $1/3$. As a solution to this, the author first considered in [18] implications over so-called probabilistic attributes in Formal Concept Analysis and showed how these can be axiomatized from a probabilistic formal context. Then in [20], his results have been extended to the probabilistic description logic $\mathsf{Prob}_1^{>}\mathcal{EL}^{\bot}$, a sublogic of $\mathsf{Prob}^{>}\mathcal{EL}^{\bot}$ that does not allow for nesting of probabilistically quantified concept descriptions. In Section 5 we shall expand on the results from [20] with the goal to constitute an effective procedure for axiomatizing CI bases in $\mathsf{Prob}^{>}\mathcal{EL}^{\bot}$, that is, we extend the procedure in [20] to allow for nesting of probabilistically quantified concept descriptions.

3 The Probabilistic Description Logic $\mathsf{Prob}^{>}\mathcal{EL}^{\bot}$

The probabilistic description logic $\mathsf{Prob}^{>}\mathcal{EL}^{\bot}$ constitutes an extension of the tractable description logic $\mathcal{EL}^{\bot}$ [2] that allows for expressing and reasoning with probabilities. More specifically, it is a sublogic of Prob-$\mathcal{EL}$ introduced by Gutiérrez-Basulto, Jung, Lutz, and Schröder in [11] in which only the relation symbols $>$ and $\geq$ are available for the probability restrictions, and in which the bottom concept description $\bot$ is present.[2] In the sequel of this section, we shall introduce the syntax and semantics of $\mathsf{Prob}^{>}\mathcal{EL}^{\bot}$. Furthermore, we will show that a common inference problem in $\mathsf{Prob}^{>}\mathcal{EL}^{\bot}$ is **EXP**-complete and, thus, more expensive than in $\mathcal{EL}^{\bot}$ where the same problem is **P**-complete.

Throughout the whole document, assume that Σ is an arbitrary but fixed *signature*, that is, Σ is a disjoint union of a set Σ_{C} of *concept names* and a set Σ_{R} of *role names*. Then, $\mathsf{Prob}^{>}\mathcal{EL}^{\bot}$ *concept descriptions* C over Σ may be inductively constructed by means of the following grammar rule (where $A \in \Sigma_{\mathsf{C}}$, $r \in \Sigma_{\mathsf{R}}$, $\gtrdot \in \{\geq, >\}$ and $p \in [0,1] \cap \mathbb{Q}$).

$$C ::= \bot \qquad \textit{(bottom concept description/contradiction)}$$

[2] We merely introduce $\bot$ as syntactic sugar; of course, it is semantically equivalent to the unsatisfiable probabilistic restriction $\mathsf{d} > 1.\, \top$.

$$
\begin{aligned}
&\mid \top && \textit{(top concept description/tautology)}\\
&\mid A && \textit{(concept name)}\\
&\mid C \sqcap C && \textit{(conjunction)}\\
&\mid \exists\, r.\, C && \textit{(existential restriction)}\\
&\mid \mathsf{d} \gtrdot p.\, C && \textit{(probability restriction)}
\end{aligned}
$$

Within this document, we stick to the default conventions and denote concept names by letters A or B, denote concept descriptions by letters C, D, E, etc., and denote role names by letters r, s, t, etc., each possibly with sub- or superscripts. Furthermore, we write $\mathsf{Prob}^{>}\mathcal{EL}^{\bot}(\Sigma)$ for the set of all $\mathsf{Prob}^{>}\mathcal{EL}^{\bot}$ concept descriptions over Σ. An $\mathcal{EL}^{\bot}$ *concept description* is a $\mathsf{Prob}^{>}\mathcal{EL}^{\bot}$ concept description not containing any subconcept of the form $\mathsf{d} \gtrdot p.\, C$,[3] and we shall write $\mathcal{EL}^{\bot}(\Sigma)$ for the set of all $\mathcal{EL}^{\bot}$ concept descriptions over Σ. If both C and D are concept descriptions, then the expression $C \sqsubseteq D$ is a *concept inclusion* (abbrv. CI), and the expression $C \equiv D$ is a *concept equivalence* (abbrv. CE). A *terminological box* (abbrv. TBox) is a finite set of CIs and CEs.

An example of a $\mathsf{Prob}^{>}\mathcal{EL}^{\bot}$ concept description is the following; it describes cats that are both alive and dead with a respective probability of at least 50 %. In particular, we could consider the below concept description as a formalization of the famous thought experiment *Schrödinger's Cat*.

$$
\begin{aligned}
\mathsf{Cat} \sqcap \mathsf{d} \geq 1/2.\, \exists\, \mathsf{hasPhysicalCondition}.\, \mathsf{Alive}&\\
\sqcap\, \mathsf{d} \geq 1/2.\, \exists\, \mathsf{hasPhysicalCondition}.\, \mathsf{Dead}&
\end{aligned}
\tag{1}
$$

The *probability depth* $\mathsf{pd}(C)$ of a $\mathsf{Prob}^{>}\mathcal{EL}^{\bot}$ concept description C is defined as the maximal nesting depth of probability restrictions within C, and we formally define it as follows: $\mathsf{pd}(A) \coloneqq 0$ for each $A \in \Sigma_{\mathsf{C}} \cup \{\bot, \top\}$, $\mathsf{pd}(C \sqcap D) \coloneqq \mathsf{pd}(C) \vee \mathsf{pd}(D)$,[4] $\mathsf{pd}(\exists\, r.\, C) \coloneqq \mathsf{pd}(C)$, and $\mathsf{pd}(\mathsf{d} \gtrdot p.\, C) \coloneqq 1 + \mathsf{pd}(C)$. Then, $\mathsf{Prob}_{n}^{>}\mathcal{EL}^{\bot}(\Sigma)$ denotes the set of all $\mathsf{Prob}^{>}\mathcal{EL}^{\bot}$ concept descriptions over Σ the probability depth of which does not exceed n.

Our considered logic $\mathsf{Prob}^{>}\mathcal{EL}^{\bot}$ possesses a model-theoretic semantics; so-called probabilistic interpretations function as models. Such a *probabilistic interpretation* over Σ is a tuple $\mathcal{I} \coloneqq (\Delta^{\mathcal{I}}, \Omega^{\mathcal{I}}, \cdot^{\mathcal{I}}, \mathbb{P}^{\mathcal{I}})$ that consists of a non-empty set $\Delta^{\mathcal{I}}$ of *objects*, called the *domain*, a non-empty, countable set $\Omega^{\mathcal{I}}$ of *worlds*, a discrete probability measure $\mathbb{P}^{\mathcal{I}}$ on $\Omega^{\mathcal{I}}$, and an *extension function* $\cdot^{\mathcal{I}}$ such that, for each world $\omega \in \Omega^{\mathcal{I}}$, any concept name $A \in \Sigma_{\mathsf{C}}$ is mapped to a subset $A^{\mathcal{I}(\omega)} \subseteq \Delta^{\mathcal{I}}$ and each role name $r \in \Sigma_{\mathsf{R}}$ is mapped to a binary relation $r^{\mathcal{I}(\omega)} \subseteq \Delta^{\mathcal{I}} \times \Delta^{\mathcal{I}}$. We remark that the discrete probability measure is a mapping $\mathbb{P}^{\mathcal{I}} \colon \wp(\Omega^{\mathcal{I}}) \to [0,1]$ which satisfies $\mathbb{P}^{\mathcal{I}}(\emptyset) = 0$ and $\mathbb{P}^{\mathcal{I}}(\Omega^{\mathcal{I}}) = 1$, and which is σ*-additive*, that is, for all countable families $(\, U_n \mid n \in \mathbb{N}\,)$ of pairwise disjoint sets $U_n \subseteq \Omega^{\mathcal{I}}$ it holds true that $\mathbb{P}^{\mathcal{I}}(\bigcup\{\, U_n \mid n \in \mathbb{N}\,\}) = \sum(\, \mathbb{P}^{\mathcal{I}}(U_n) \mid n \in \mathbb{N}\,)$.

We shall follow the assumption in [11, Section 2.6] and consider only probabilistic interpretations without any infinitely improbable worlds, i.e., which do not contain any world $\omega \in \Omega^{\mathcal{I}}$ with $\mathbb{P}^{\mathcal{I}}\{\omega\} = 0$. Furthermore, a probabilistic interpretation $\mathcal{I}$ is

[3] The author does not use the denotation $P_{\gtrdot p}C$ for probability restrictions as in [11], since quantifiers are usually single letters rotated by 180 degrees.

[4] Note that $\vee$ denotes the binary supremum operator for numbers, which here coincides with the maximum operator, since there are only finitely many arguments.

finitely representable if $\Delta^{\mathcal{I}}$ is finite, $\Omega^{\mathcal{I}}$ is finite, the *active signature*

$$\Sigma^{\mathcal{I}} \coloneqq \{\, \sigma \mid \sigma \in \Sigma \text{ and } \sigma^{\mathcal{I}(\omega)} \neq \emptyset \text{ for some } \omega \in \Omega^{\mathcal{I}} \,\}$$

is finite, and if $\mathbb{P}^{\mathcal{I}}$ has only rational values.

It is easy to see that, for any probabilistic interpretation $\mathcal{I}$, each world $\omega \in \Omega^{\mathcal{I}}$ can be represented as a labeled, directed graph: the node set is the domain $\Delta^{\mathcal{I}}$, the edge set is $\bigcup\{\, r^{\mathcal{I}(\omega)} \mid r \in \Sigma_{\mathsf{R}} \,\}$, any node δ is labeled with all those concept names A that satisfy $\delta \in A^{\mathcal{I}(\omega)}$, and any edge (δ, ϵ) has a role name r as a label if $(\delta, \epsilon) \in r^{\mathcal{I}(\omega)}$ holds true. That way, we can regard probabilistic interpretations also as discrete probability distributions over description graphs.

Later, we will also use the notion of *interpretations*, which are the models upon which the semantics of $\mathcal{EL}^{\bot}$ is built. Put simply, these are probabilistic interpretations with only one world, that is, these are tuples $\mathcal{I} \coloneqq (\Delta^{\mathcal{I}}, \cdot^{\mathcal{I}})$ where $\Delta^{\mathcal{I}}$ is a non-empty set of *objects*, called *domain*, and where $\cdot^{\mathcal{I}}$ is an *extension function* that maps concept names $A \in \Sigma_{\mathsf{C}}$ to subsets $A^{\mathcal{I}} \subseteq \Delta^{\mathcal{I}}$ and maps role names $r \in \Sigma_{\mathsf{R}}$ to binary relations $r^{\mathcal{I}} \subseteq \Delta^{\mathcal{I}} \times \Delta^{\mathcal{I}}$.

Let $\mathcal{I}$ be a probabilistic interpretation. Then, the *extension* $C^{\mathcal{I}(\omega)}$ of a $\mathsf{Prob}^{>}\mathcal{EL}^{\bot}$ concept description C in a world ω of $\mathcal{I}$ is recursively defined as follows.

$$\bot^{\mathcal{I}(\omega)} \coloneqq \emptyset \qquad \top^{\mathcal{I}(\omega)} \coloneqq \Delta^{\mathcal{I}} \qquad (C \sqcap D)^{\mathcal{I}(\omega)} \coloneqq C^{\mathcal{I}(\omega)} \cap D^{\mathcal{I}(\omega)}$$

$$(\exists\, r.\, C)^{\mathcal{I}(\omega)} \coloneqq \{\, \delta \mid \delta \in \Delta^{\mathcal{I}},\ (\delta, \epsilon) \in r^{\mathcal{I}(\omega)},\ \text{and } \epsilon \in C^{\mathcal{I}(\omega)} \text{ for some } \epsilon \in \Delta^{\mathcal{I}} \,\}$$

$$(\mathsf{d} \geqslant p.\, C)^{\mathcal{I}(\omega)} \coloneqq \{\, \delta \mid \delta \in \Delta^{\mathcal{I}} \text{ and } \mathbb{P}^{\mathcal{I}}\{\delta \in C^{\mathcal{I}}\} \geqslant p \,\}$$

In the last of the above definitions we use the abbreviation

$$\{\delta \in C^{\mathcal{I}}\} \coloneqq \{\, \omega \mid \omega \in \Omega^{\mathcal{I}} \text{ and } \delta \in C^{\mathcal{I}(\omega)} \,\}.$$

All but the last formula can be used in a similar manner to define the *extension* $C^{\mathcal{I}}$ of an $\mathcal{EL}^{\bot}$ concept description C in an interpretation $\mathcal{I}$. Please note that, in accordance with [11], there is nothing wrong with the above definition of extensions; in particular, it is true that the extension $(\mathsf{d} \geqslant p.\, C)^{\mathcal{I}(\omega)}$ of a probabilistic restriction $\mathsf{d} \geqslant p.\, C$ is indeed independent of the concrete world ω, i.e., it holds true that $(\mathsf{d} \geqslant p.\, C)^{\mathcal{I}(\omega)} = (\mathsf{d} \geqslant p.\, C)^{\mathcal{I}(\psi)}$ whenever ω and ψ are arbitrary worlds in $\Omega^{\mathcal{I}}$. This is due to the intended meaning of $\mathsf{d} \geqslant p.\, C$: it describes the class of objects for which the probability of being a C is $\geqslant p$. As a probabilistic interpretation $\mathcal{I}$ provides a multi-world view where probabilities can be assigned to sets of worlds, the probability of an object $\delta \in \Delta^{\mathcal{I}}$ being a C is defined as the probability of the set of all those worlds in which δ is some C, just like we have defined it above. We shall elaborate on this again as soon as we have defined validity of concept inclusions in probabilistic interpretations, and mind that extensions of a fixed probabilistic quantification are equal in all worlds.

A toy example of a probabilistic interpretation is $\mathcal{I}_{\mathsf{Cat}}$ shown in Figure 1. As one quickly verifies, only the object SchrödingersCat belongs to the extension of the concept description from Equation (1).

A concept inclusion $C \sqsubseteq D$ or a concept equivalence $C \equiv D$ is *valid* in $\mathcal{I}$ if, for each world $\omega \in \Omega^{\mathcal{I}}$, it holds true that $C^{\mathcal{I}(\omega)} \subseteq D^{\mathcal{I}(\omega)}$ or $C^{\mathcal{I}(\omega)} = D^{\mathcal{I}(\omega)}$, respectively, and we shall then write $\mathcal{I} \models C \sqsubseteq D$ or $\mathcal{I} \models C \equiv D$, respectively. Furthermore, $\mathcal{I}$ is a *model* of a TBox $\mathcal{T}$, denoted as $\mathcal{I} \models \mathcal{T}$, if every concept inclusion or concept equivalence in $\mathcal{T}$ is valid in $\mathcal{I}$. A TBox $\mathcal{T}$ *entails* a concept inclusion or concept equivalence α,

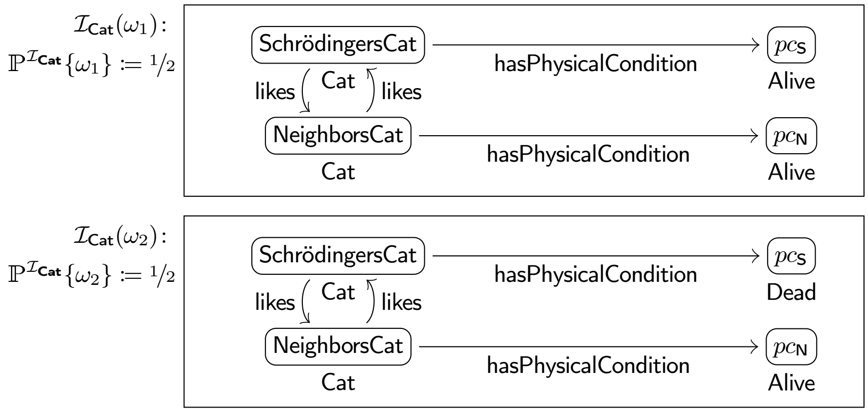

Figure 1. An exemplary probabilistic interpretation

symbolized by $\mathcal{T} \models \alpha$, if α is valid in every model of $\mathcal{T}$. In case $\mathcal{T} \models C \sqsubseteq D$ we say that C is *subsumed* by D with respect to $\mathcal{T}$, and if $\mathcal{T} \models C \equiv D$, then we call C and D *equivalent* to each other with respect to $\mathcal{T}$.

If $\mathcal{Y}$ is either an interpretation or a terminological box and $\leq$ is a suitable relation symbol, e.g., one of $\sqsubseteq$, $\equiv$, $\sqsupseteq$, then we may also use the denotation $C \leq_{\mathcal{Y}} D$ instead of $\mathcal{Y} \models C \leq D$ and, analogously, we may write $C \not\leq_{\mathcal{Y}} D$ for $\mathcal{Y} \not\models C \leq D$.

Considering again the above definition of extensions of concept descriptions together with the just defined validity of concept inclusions, we can also justify the independence of $(\mathsf{d} \mathrel{\geqslant} p.\, C)^{\mathcal{I}(\omega)}$ from world ω in the following way. Fix some probabilistic interpretation as well as some concept inclusion $C \sqsubseteq D$. Since concept inclusions are terminological axioms, and as such represent knowledge that globally holds true, it is only natural to say that $C \sqsubseteq D$ is valid in $\mathcal{I}$ if, and only if, $C \sqsubseteq D$ is valid in each *slice* $\mathcal{I}(\omega)$ for any world $\omega \in \Omega^{\mathcal{I}}$—apparently, this is what we have defined above. If $C = \mathsf{d} \geq p.\, C'$ and $D = \mathsf{d} \geq q.\, D'$ are probabilistic restrictions, then the intended meaning of the concept inclusion $\mathsf{d} \geq p.\, C' \sqsubseteq \mathsf{d} \geq q.\, D'$ is that any object being a C' with probability at least p is also a D' with probability q or greater. Of course, this is equivalent to $(\mathsf{d} \geq p.\, C')^{\mathcal{I}(\omega)} \subseteq (\mathsf{d} \geq q.\, D')^{\mathcal{I}(\omega)}$ for each world $\omega \in \Omega^{\mathcal{I}}$, that is, to $\mathcal{I} \models \mathsf{d} \geq p.\, C' \sqsubseteq \mathsf{d} \geq q.\, D'$. This argumentation can now be extended to the general case where C and D are arbitrary $\mathsf{Prob}^{>}\mathcal{EL}^{\perp}$ concept descriptions.

For some of the upcoming proofs we need the following lemma, which expresses the fact that the probabilistic restriction constructor—more specifically, each mapping $C \mapsto \mathsf{d} \mathrel{\geqslant} p.\, C$ for $\mathrel{\geqslant} \in \{\geq, >\}$ and $p \in [0,1] \cap \mathbb{Q}$—is monotonic.

Lemma 1. *Consider a $\mathsf{Prob}^{>}\mathcal{EL}^{\perp}$ terminological box $\mathcal{T}$ and a $\mathsf{Prob}^{>}\mathcal{EL}^{\perp}$ concept inclusion $C \sqsubseteq D$. Then, $C \sqsubseteq_{\mathcal{T}} D$ implies $\mathsf{d} \mathrel{\geqslant} p.\, C \sqsubseteq_{\mathcal{T}} \mathsf{d} \mathrel{\geqslant} p.\, D$ for any $\mathrel{\geqslant} \in \{\geq, >\}$ and for each $p \in [0,1] \cap \mathbb{Q}$.*

Proof. Fix some model $\mathcal{I}$ of $\mathcal{T}$ and let $\mathbb{P}^{\mathcal{I}}\{\delta \in C^{\mathcal{I}}\} \mathrel{\geqslant} p$ for an object $\delta \in \Delta^{\mathcal{I}}$. From $\mathcal{T} \models C \sqsubseteq D$ we infer that, for each world $\omega \in \Omega^{\mathcal{I}}$, it holds true that $\delta \in C^{\mathcal{I}(\omega)}$ implies $\delta \in D^{\mathcal{I}(\omega)}$. Consequently, we have that $\{\delta \in C^{\mathcal{I}}\} \subseteq \{\delta \in D^{\mathcal{I}}\}$ and, thus, $\mathbb{P}^{\mathcal{I}}\{\delta \in D^{\mathcal{I}}\} \mathrel{\geqslant} p$ due to the monotonicity of the probability measure $\mathbb{P}^{\mathcal{I}}$. $\square$

For a complexity analysis, we consider the following *subsumption problem*.

Instance: Let $\mathcal{T}$ be a TBox and let $C \sqsubseteq D$ be a concept inclusion.
Question: Is C subsumed by D w.r.t. $\mathcal{T}$, i.e., does $C \sqsubseteq_{\mathcal{T}} D$ hold true?

The next proposition shows that this problem is **EXP**-complete and, consequently, more expensive than deciding subsumption w.r.t. a TBox in its non-probabilistic sibling $\mathcal{EL}^{\bot}$—a problem which is well-known to be **P**-complete. We conclude that reasoning in $\mathsf{Prob}^{>}\mathcal{EL}^{\bot}$ is worst-case intractable, while reasoning in $\mathcal{EL}^{\bot}$ is always tractable.

Proposition 2. *In $\mathsf{Prob}^{>}\mathcal{EL}^{\bot}$, the subsumption problem is* **EXP***-complete.*

Proof. Containment in **EXP** follows from [11, Theorem 3] and the fact that $\mathsf{Prob}^{>}\mathcal{EL}^{\bot}$ is a sublogic of $\mathsf{Prob}\text{-}\mathcal{ALC}$. **EXP**-hardness is a consequence of [11, Theorem 13 and Sections A.1, A.2, and A.3], where **EXP**-hardness of the logics $\mathsf{Prob}\text{-}\mathcal{EL}^{\sim p}$ for $\sim \in \{\geq, >\}$, that is, of sublogics of $\mathsf{Prob}^{>}\mathcal{EL}^{\bot}$, is demonstrated. □

4 Concept Inclusion Bases in $\mathcal{EL}^{\bot}$

When developing a method for axiomatizing $\mathsf{Prob}^{>}\mathcal{EL}^{\bot}$ concept inclusions valid in a given probabilistic interpretation in the next section, we will use techniques for axiomatizing $\mathcal{EL}^{\bot}$ CIs valid in an interpretation as developed by Baader and Distel in [1,6] for greatest fixed-point semantics, and as adjusted by Borchmann, Distel, and the author in [5] for the role-depth-bounded case. A brief introduction is as follows. A *concept inclusion base* for an interpretation $\mathcal{I}$ is a TBox $\mathcal{T}$ such that, for each CI $C \sqsubseteq D$, it holds true that $C \sqsubseteq_{\mathcal{I}} D$ if, and only if, $C \sqsubseteq_{\mathcal{T}} D$. For each finite interpretation $\mathcal{I}$ with finite active signature, there is a *canonical base* $\mathsf{Can}(\mathcal{I})$ with respect to greatest fixed-point semantics, which contains a minimal number of CIs among all concept inclusion bases for $\mathcal{I}$, cf. [6, Corollary 5.13 and Theorem 5.18], and similarly there is a minimal *canonical base* $\mathsf{Can}(\mathcal{I}, d)$ with respect to an upper bound $d \in \mathbb{N}$ on the role depths, cf. [5, Theorem 4.32]. The construction of both canonical bases is built upon the notion of a *model-based most specific concept description* (abbrv. MMSC), which, for an interpretation $\mathcal{I}$ and some subset $\Xi \subseteq \Delta^{\mathcal{I}}$, is a concept description C such that $\Xi \subseteq C^{\mathcal{I}}$ and, for each concept description D, it holds true that $\Xi \subseteq D^{\mathcal{I}}$ implies $C \sqsubseteq_{\emptyset} D$. These exist either if greatest fixed-point semantics is applied (in order to be able to express cycles present in $\mathcal{I}$) or if the role depth of C is bounded by some $d \in \mathbb{N}$, and these are then denoted as $\Xi^{\mathcal{I}}$ or $\Xi^{\mathcal{I}_d}$, respectively. These mappings $\cdot^{\mathcal{I}} \colon \wp(\Delta^{\mathcal{I}}) \to \mathcal{EL}^{\bot}_{\mathsf{gfp}}(\Sigma)$ and $\cdot^{\mathcal{I}_d} \colon \wp(\Delta^{\mathcal{I}}) \to \mathcal{EL}^{\bot}_{d}(\Sigma)$ are the adjoints of the respective extension functions $\cdot^{\mathcal{I}} \colon \mathcal{EL}^{\bot}_{\mathsf{gfp}}(\Sigma) \to \wp(\Delta^{\mathcal{I}})$ and $\cdot^{\mathcal{I}} \colon \mathcal{EL}^{\bot}_{d}(\Sigma) \to \wp(\Delta^{\mathcal{I}})$, and the pair of both constitutes a *Galois connection*, cf. [6, Lemma 4.1] and [5, Lemmas 4.3 and 4.4], respectively.

As a variant of these two approaches, the author presented in [16] a method for constructing canonical bases relative to an existing TBox. If $\mathcal{I}$ is an interpretation and $\mathcal{B}$ is a TBox such that $\mathcal{I} \models \mathcal{B}$, then a *concept inclusion base* for $\mathcal{I}$ *relative* to $\mathcal{B}$ is a TBox $\mathcal{T}$ such that, for each CI $C \sqsubseteq D$, it holds true that $C \sqsubseteq_{\mathcal{I}} D$ if, and only if, $C \sqsubseteq_{\mathcal{T} \cup \mathcal{B}} D$. The corresponding *canonical base* is denoted as $\mathsf{Can}(\mathcal{I}, \mathcal{B})$, cf. [16, Theorem 1].

So far, the complexity of computing CI bases in the description logic $\mathcal{EL}^{\bot}$ has not been determined. Using simple arguments, one could only infer that the canonical base $\mathsf{Can}(\mathcal{I})$ can be computed in double exponential time with respect to the cardinality of the domain $\Delta^{\mathcal{I}}$. However, since we want to determine the computational complexity of the task of constructing CI bases in the probabilistic description logic $\mathsf{Prob}^{>}\mathcal{EL}^{\bot}$, which we will describe and prove in the next section and which we will build on top of means for computing such bases in $\mathcal{EL}^{\bot}$, we cite a recent answer from the author to this open question in the following proposition.

[22, Proposition 2]. *For each finitely representable interpretation $\mathcal{I}$, its canonical base $\mathsf{Can}(\mathcal{I})$ can be computed in deterministic exponential time with respect to the cardinality of the domain $\Delta^{\mathcal{I}}$. Furthermore, there are finitely representable interpretations $\mathcal{I}$ for which a concept inclusion base cannot be encoded in polynomial space w.r.t. $|\Delta^{\mathcal{I}}|$.*

It is not hard to adapt this result to the role-depth-bounded case; one can show that computing $\mathsf{Can}(\mathcal{I}, d)$ can be done in deterministic exponential time w.r.t. $|\Delta^{\mathcal{I}}|$ and d.

5 Axiomatization of Concept Inclusions in $\mathsf{Prob}_n^{>}\mathcal{EL}^{\bot}$

In this section, we shall develop an effective method for axiomatizing $\mathsf{Prob}_n^{>}\mathcal{EL}^{\bot}$ concept inclusions which are valid in a given finitely representable probabilistic interpretation. After defining the appropriate notion of a *concept inclusion base*, we show how this problem can be tackled using the aforementioned existing results on computing concept inclusion bases in $\mathcal{EL}^{\bot}$ from Section 4. More specifically, we devise an extension of the given signature by finitely many probability restrictions $\mathsf{d} \geqslant p.\, C$ that are treated as additional concept names, and we define so-called *scalings* $\mathcal{I}_n$ of the input probabilistic interpretation $\mathcal{I}$ which are (single-world) interpretations that suitably interpret these new concept names and, furthermore, such that there is a correspondence between $\mathsf{Prob}_n^{>}\mathcal{EL}^{\bot}$ CIs valid in $\mathcal{I}$ and CIs valid in $\mathcal{I}_n$. This very correspondence makes it possible to utilize the above mentioned techniques for axiomatizing CIs in $\mathcal{EL}^{\bot}$.

Definition 3. *A* $\mathsf{Prob}_n^{>}\mathcal{EL}^{\bot}$ concept inclusion base *for a probabilistic interpretation $\mathcal{I}$ is a* $\mathsf{Prob}_n^{>}\mathcal{EL}^{\bot}$ *terminological box $\mathcal{T}$ which is* sound *for $\mathcal{I}$, that is, $C \sqsubseteq_{\mathcal{T}} D$ implies $C \sqsubseteq_{\mathcal{I}} D$ for each* $\mathsf{Prob}_n^{>}\mathcal{EL}^{\bot}$ *concept inclusion $C \sqsubseteq D$,*[5] *and* complete *for $\mathcal{I}$, that is, $C \sqsubseteq_{\mathcal{I}} D$ only if $C \sqsubseteq_{\mathcal{T}} D$ for any* $\mathsf{Prob}_n^{>}\mathcal{EL}^{\bot}$ *concept inclusion $C \sqsubseteq D$.*

The following definition is to be read inductively, that is, initially some objects are defined for the probability depth $n = 0$, and if the objects are defined for the probability depth n, then these are used to define the next objects for the probability depth $n + 1$.

A first important step is to significantly reduce the possibilities of concept descriptions occurring as a filler in the probability restrictions, that is, of fillers C in expressions $\mathsf{d} \geqslant p.\, C$. As it turns out, it suffices to consider only those fillers that are model-based most specific concept descriptions of some suitable *scaling* of the given probabilistic interpretation $\mathcal{I}$. We shall demonstrate that there are only finitely many such fillers—provided that the given probabilistic interpretation $\mathcal{I}$ is finitely representable.

As next step, we restrict the probability bounds p occurring in probability restrictions $\mathsf{d} \geqslant p.\, C$. Apparently, it is sufficient to consider only those values p that can occur when evaluating the extension of $\mathsf{Prob}_{n+1}^{>}\mathcal{EL}^{\bot}$ concept descriptions in $\mathcal{I}$, which, obviously, are the values $\mathbb{P}^{\mathcal{I}}\{\delta \in C^{\mathcal{I}}\}$ for any $\delta \in \Delta^{\mathcal{I}}$ and any $C \in \mathsf{Prob}_n^{>}\mathcal{EL}^{\bot}(\Sigma)$. In the sequel of this section we will see that there are only finitely many such probability bounds if $\mathcal{I}$ is finitely representable.

[5] Of course, soundness is equivalent to $\mathcal{I} \models \mathcal{T}$.

Having found a finite number of representatives for probability bounds as well as a finite number of fillers to be used in probability restrictions for each probability depth n, we now show that we can treat these finitely many concept descriptions as concept names of a signature Γ_n extending Σ in a way such that any $\mathsf{Prob}_n^{>}\mathcal{EL}^{\bot}$ concept inclusion is valid in $\mathcal{I}$ if, and only if, that concept inclusion projected onto the extended signature Γ_n is valid in a suitable *scaling* of $\mathcal{I}$ that interprets Γ_n.

Definition 4. *Fix some probabilistic interpretation $\mathcal{I}$ over a signature Σ. Then, we define the following objects Γ_n, $\mathcal{I}_n$, and $P_{\mathcal{I},n}$ by simultaneous induction over $n \in \mathbb{N}$.*

1. *The nth signature Γ_n is inductively defined as follows. We set $(\Gamma_0)_\mathsf{C} \coloneqq \Sigma_\mathsf{C}$ and $(\Gamma_0)_\mathsf{R} \coloneqq \Sigma_\mathsf{R}$. The subsequent signatures are then obtained in the following way.*[6]

$$(\Gamma_{n+1})_\mathsf{C} \coloneqq (\Gamma_n)_\mathsf{C} \cup \left\{ \mathsf{d} \geq p.\, X^{\mathcal{I}_n} \;\middle|\; \begin{array}{l} p \in P_{\mathcal{I},n} \setminus \{0\},\ X \subseteq \Delta^{\mathcal{I}} \times \Omega^{\mathcal{I}}, \\ \text{and } \bot \not\equiv_\emptyset X^{\mathcal{I}_n} \not\equiv_\emptyset \top \end{array} \right\}$$
$$(\Gamma_{n+1})_\mathsf{R} \coloneqq \Sigma_\mathsf{R}$$

2. *The nth* scaling *of $\mathcal{I}$ is defined as the interpretation $\mathcal{I}_n$ over Γ_n that has the following components.*

$$\Delta^{\mathcal{I}_n} \coloneqq \Delta^{\mathcal{I}} \times \Omega^{\mathcal{I}}$$
$$\cdot^{\mathcal{I}_n} \colon \begin{cases} A \mapsto \{\, (\delta,\omega) \mid \delta \in A^{\mathcal{I}(\omega)} \,\} & \text{for each } A \in (\Gamma_n)_\mathsf{C} \\ r \mapsto \{\, ((\delta,\omega),(\epsilon,\omega)) \mid (\delta,\epsilon) \in r^{\mathcal{I}(\omega)} \,\} & \text{for each } r \in (\Gamma_n)_\mathsf{R} \end{cases}$$

3. *The nth set $P_{\mathcal{I},n}$ of probability values for $\mathcal{I}$ is given as follows.*

$$P_{\mathcal{I},n} \coloneqq \{\, \mathbb{P}^{\mathcal{I}}\{\delta \in C^{\mathcal{I}}\} \mid \delta \in \Delta^{\mathcal{I}} \text{ and } C \in \mathsf{Prob}_n^{>}\mathcal{EL}^{\bot}(\Sigma) \,\}$$

Furthermore, for each $p \in [0,1)$, we define $(p)^+_{\mathcal{I},n}$ as the next value in $P_{\mathcal{I},n}$ above p, that is, we set

$$(p)^+_{\mathcal{I},n} \coloneqq \bigwedge \{\, q \mid q \in P_{\mathcal{I},n} \text{ and } q > p \,\}.$$

Of course, we have that $\{0,1\} \subseteq P_{\mathcal{I},n}$ for each $n \in \mathbb{N}$. Note that $\mathcal{I}_{n+1}$ extends $\mathcal{I}_n$ by also interpreting the additional concept names in $(\Gamma_{n+1})_\mathsf{C} \setminus (\Gamma_n)_\mathsf{C}$, that is, the restriction $\mathcal{I}_{n+1}\restriction_{\Gamma_n}$ equals $\mathcal{I}_n$. Similarly, $\mathcal{I}_n\restriction_{\Gamma_m}$ and $\mathcal{I}_m$ are equal if $m \leq n$.

As explained earlier, it suffices to only consider fillers in probabilistic restrictions that are model-based most specific concept descriptions. More specifically, the following holds true.

Lemma 5. *Consider a probabilistic interpretation $\mathcal{I}$ and a concept description $\mathsf{d} \gtrdot p.\, C$ such that $C \in \mathcal{EL}^{\bot}(\Gamma_n)$ for some $n \in \mathbb{N}$. Then, the concept equivalence $\mathsf{d} \gtrdot p.\, C \equiv \mathsf{d} \gtrdot p.\, C^{\mathcal{I}_n\mathcal{I}_n}$ is valid in $\mathcal{I}$.*

[6] The mapping $\cdot^{\mathcal{I}} \colon \wp(\Delta^{\mathcal{I}}) \to \mathcal{EL}^{\bot}(\Sigma)$ for some (non-probabilistic) interpretation $\mathcal{I}$ has been introduced in Section 4.

Proof. Using structural induction on $\mathcal{EL}^\bot$ concept descriptions C over Σ, it can be proven that $C^{\mathcal{I}(\omega)} \times \{\omega\} = C^{\mathcal{I}_n} \cap (\Delta^{\mathcal{I}} \times \{\omega\})$ is satisfied for each world $\omega \in \Omega^{\mathcal{I}}$ and for every $n \in \mathbb{N}$, cf. [19, Lemma 16]. For extending this result to $\mathsf{Prob}_n^{>}\mathcal{EL}^\bot$ concept descriptions that are in $\mathcal{EL}^\bot(\Gamma_n)$, we need to show a further inductive case for the probability restrictions $\mathsf{d} \mathbin{⩾} p.\, C$. As one quickly verifies, the following equalities hold true for all probability restrictions $\mathsf{d} \mathbin{⩾} p.\, C \in (\Gamma_n)_{\mathsf{C}}$.

$$(\mathsf{d} \mathbin{⩾} p.\, C)^{\mathcal{I}(\omega)} \times \{\omega\} = \{\, (\delta, \omega) \mid \mathbb{P}^{\mathcal{I}}\{\delta \in C^{\mathcal{I}}\} \mathbin{⩾} p \,\} = (\mathsf{d} \mathbin{⩾} p.\, C)^{\mathcal{I}_n} \cap (\Delta^{\mathcal{I}} \times \{\omega\})$$

It follows that, for any $n \in \mathbb{N}$ and for each concept description $C \in \mathcal{EL}^\bot(\Gamma_n)$, it holds true that $C^{\mathcal{I}(\omega)} = \pi_1(C^{\mathcal{I}_n} \cap (\Delta^{\mathcal{I}} \times \{\omega\}))$ (where π_1 projects pairs to their first components). By applying well-known properties of Galois connections we obtain that $C^{\mathcal{I}(\omega)} = C^{\mathcal{I}_n\mathcal{I}_n\mathcal{I}(\omega)}$, and so $\mathbb{P}^{\mathcal{I}}\{\delta \in C^{\mathcal{I}}\} = \mathbb{P}^{\mathcal{I}}\{\delta \in C^{\mathcal{I}_n\mathcal{I}_n\mathcal{I}}\}$ holds true. □

The above lemma does not hold true for arbitrary fillers C, but only for fillers that can (syntactically) also be seen as $\mathcal{EL}^\bot$ concept descriptions over Γ_n. However, this does not cause any problems, since we can simply project any other filler onto this signature Γ_n. In particular, we define projections of arbitrary $\mathsf{Prob}_n^{>}\mathcal{EL}^\bot$ concept descriptions onto the signature Γ_n in the following manner.

Definition 6. *Fix some $n \in \mathbb{N}$ as well as a probabilistic interpretation $\mathcal{I}$. The* nth projection $\pi_{\mathcal{I},n}(C)$ *of a $\mathsf{Prob}_n^{>}\mathcal{EL}^\bot$ concept description C with respect to $\mathcal{I}$ is obtained from C by replacing subconcepts of the form $\mathsf{d} \mathbin{⩾} p.\, D$ with suitable elements from $(\Gamma_n)_{\mathsf{C}}$ and, more specifically, we recursively define it as follows. We set $\pi_{\mathcal{I},0}(C) \coloneqq C$ for each concept description $C \in \mathcal{EL}^\bot(\Sigma)$. The subsequent projections are then given in the following manner.*

$$\begin{aligned}
\pi_{\mathcal{I},n+1}(A) &\coloneqq A && \text{if } A \in \Sigma_{\mathsf{C}} \cup \{\bot, \top\} \\
\pi_{\mathcal{I},n+1}(C \sqcap D) &\coloneqq \pi_{\mathcal{I},n+1}(C) \sqcap \pi_{\mathcal{I},n+1}(D) \\
\pi_{\mathcal{I},n+1}(\exists\, r.\, C) &\coloneqq \exists\, r.\, \pi_{\mathcal{I},n+1}(C) \\
\pi_{\mathcal{I},n+1}(\mathsf{d} \mathbin{⩾} p.\, C) &\coloneqq \begin{cases}
\bot & \text{if } {⩾}\, p = {>}\, 1 \\
\top & \text{else if } {⩾}\, p = {\geq}\, 0 \\
\bot & \text{else if } (\pi_{\mathcal{I},n}(C))^{\mathcal{I}_{n+1}\mathcal{I}_{n+1}} \equiv_\emptyset \bot \\
\top & \text{else if } (\pi_{\mathcal{I},n}(C))^{\mathcal{I}_{n+1}\mathcal{I}_{n+1}} \equiv_\emptyset \top \\
\mathsf{d} \geq p.(\pi_{\mathcal{I},n}(C))^{\mathcal{I}_{n+1}\mathcal{I}_{n+1}} & \text{else if } {⩾} = {\geq} \text{ and } p \in P_{\mathcal{I},n+1} \\
\mathsf{d} \geq (p)^{+}_{\mathcal{I},n+1}.(\pi_{\mathcal{I},n}(C))^{\mathcal{I}_{n+1}\mathcal{I}_{n+1}} & \text{else}
\end{cases}
\end{aligned}$$

For technical details, we introduce further notation: we denote by $\pi'_{\mathcal{I},n+1}(\mathsf{d} \mathbin{⩾} p.\, C)$ and $\pi''_{\mathcal{I},n+1}(\mathsf{d} \mathbin{⩾} p.\, C)$ the concept description that is obtained from the projection $\pi_{\mathcal{I},n+1}(\mathsf{d} \mathbin{⩾} p.\, C)$ by replacing $(\pi_{\mathcal{I},n}(C))^{\mathcal{I}_{n+1}\mathcal{I}_{n+1}}$ with $\pi_{\mathcal{I},n}(C)$ and C, respectively.

Usually, projection mappings in mathematics are *idempotent*. It is easy to verify by induction over n that this also holds true for our projection mappings $\pi_{\mathcal{I},n}$ which we have just defined. This justifies the naming choice. Furthermore, we can show that the mappings $\pi_{\mathcal{I},n}$ are *intensive*, i.e., projecting some $\mathsf{Prob}_n^{>}\mathcal{EL}^\bot$ concept description C onto the nth signature Γ_n yields a more specific concept description, cf. the next lemma. Furthermore, the mappings $\pi_{\mathcal{I},n}$ are monotonic—a fact that can be proven by induction over n as well. As a corollary, it follows that each mapping $\pi_{\mathcal{I},n}$ is a *kernel operator*. However, please just take this as a side note, since we do not need the two additional properties of idempotency and monotonicity within this document.

Lemma 7. *Assume that $\mathcal{I}$ is a probabilistic interpretation, let $n \in \mathbb{N}$, and fix some $\mathsf{Prob}_n^{>}\mathcal{EL}^{\bot}$ concept description C. Then, it holds true that $\pi_{\mathcal{I},n}(C) \sqsubseteq_\emptyset C$.*

Proof. We prove by induction over k that $\pi_{\mathcal{I},k}(C) \sqsubseteq_\emptyset C$ for any $k \leq n$ and every $C \in \mathsf{Prob}_k^{>}\mathcal{EL}^{\bot}(\Sigma)$. Due to equality of C and its 0th projection $\pi_{\mathcal{I},0}(C)$, the base case for $k = 0$ is obvious. For the inductive step for $k+1$, we continue with an (inner) induction on the structure of C. All cases, except the case for a probability restriction $\mathsf{d} \mathrel{\triangleright} p.\, D$, are easy. We claim that $\pi_{\mathcal{I},k+1}(\mathsf{d} \mathrel{\triangleright} p.\, D) \sqsubseteq_\emptyset \mathsf{d} \mathrel{\triangleright} p.\, D$. Since $(\pi_{\mathcal{I},k}(D))^{\mathcal{I}_{k+1}\mathcal{I}_{k+1}} \sqsubseteq_\emptyset \pi_{\mathcal{I},k}(D)$, it follows by means of Lemma 1 that

$$\pi_{\mathcal{I},k+1}(\mathsf{d} \mathrel{\triangleright} p.\, D) \sqsubseteq_\emptyset \pi'_{\mathcal{I},k+1}(\mathsf{d} \mathrel{\triangleright} p.\, D). \tag{2}$$

The induction hypothesis together with Lemma 1 implies that

$$\pi'_{\mathcal{I},k+1}(\mathsf{d} \mathrel{\triangleright} p.\, D) \sqsubseteq_\emptyset \pi''_{\mathcal{I},k+1}(\mathsf{d} \mathrel{\triangleright} p.\, D) \tag{3}$$

and, furthermore, it is apparent that

$$\pi''_{\mathcal{I},k+1}(\mathsf{d} \mathrel{\triangleright} p.\, D) \sqsubseteq_\emptyset \mathsf{d} \mathrel{\triangleright} p.\, D. \tag{4}$$

In summary, Equations (2) – (4) show that $\pi_{\mathcal{I},k+1}(\mathsf{d} \mathrel{\triangleright} p.\, D)$ is subsumed by $\mathsf{d} \mathrel{\triangleright} p.\, D$ with respect to the empty TBox. □

As a crucial observation regarding projections, we see that—within our given probabilistic interpretation $\mathcal{I}$—we do not have to distinguish between any $\mathsf{Prob}_n^{>}\mathcal{EL}^{\bot}$ concept description C and its nth projection $\pi_{\mathcal{I},n}(C)$, since the upcoming lemma shows that both always possess the same extension in each world of $\mathcal{I}$. Simply speaking, the signatures Γ_n contain enough building bricks to describe anything that happens within $\mathcal{I}$ up to a probability depth of n.

Lemma 8. *Assume that $\mathcal{I}$ is a probabilistic interpretation, let $n \in \mathbb{N}$, and consider some $\mathsf{Prob}_n^{>}\mathcal{EL}^{\bot}$ concept description C. Then, C and its nth projection $\pi_{\mathcal{I},n}(C)$ have the same extension in any world of $\mathcal{I}$.*

Proof. We show the claim by means of an outer induction on n and an inner induction on the structure of C. The outer base case for $n = 0$ is trivial, since then C and its projection $\pi_{\mathcal{I},0}(C)$ are equal. We proceed with the outer inductive case for $n+1$ and a structural induction on C. Then, according to the definition of an $(n+1)$th projection, the only nontrivial case considers probabilistic restrictions occurring in C. It is readily verified that $\mathsf{d} \mathrel{\triangleright} p.\, E$ and $\pi''_{\mathcal{I},n+1}(\mathsf{d} \mathrel{\triangleright} p.\, E)$ have the same extension in each world of $\mathcal{I}$. Using the fact that E is a $\mathsf{Prob}_n^{>}\mathcal{EL}^{\bot}$ concept description together with the outer induction hypothesis, we infer that $\pi''_{\mathcal{I},n+1}(\mathsf{d} \mathrel{\triangleright} p.\, E)$ and $\pi'_{\mathcal{I},n+1}(\mathsf{d} \mathrel{\triangleright} p.\, E)$ have the same extension in each world of $\mathcal{I}$ too. An application of Lemma 5 now yields that, in every world of $\mathcal{I}$, also the extensions of $\pi'_{\mathcal{I},n+1}(\mathsf{d} \mathrel{\triangleright} p.\, E)$ and $\pi_{\mathcal{I},n+1}(\mathsf{d} \mathrel{\triangleright} p.\, E)$ are the same. □

As a last important statement on the properties of the projection mappings, we now demonstrate that validity of some concept inclusion $C \sqsubseteq D$ with a probability depth not exceeding n is equivalent to validity of the projected concept inclusion $\pi_{\mathcal{I},n}(C) \sqsubseteq \pi_{\mathcal{I},n}(D)$ in the scaling $\mathcal{I}_n$. This is a key lemma for the upcoming construction of a concept inclusion base for $\mathcal{I}$.

Lemma 9. *Let $n \in \mathbb{N}$, and consider a probabilistic interpretation $\mathcal{I}$ as well as some $\mathsf{Prob}_n^{>}\mathcal{EL}^{\bot}$ concept inclusion $C \sqsubseteq D$. Then, $C \sqsubseteq D$ is valid in $\mathcal{I}$ if, and only if, the nth projected concept inclusion $\pi_{\mathcal{I},n}(C) \sqsubseteq \pi_{\mathcal{I},n}(D)$ is valid in the nth scaling $\mathcal{I}_n$.*

Proof. We start with observing that, according to Lemma 8, $C \sqsubseteq D$ is valid in $\mathcal{I}$ if, and only if, $\pi_{\mathcal{I},n}(C) \sqsubseteq \pi_{\mathcal{I},n}(D)$ is valid in $\mathcal{I}$. Then, the equivalence of $\mathcal{I} \models \pi_{\mathcal{I},n}(C) \sqsubseteq \pi_{\mathcal{I},n}(D)$ and $\mathcal{I}_n \models \pi_{\mathcal{I},n}(C) \sqsubseteq \pi_{\mathcal{I},n}(D)$ follows from the very definition of the nth scaling $\mathcal{I}_n$ and the fact that the projections $\pi_{\mathcal{I},n}(C)$ and $\pi_{\mathcal{I},n}(D)$ can be interpreted as $\mathcal{EL}^{\bot}$ concept descriptions over Γ_n. ☐

Now we go on to considering the sets $P_{\mathcal{I},n}$ of essential probability values. As we have already claimed, these sets are always finite—provided that the fixed probabilistic interpretation is finitely representable. In order to prove this, we need the following statement.

Lemma 10. *For each probabilistic interpretation $\mathcal{I}$ and any $n \in \mathbb{N}$, the following equation is satisfied.*

$$P_{\mathcal{I},n} = \{\, \mathbb{P}^{\mathcal{I}}\{\delta \in X^{\mathcal{I}_n \mathcal{I}}\} \mid \delta \in \Delta^{\mathcal{I}} \text{ and } X \subseteq \Delta^{\mathcal{I}} \times \Omega^{\mathcal{I}} \,\}$$

Proof. Fix some $\mathsf{Prob}_n^{>}\mathcal{EL}^{\bot}$ concept description C. From Lemma 8 we infer that

$$\mathbb{P}^{\mathcal{I}}\{\delta \in C^{\mathcal{I}}\} = \mathbb{P}^{\mathcal{I}}\{\delta \in (\pi_{\mathcal{I},n}(C))^{\mathcal{I}}\}$$

holds true. In the proof of Lemma 5 we have shown that $\mathbb{P}^{\mathcal{I}}\{\delta \in D^{\mathcal{I}}\} = \mathbb{P}^{\mathcal{I}}\{\delta \in D^{\mathcal{I}_n \mathcal{I}_n \mathcal{I}}\}$ holds true for each $\mathsf{Prob}^{>}\mathcal{EL}^{\bot}$ concept description D over Σ that (syntactically) is also an $\mathcal{EL}^{\bot}$ concept description over Γ_n. Hence, we can use this identity for $D \coloneqq \pi_{\mathcal{I},n}(C)$, which yields that

$$\mathbb{P}^{\mathcal{I}}\{\delta \in (\pi_{\mathcal{I},n}(C))^{\mathcal{I}}\} = \mathbb{P}^{\mathcal{I}}\{\delta \in (\pi_{\mathcal{I},n}(C))^{\mathcal{I}_n \mathcal{I}_n \mathcal{I}}\}.$$

Of course, we have that $(\pi_{\mathcal{I},n}(C))^{\mathcal{I}_n} \subseteq \Delta^{\mathcal{I}} \times \Omega^{\mathcal{I}}$.

Since C is an arbitrary concept description, we conclude that $P_{\mathcal{I},n}$ is a subset of $\{\, \mathbb{P}^{\mathcal{I}}\{\delta \in X^{\mathcal{I}_n \mathcal{I}}\} \mid \delta \in \Delta^{\mathcal{I}}$ and $X \subseteq \Delta^{\mathcal{I}} \times \Omega^{\mathcal{I}} \,\}$. The reverse set inclusion is trivial. ☐

For most, if not all, practical use case we can argue that the given probabilistic interpretation $\mathcal{I}$ can be assumed as finitely representable. Utilizing some of our previous results then implies that each nth scaling of $\mathcal{I}$ is finitely representable as well. More specifically, the following is satisfied.

Corollary 11. *If $\mathcal{I}$ is a finitely representable probabilistic interpretation, then it holds true that, for each $n \in \mathbb{N}$, the subset $\Gamma_n \setminus \Sigma$ of the nth signature is finite, the nth scaling $\mathcal{I}_n$ is finite and has a finite active signature, and the nth set $P_{\mathcal{I},n}$ of probability values is finite and satisfies $P_{\mathcal{I},n} \subseteq \mathbb{Q}$.*

As already mentioned in Sections 2 and 4, we want to make use of existing techniques that allow for axiomatizing interpretations in the description logic $\mathcal{EL}^{\bot}$. In order to do so, we need to be sure that the semantics of $\mathcal{EL}^{\bot}$ and its probabilistic sibling $\mathsf{Prob}^{>}\mathcal{EL}^{\bot}$ are not too different, or expressed alternatively, that there is a suitable correspondence between (non-probabilistic) entailment in $\mathcal{EL}^{\bot}$ and (probabilistic) entailment in $\mathsf{Prob}^{>}\mathcal{EL}^{\bot}$. A more sophisticated formulation is presented in the following lemma.

Lemma 12. *Let $\mathcal{T}$ be a $\mathsf{Prob}^{>}\mathcal{EL}^{\bot}$ TBox, and assume that $\mathcal{B}$ is a set that consists of tautological $\mathsf{Prob}^{>}\mathcal{EL}^{\bot}$ concept inclusions, i.e., $\emptyset \models \mathcal{B}$. If $C \sqsubseteq D$ is a $\mathsf{Prob}^{>}\mathcal{EL}^{\bot}$ concept inclusion that is entailed by $\mathcal{T} \cup \mathcal{B}$ with respect to non-probabilistic entailment, then $C \sqsubseteq D$ is also entailed by $\mathcal{T}$ with respect to probabilistic entailment.*

Proof. Fix some signature Σ, let $\mathcal{T} \cup \mathcal{B} \models C \sqsubseteq D$ (non-probabilistically), and consider some probabilistic interpretation $\mathcal{I}$ such that $\mathcal{I} \models \mathcal{T}$. Of course, it also holds true that $\mathcal{I} \models \mathcal{B}$. We extend Σ to the signature Γ defined as follows: $\Gamma_{\mathsf{C}} := \Sigma_{\mathsf{C}} \cup \{\, \mathsf{d} \gtrdot p.\, C \mid \gtrdot \in \{\geq, >\},\ p \in [0,1] \cap \mathbb{Q},\ \text{and } C \in \mathsf{Prob}^{>}\mathcal{EL}^{\bot}(\Sigma) \,\}$ and $\Gamma_{\mathsf{R}} := \Sigma_{\mathsf{R}}$. It is apparent that, syntactically, $\mathcal{EL}^{\bot}(\Gamma) = \mathsf{Prob}^{>}\mathcal{EL}^{\bot}(\Sigma)$ holds true. Furthermore, we define the interpretation $\mathcal{J}$ where $\Delta^{\mathcal{J}} := \Delta^{\mathcal{I}} \times \Omega^{\mathcal{I}}$, $A^{\mathcal{J}} := \{\, (\delta, \omega) \mid \delta \in A^{\mathcal{I}(\omega)} \,\}$ for each $A \in \Gamma_{\mathsf{C}}$, and $r^{\mathcal{J}} := \{\, ((\delta, \omega), (\epsilon, \omega)) \mid (\delta, \epsilon) \in r^{\mathcal{I}(\omega)} \,\}$ for each $r \in \Gamma_{\mathsf{R}}$. We can show with structural induction that $C^{\mathcal{J}} = \bigcup \{\, C^{\mathcal{I}(\omega)} \times \{\omega\} \mid \omega \in \Omega^{\mathcal{I}} \,\}$ for any $C \in \mathcal{EL}^{\bot}(\Gamma)$. Consequently, $\mathcal{I} \models E \sqsubseteq F$ is equivalent to $\mathcal{J} \models E \sqsubseteq F$ for each $\mathsf{Prob}^{>}\mathcal{EL}^{\bot}$ concept inclusion $E \sqsubseteq F$. It follows that $\mathcal{J} \models \mathcal{T} \cup \mathcal{B}$, and we infer that $\mathcal{J} \models C \sqsubseteq D$, which implies that $\mathcal{I} \models C \sqsubseteq D$. As $\mathcal{I}$ is an arbitrary model of $\mathcal{T}$, we can safely conclude that $\mathcal{T} \models C \sqsubseteq D$ (probabilistically). □

As final step, we show that each concept inclusion base of the probabilistic scaling $\mathcal{I}_n$ induces a $\mathsf{Prob}_n^{>}\mathcal{EL}^{\bot}$ concept inclusion base of $\mathcal{I}$. While soundness is easily verified, completeness follows from the fact that $C \sqsubseteq_{\mathcal{T}} \pi_{\mathcal{I},n}(C) \sqsubseteq_{\mathcal{T}} \pi_{\mathcal{I},n}(D) \sqsubseteq_{\emptyset} D$ holds true for every valid $\mathsf{Prob}_n^{>}\mathcal{EL}^{\bot}$ concept inclusion $C \sqsubseteq D$ of $\mathcal{I}$.

Theorem 13. *Fix a number $n \in \mathbb{N}$ and some finitely representable probabilistic interpretation $\mathcal{I}$. If $\mathcal{T}_n$ is a concept inclusion base for the nth scaling $\mathcal{I}_n$ with respect to some set $\mathcal{B}_n$ of tautological $\mathsf{Prob}_n^{>}\mathcal{EL}^{\bot}$ concept inclusions used as background knowledge, then the following terminological box $\mathcal{T}$ is a $\mathsf{Prob}_n^{>}\mathcal{EL}^{\bot}$ concept inclusion base for $\mathcal{I}$.*

$$\begin{aligned} \mathcal{T} &:= \mathcal{T}_n \cup \bigcup \{\, \mathcal{U}_{\mathcal{I},\ell} \mid \ell \in \{1, \ldots, n\} \,\} \quad \textit{where} \\ \mathcal{U}_{\mathcal{I},\ell} &:= \{\, \mathsf{d} > p.\, X^{\mathcal{I}_\ell \mathcal{I}_\ell} \sqsubseteq \mathsf{d} \geq (p)^{+}_{\mathcal{I},\ell}.\, X^{\mathcal{I}_\ell \mathcal{I}_\ell} \mid p \in P_{\mathcal{I},\ell} \setminus \{1\} \textit{ and } X \subseteq \Delta^{\mathcal{I}} \times \Omega^{\mathcal{I}} \,\} \end{aligned}$$

Proof. Soundness is apparently satisfied. We proceed with showing completeness; thus, fix some $\mathsf{Prob}_n^{>}\mathcal{EL}^{\bot}$ concept inclusion $C \sqsubseteq D$ which is valid in $\mathcal{I}$. We shall demonstrate the validity of the following subsumptions.

$$C \sqsubseteq_{\mathcal{T}} \pi_{\mathcal{I},n}(C) \sqsubseteq_{\mathcal{T}} \pi_{\mathcal{I},n}(D) \sqsubseteq_{\emptyset} D$$

According to Lemma 7, it holds true that $\pi_{\mathcal{I},n}(D) \sqsubseteq_{\emptyset} D$. Lemma 9 immediately yields that $\pi_{\mathcal{I},n}(C) \sqsubseteq \pi_{\mathcal{I},n}(D)$ is valid in the nth scaling $\mathcal{I}_n$. Since $\mathcal{T}_n$ is complete for $\mathcal{I}_n$ relative to $\mathcal{B}_n$, it follows that $\mathcal{T}_n \cup \mathcal{B}_n$ entails $\pi_{\mathcal{I},n}(C) \sqsubseteq \pi_{\mathcal{I},n}(D)$ with respect to non-probabilistic entailment and, thus, $\mathcal{T}$ entails $\pi_{\mathcal{I},n}(C) \sqsubseteq \pi_{\mathcal{I},n}(D)$ with respect to probabilistic entailment.

It remains to show that $C \sqsubseteq_{\mathcal{T}} \pi_{\mathcal{I},n}(C)$ holds true; we do so by proving with an induction on k that $C \sqsubseteq_{\mathcal{T}} \pi_{\mathcal{I},k}(C)$ holds true for each $k \leq n$ and $C \in \mathsf{Prob}_k^{>}\mathcal{EL}^{\bot}(\Sigma)$. The base case where $k = 0$ is obvious, since each $\mathsf{Prob}_0^{>}\mathcal{EL}^{\bot}$ concept description C equals its 0th projection $\pi_{\mathcal{I},0}(C)$. For the inductive case for $k+1$, we proceed with an inner induction on the structure of C. The only non-trivial case considers probability restrictions $\mathsf{d} \gtrdot p.\,E$. Of course, E is then a $\mathsf{Prob}_k^{>}\mathcal{EL}^{\bot}$ concept description, and the induction hypothesis yields that $E \sqsubseteq_{\mathcal{T}} \pi_{\mathcal{I},k}(E)$. As an immediate consequence from Lemma 1 we infer that

$$\mathsf{d} \gtrdot p.\,E \sqsubseteq_{\mathcal{T}} \mathsf{d} \gtrdot p.\,\pi_{\mathcal{I},k}(E). \tag{5}$$

Furthermore, the concept inclusion $\pi_{\mathcal{I},k}(E) \sqsubseteq (\pi_{\mathcal{I},k}(E))^{\mathcal{I}_{k+1}\mathcal{I}_{k+1}}$ is valid in $\mathcal{I}_{k+1}$. Since $\mathcal{I}_n\restriction_{\Gamma_{k+1}} = \mathcal{I}_{k+1}$ holds true, and both $\pi_{\mathcal{I},k}(E)$ and $(\pi_{\mathcal{I},k}(E))^{\mathcal{I}_{k+1}\mathcal{I}_{k+1}}$ (syntactically) are $\mathcal{EL}^{\bot}$ concept descriptions over $\Gamma_{k+1} \subseteq \Gamma_n$, we conclude that the considered concept inclusion $\pi_{\mathcal{I},k}(E) \sqsubseteq (\pi_{\mathcal{I},k}(E))^{\mathcal{I}_{k+1}\mathcal{I}_{k+1}}$ is valid in $\mathcal{I}_n$. Consequently, this CI is (non-probabilistically) entailed by $\mathcal{T}_n \cup \mathcal{B}_n$ and, according to Lemma 12, it is hence (probabilistically) entailed by $\mathcal{T}$. An application of Lemma 1 now shows that

$$\mathsf{d} \gtrdot p.\,\pi_{\mathcal{I},k}(E) \sqsubseteq_{\mathcal{T}} \mathsf{d} \gtrdot p.(\pi_{\mathcal{I},k}(E))^{\mathcal{I}_{k+1}\mathcal{I}_{k+1}}. \tag{6}$$

Obviously, the subset $\mathcal{U}_{\mathcal{I},k+1}$ of $\mathcal{T}$ entails the concept inclusion

$$\mathsf{d} \gtrdot p.(\pi_{\mathcal{I},k}(E))^{\mathcal{I}_{k+1}\mathcal{I}_{k+1}} \sqsubseteq \pi''_{\mathcal{I},k+1}(\mathsf{d} \gtrdot p.(\pi_{\mathcal{I},k}(E))^{\mathcal{I}_{k+1}\mathcal{I}_{k+1}}),$$

and since the latter concept description is exactly $\pi_{\mathcal{I},k+1}(\mathsf{d} \gtrdot p.\,E)$ we infer that

$$\mathsf{d} \gtrdot p.(\pi_{\mathcal{I},k}(E))^{\mathcal{I}_{k+1}\mathcal{I}_{k+1}} \sqsubseteq_{\mathcal{T}} \pi_{\mathcal{I},k+1}(\mathsf{d} \gtrdot p.\,E). \tag{7}$$

Putting the results from Equations (5) – (7) together now demonstrates the truth of the claim that $\mathsf{d} \gtrdot p.\,E$ is subsumed by $\pi_{\mathcal{I},k+1}(\mathsf{d} \gtrdot p.\,E)$ with respect to $\mathcal{T}$. □

As already mentioned in Section 4 and according to [16], a suitable such concept inclusion base $\mathcal{T}_n$ for the nth scaling $\mathcal{I}_n$ with respect to background knowledge $\mathcal{B}_n$ exists and can be computed effectively, namely the canonical base $\mathsf{Can}(\mathcal{I}_n, \mathcal{B}_n)$. This enables us to immediately draw the following conclusion.

Corollary 14. *Let $\mathcal{I}$ be a finitely representable probabilistic interpretation, fix some $n \in \mathbb{N}$, and let $\mathcal{B}_n$ denote the set of all $\mathcal{EL}^{\bot}$ concept inclusions over Γ_n that are tautological with respect to probabilistic entailment, i.e., are valid in every probabilistic interpretation. Then, the* canonical base *for $\mathcal{I}$ and probability depth n that is defined as*

$$\mathsf{Can}(\mathcal{I}, n) \coloneqq \mathsf{Can}(\mathcal{I}_n, \mathcal{B}_n) \cup \bigcup \{\, \mathcal{U}_{\mathcal{I},\ell} \mid \ell \in \{1, \ldots, n\} \,\}$$

is a $\mathsf{Prob}_n^{>}\mathcal{EL}^{\bot}$ concept inclusion base for $\mathcal{I}$, and it can be computed effectively.

Eventually, we close our investigations with a complexity analysis of the problem of actually computing the canonical base $\mathsf{Can}(\mathcal{I}, n)$. As it turns out, this computation is—in terms of computational complexity—not more expensive than the corresponding axiomatization task in $\mathcal{EL}^{\bot}$, cf. [22, Proposition 2]; both in $\mathcal{EL}^{\bot}$ and in $\mathsf{Prob}_n^{>}\mathcal{EL}^{\bot}$ concept inclusion bases can be computed in exponential time.

However, this result only holds true if we dispense with the pre-computation of the tautological background knowledge $\mathcal{B}_n$ at all. First of all, Γ_n can have exponential size, and there are d-exponentially many $\mathcal{EL}^\bot$ concept descriptions over some fixed signature. Thus, a naïve enumeration of $\mathcal{B}_n$ is too expensive. However, also computing the implicative background knowledge on the FCA side with utilizing some $\mathsf{Prob}^{>}\mathcal{EL}^\bot$ reasoner on demand is too expensive as well. This is due to the fact that one needs to enumerate all implications $\mathbf{C} \to \{D\}$ where $\mathbf{C} \cup \{D\}$ is a subset of the attribute set of the induced formal context of $\mathcal{I}_d$. On the one hand, the number of such implications is exponential in the size of the attribute set and this attribute set can contain exponentially many concept descriptions that can each have an exponential size. On the other hand, we have already seen that deciding subsumption in $\mathsf{Prob}^{>}\mathcal{EL}^\bot$ is an **EXP**-complete problem. Even a more sophisticated approach that directly uses a $\mathsf{Prob}^{>}\mathcal{EL}^\bot$ reasoner to close a pseudo-intent against the tautological $\mathsf{Prob}^{>}\mathcal{EL}^\bot$ concept inclusions does not solve this problem due to the exponential size of the attributes of the induced context and subsumption being **EXP**-complete for $\mathsf{Prob}^{>}\mathcal{EL}^\bot$.

Hence, if we define $\mathsf{Can}^*(\mathcal{I}, n) \coloneqq \mathsf{Can}(\mathcal{I}_n) \cup \bigcup\{\, \mathcal{U}_{\mathcal{I},\ell} \mid \ell \in \{1, \ldots, n\} \,\}$, then $\mathsf{Can}^*(\mathcal{I}, n)$ is still a $\mathsf{Prob}_n^{>}\mathcal{EL}^\bot$ concept inclusion base for $\mathcal{I}$ but, as a drawback, might contain tautological axioms. Its advantage is that it can always be computed in exponential time.

Proposition 15. *For any finitely representable probabilistic interpretation $\mathcal{I}$ and any $n \in \mathbb{N}$, the canonical base $\mathsf{Can}^*(\mathcal{I}, n)$ can be computed in deterministic time that is exponential in $|\Delta^\mathcal{I}| \cdot |\Omega^\mathcal{I}|$ and polynomial in n. Furthermore, there are finitely representable probabilistic interpretations $\mathcal{I}$ for which a concept inclusion base cannot be encoded in polynomial space with respect to $|\Delta^\mathcal{I}| \cdot |\Omega^\mathcal{I}| \cdot n$.*

Proof. The statements are obtained as corollaries of [22, Proposition 2] for the following reasons. The sum of two rational numbers can be computed in polynomial time. This result is necessary for determining the complexity of evaluating a $\mathsf{Prob}^{>}\mathcal{EL}^\bot$ concept description C in some world of a probabilistic interpretation $\mathcal{I}$, which is polynomial in $|C| + |\Delta^\mathcal{I}| \cdot |\Omega^\mathcal{I}|$. For each $n \in \mathbb{N}$, it holds true that the cardinality of $P_{\mathcal{I},n}$ is bounded by $|\Delta^\mathcal{I}| \cdot 2^{|\Delta^\mathcal{I}| \cdot |\Omega^\mathcal{I}|}$, i.e., $|P_{\mathcal{I},n}|$ is exponential in $|\Delta^\mathcal{I}| \cdot |\Omega^\mathcal{I}|$. For each $n \in \mathbb{N}$, we have that the cardinality of $\Gamma_n \setminus \Sigma$ is bounded by $n \cdot |\Delta^\mathcal{I}| \cdot 2^{2 \cdot |\Delta^\mathcal{I}| \cdot |\Omega^\mathcal{I}|}$, i.e., $|\Gamma_n \setminus \Sigma|$ is exponential in $|\Delta^\mathcal{I}| \cdot |\Omega^\mathcal{I}|$. Furthermore, each element in $\Gamma_n \setminus \Sigma$ has an encoding of exponential size, and we conclude that $\Gamma_n \setminus \Sigma$ also has an encoding of exponential size. $\square$

6 Conclusion

We have devised an effective procedure for computing finite axiomatizations of observations that are represented as probabilistic interpretations. More specifically, we have shown how concept inclusion bases—TBoxes that are sound and complete for the input data set—can be constructed in the probabilistic description logic $\mathsf{Prob}^{>}\mathcal{EL}^\bot$. In a complexity analysis we found that we can always compute a canonical base in exponential time.

Future research is possible in various directions. One could extend the results to a more expressive probabilistic DL, e.g., to $\mathsf{Prob}^{>}\mathcal{M}$, or one could include upper probability bounds. Furthermore, for increasing the practicability of the approach, it could be investigated how the construction of a concept inclusion base can be made *incremental* or *interactive*. It might be the case that there already exists a TBox and there are new observations in form of a probabilistic interpretation; the goal is then to construct a TBox being a base for the CIs that are entailed by the existing knowledge as well as hold true in the new observations. While this would represent a *push* approach of learning, future research could tackle the *pull* approach as well, i.e., equip the procedure with expert interaction such that an exploration of partial observations is made possible.

Additionally, it is worth investigating whether the proposed approach could be optimized; for instance, one could check if equivalent results can be obtained with a subset of Γ_n or with another extended signature. Currently, it is also unknown whether, for each finitely representable probabilistic interpretation $\mathcal{I}$, there is some finite bound n on the probability depth such that each $\mathsf{Prob}_n^{>}\mathcal{EL}^{\bot}$ concept inclusion base for $\mathcal{I}$ is also sound and complete for *all* $\mathsf{Prob}^{>}\mathcal{EL}^{\bot}$ concept inclusions that are valid in $\mathcal{I}$—much like this is the case for the role depth in $\mathcal{EL}^{\bot}$.

Acknowledgments The author gratefully thanks Franz Baader for drawing attention to the issue in [15], and furthermore thanks the reviewers for their constructive remarks.

References

1. Baader, F., Distel, F.: A Finite Basis for the Set of $\mathcal{EL}$-Implications Holding in a Finite Model. In: Medina, R., Obiedkov, S.A. (eds.) Formal Concept Analysis, 6th International Conference, ICFCA 2008, Montreal, Canada, February 25-28, 2008, Proceedings. Lecture Notes in Computer Science, vol. 4933, pp. 46–61. Springer (2008)
2. Baader, F., Horrocks, I., Lutz, C., Sattler, U.: An Introduction to Description Logic. Cambridge University Press (2017)
3. Bacchus, F.: Representing and Reasoning with Probabilistic Knowledge. Phd thesis, The University of Alberta, Edmonton, Alberta (1988)
4. Borchmann, D.: Learning Terminological Knowledge with High Confidence from Erroneous Data. Doctoral thesis, Technische Universität Dresden, Dresden, Germany (2014)
5. Borchmann, D., Distel, F., Kriegel, F.: Axiomatisation of General Concept Inclusions from Finite Interpretations. Journal of Applied Non-Classical Logics **26**(1), 1–46 (2016)
6. Distel, F.: Learning Description Logic Knowledge Bases from Data using Methods from Formal Concept Analysis. Doctoral thesis, Technische Universität Dresden, Dresden, Germany (2011)
7. Ganter, B.: Two Basic Algorithms in Concept Analysis. FB4-Preprint 831, Technische Hochschule Darmstadt, Darmstadt, Germany (1984)
8. Ganter, B., Obiedkov, S.A.: Conceptual Exploration. Springer (2016)
9. Ganter, B., Wille, R.: Formal Concept Analysis: Mathematical Foundations. Springer (1999)
10. Guigues, J.L., Duquenne, V.: Famille minimale d'implications informatives résultant d'un tableau de données binaires. Mathématiques et Sciences Humaines **95**, 5–18 (1986)
11. Gutiérrez-Basulto, V., Jung, J.C., Lutz, C., Schröder, L.: Probabilistic Description Logics for Subjective Uncertainty. Journal of Artificial Intelligence Research **58**, 1–66 (2017)
12. Halpern, J.Y.: An Analysis of First-Order Logics of Probability. Artificial Intelligence **46**(3), 311–350 (1990)
13. Heinsohn, J.: Probabilistic Description Logics. In: López de Mántaras, R., Poole, D. (eds.) UAI '94: Proceedings of the Tenth Annual Conference on Uncertainty in Artificial Intelligence, Seattle, Washington, USA, July 29-31, 1994. pp. 311–318. Morgan Kaufmann (1994)

14. Jaeger, M.: Probabilistic Reasoning in Terminological Logics. In: Doyle, J., Sandewall, E., Torasso, P. (eds.) Proceedings of the 4th International Conference on Principles of Knowledge Representation and Reasoning (KR'94). Bonn, Germany, May 24-27, 1994. pp. 305–316. Morgan Kaufmann (1994)
15. Kriegel, F.: Axiomatization of General Concept Inclusions in Probabilistic Description Logics. In: Hölldobler, S., Krötzsch, M., Peñaloza, R., Rudolph, S. (eds.) KI 2015: Advances in Artificial Intelligence - 38th Annual German Conference on AI, Dresden, Germany, September 21-25, 2015, Proceedings. Lecture Notes in Computer Science, vol. 9324, pp. 124–136. Springer (2015)
16. Kriegel, F.: Incremental Learning of TBoxes from Interpretation Sequences with Methods of Formal Concept Analysis. In: Calvanese, D., Konev, B. (eds.) Proceedings of the 28th International Workshop on Description Logics, Athens, Greece, June 7-10, 2015. CEUR Workshop Proceedings, vol. 1350. CEUR-WS.org (2015)
17. Kriegel, F.: Acquisition of Terminological Knowledge from Social Networks in Description Logic. In: Missaoui, R., Kuznetsov, S.O., Obiedkov, S.A. (eds.) Formal Concept Analysis of Social Networks, pp. 97–142. Lecture Notes in Social Networks, Springer (2017)
18. Kriegel, F.: Implications over Probabilistic Attributes. In: Bertet, K., Borchmann, D., Cellier, P., Ferré, S. (eds.) Formal Concept Analysis - 14th International Conference, ICFCA 2017, Rennes, France, June 13-16, 2017, Proceedings. Lecture Notes in Computer Science, vol. 10308, pp. 168–183. Springer (2017)
19. Kriegel, F.: Probabilistic Implication Bases in FCA and Probabilistic Bases of GCIs in $\mathcal{EL}^{\perp}$. International Journal of General Systems **46**(5), 511–546 (2017)
20. Kriegel, F.: Acquisition of Terminological Knowledge in Probabilistic Description Logic. In: Trollmann, F., Turhan, A. (eds.) KI 2018: Advances in Artificial Intelligence - 41st German Conference on AI, Berlin, Germany, September 24-28, 2018, Proceedings. Lecture Notes in Computer Science, vol. 11117, pp. 46–53. Springer (2018)
21. Kriegel, F.: Learning Description Logic Axioms from Discrete Probability Distributions over Description Graphs (Extended Version). LTCS-Report 18-12, Chair of Automata Theory, Institute of Theoretical Computer Science, Technische Universität Dresden, Dresden, Germany (2018), `https://tu-dresden.de/inf/lat/reports#Kr-LTCS-18-12`
22. Kriegel, F.: Most Specific Consequences in the Description Logic $\mathcal{EL}$. LTCS-Report 18-11, Chair of Automata Theory, Institute of Theoretical Computer Science, Technische Universität Dresden, Dresden, Germany (2018), `https://tu-dresden.de/inf/lat/reports#Kr-LTCS-18-11`, accepted for publication in Discrete Applied Mathematics
23. Lukasiewicz, T.: Expressive Probabilistic Description Logics. Artificial Intelligence **172**(6-7), 852–883 (2008)
24. Rudolph, S.: Relational Exploration: Combining Description Logics and Formal Concept Analysis for Knowledge Specification. Doctoral thesis, Technische Universität Dresden, Dresden, Germany (2006)