Relationships between Probabilistic Description and First-Order Logics

Pavel Klinov and Bijan Parsia

School of Computer Science
University of Manchester, United Kingdom
{pklinov|bparsia}@cs.man.ac.uk

Abstract This paper analyzes the probabilistic description logic P-$\mathcal{SHIQ}$ as a fragment of first-order probabilistic logic (FOPL). P-$\mathcal{SHIQ}$ was suggested as a language that is capable of representing and reasoning about different kinds of uncertainty in ontologies, namely generic probabilistic relationships between concepts and probabilistic facts about individuals. However, some semantic properties of P-$\mathcal{SHIQ}$ have been unclear which raised concerns regarding whether it could be used for representing probabilistic ontologies. In this paper we provide an insight into its semantics by translating P-$\mathcal{SHIQ}$ into FOPL with a specific semantics based on possible worlds. From that reduction, we show that some of the restrictions of P-$\mathcal{SHIQ}$ are fundamental and sketch alternative semantic foundations for a probabilistic description logic.

1 Introduction and Motivation

One common complaint about description logic (DL) based ontology languages, such as the Web Ontology Language (OWL), is they fail to support non-classical uncertainty, in particular, probability. One answer to this complaint is the P-$\mathcal{SH}$ family of logics which allow for the incorporation of probabilistic formulae as an extension of the familiar and widely used $\mathcal{SH}$ DLs [1] [2]. Unlike Bayesian extensions to DLs and OWL, the P-$\mathcal{SH}$ family consists of proper extensions to the syntax and semantics of the underlying logic and inference services. These logics are also decidable, generally of the same worst case complexity as the base logic, and can be implemented on top of existing DL reasoners.

However, there are several issues with the P-$\mathcal{SH}$ family both from an expressivity and from a theoretical point of view. First, it has not been fully clear how it actually combines statistical and subjective probabilities. Second, probabilistic ABoxes have a number of strong restrictions (including no support of roles assertions between probabilistic individuals and only one probabilistic individual per ABox).

Often, insight into a DL (and associated extensions and reasoning techniques) has followed by considering its standard first-order translation, that is, in considering it as a fragment of first order-logics. In this paper, we attempt to apply this methodology to the P-$\mathcal{SH}$ family by considering them as fragments of a first-order logic extended with various forms of probability (FOPL). We show that we can understand P-$\mathcal{SH}$ logics as fragments of FOPL and explain its limitations on the basis of the known properties of FOPL with semantics based on possible worlds. Finally, we sketch another fragment of FOPL which has different semantics and allows lifting of the P-$\mathcal{SH}$ restrictions.

2 Preliminaries

P-$\mathcal{SHIQ}$ We consider a particular representative of the P-$\mathcal{SH}$ family, named P-$\mathcal{SHIQ}$, whose syntactic constructs include those of $\mathcal{SHIQ}$ together with *conditional constraints*. Constraints are expressions of the form $(D|C)[l,u]$ where D, C are $\mathcal{SHIQ}$ concept expressions (called *conclusion* and *evidence* respectively) and $[l,u] \subseteq [0,1]$.

A probabilistic TBox (PTBox) is a 2-tuple $PT = (\mathcal{O}, \mathcal{P})$ where $\mathcal{O}$ is a classical DL ontology and P is a finite set of conditional constraints. Informally, a PTBox axiom $(D|C)[l,u]$ means that "if a randomly chosen individual belongs to C, its probability of belonging to D is in $[l,u]$". A probabilistic ABox (PABox) is a finite set of conditional constraints pertaining to a single probabilistic individual o. Set of all probabilistic individuals is denoted as $I_{\mathcal{P}}$. A probabilistic ontology $PO = (\mathcal{O}, \mathcal{P}, (\mathcal{P}_o)_{o \in I_P})$ is a combination of one PTBox and a set of PABoxes, one for each probabilistic individual.

The semantics of P-$\mathcal{SHIQ}$ is standardly explained in terms of the notion of a *possible world* which is defined with respect to a set of basic concepts Φ [2]. A possible world I is a set of DL concepts from Φ such that $\{a : C | C \in I\} \cup \{a : \neg C | C \notin I\}$ is satisfiable for a fresh individual a (in other words, possible worlds correspond to *realizable* concept types). The set of all possible worlds with respect to Φ is denoted as $\mathcal{I}_{\Phi}$. A world I satisfies a concept C denoted as $I \models C$ if $C \in I$. Satisfiability of basic concepts is inductively extended to concept expressions as usual.

A world I is said to be a *model* of a DL axiom η denoted as $I \models \eta$ if $\eta \cup \{a : C | C \in I\} \cup \{a : \neg C | C \notin I\}$ is satisfiable for a fresh individual a. A world I is a model of a classical DL knowledge base $\mathcal{O}$ denoted as $I \models \mathcal{O}$ if it is a model of all axioms of $\mathcal{O}$. Existence of such world is equivalent to the standard satisfiability in DL [2].

A probabilistic interpretation Pr is a discrete probability distribution over $\mathcal{I}_{\Phi}$. Pr is said to *satisfy* a DL knowledge base $\mathcal{O}$ denoted as $Pr \models KB$ iff $\forall I \in \mathcal{I}_{\Phi}, Pr(I) > 0 \Rightarrow I \models KB$. The probability of a concept C, denoted as $Pr(C)$, is defined as $\sum_{I \models C} Pr(I)$. $Pr(D|C)$ is used as an abbreviation for $Pr(C \sqcap D)/Pr(C)$ given $Pr(C) > 0$. A probabilistic interpretation Pr satisfies a conditional constraint $(D|C)[l,u]$, denoted as $Pr \models (D|C)[l,u]$, iff $Pr(C) = 0$ or $Pr(D|C) \in [l,u]$. Pr satisfies a set of conditional constraints F iff it satisfies each of the constraints. A PTBox $PT = (\mathcal{O}, \mathcal{P})$ is called *satisfiable* iff there exists an interpretation that satisfies $\mathcal{O} \cup \mathcal{P}$. Logical entailment is defined in a standard way [2][1].

First-Order Probabilistic Logic FOPL$_2$ is a probabilistic generalization of first-order logic aimed at capturing belief statements (the subscript 2 stands for the Type 2 semantics [3]), like "the probability that Tweety (a particular bird) flies is over 90%". It is very expressive allowing to attach probabilities to arbitrary first-order formulas. Its representational and computational properties have been thoroughly investigated, and the results of these investigations are applicable to its fragments.

The *syntax* of FOPL$_2$ is defined as follows [3]: assume a first-order alphabet Φ of function and predicate names, and a countable set of object variables X^o. *Object*

[1] P-$\mathcal{SHIQ}$, as it is presented in [2], is a non-monotonic formalism. However, we consider only its monotonic basis in this paper. Our position is that it must be clarified first, before proceeding to non-monotonic machinery, such as lexicographic entailment, built on its top.

terms are formed by closing X^o off under function application as usual. In addition, the language contains *field terms*, which range over reals (with 0 and 1 being distinguished constants) and probability terms of the form $w(\phi)$, where ϕ is a first-order formula. Field terms are closed off under applications of functions $\times, +$ on reals (the denotation $w(\phi|\psi) \leq t$ is the abbreviation of $w(\phi \wedge \psi) \leq t \times w(\psi)$). Then FOPL$_2$ formulas are defined as follows:

- If P is an n-ary predicate name in Φ and $t_1, \ldots, t_n$ are object terms, then $P(t_1, \ldots, t_n)$ is an atomic formula.
- If t_1, t_2 are field terms, then $t_1 \leq t_2$, $t_1 \geq t_2$, $t_1 < t_2$, $t_1 > t_2$, $t_1 = t_2$ are atomic formulas. Standard relationships between (in)equality relations are assumed.
- If ϕ, ψ are formulas and $x \in X^o$, then $\phi \wedge \psi$, $\phi \vee \psi$, $\forall(x)\phi$, $\exists(x)\phi$, $\neg\phi$ are formulas. Standard relationships between logical connectives and quantifiers are assumed.

A *probabilistic interpretation* (Type 2 probability structure in [3]) M is a tuple (D, S, π, μ), where D is a domain, S is a set of states, π is a function $S \times \Phi \rightarrow \Phi_D$ (where Φ_D is a set of predicates and functions over D) which preserves arity, and μ is a probability distribution over S. M together with a state s and a valuation v associates each object term o with an element of D ($[o]^{(M,s,v)} \in D$) and each field term f with a real number. (M, s, v) associates formulas with truth values (we write $(M, s, v) \models \phi$ if ϕ is true in (M, s, v)) as follows:

- $(M, s, v) \models P(x)$ if $v(x) \in \pi(s, P)$.
- $(M, s, v) \models t_1 < t_2$ if $[t_1]^{M,s,v} < [t_2]^{(M,s,v)}$.
- $(M, s, v) \models \forall(x)\phi$ if $(M, s, v[x/d]) \models \phi$ for all $d \in D$.

Other formulas, e.g. $\phi \wedge \psi$, $\neg\psi$, $t_1 = t_2$, etc. are defined as usual. It remains to define the mapping for the probability terms of the form $w(\phi)$: $[w(\phi)]^{(M,s,v)} = \mu\{s' \in S|(M, s', v) \models \phi\}$. As usual, a FOPL$_2$ formula is called *satisfiable* if there exists a tuple (M, s, v) in which the formula is true.

Note that, although FOPL$_2$ does not impose any restrictions on the set S (i.e. it can be any set over which a probability distribution can be defined). However, it is natural to associate states with possible interpretations of symbols in Φ over D (see [4]). Then the model structure can be simplified to (D, S, μ) since the interpretations are implicitly encoded in the states.

3 Mapping between P-$\mathcal{SHIQ}$ and FOPL$_2$

This section presents a mapping between P-$\mathcal{SHIQ}$ and FOPL$_2$. For brevity we will limit our attention to $\mathcal{ALC}$ concepts (calling the resulting logic P-$\mathcal{ALC}$) as the translation can be easily extended to more expressive DLs. We will show that it preserves entailments so that P-$\mathcal{SHIQ}$ can be viewed as a fragment of FOPL$_2$.

Basic Translation We define the injective function κ to be the mapping of syntactic constructs of P-$\mathcal{ALC}$ to FOPL$_2$[2]. It is a superset of the standard translation of $\mathcal{ALC}$ into

Table 1. Translation of P-$\mathcal{ALC}$ formulae into FOPL_2

P-$\mathcal{ALC}$	**FOPL_2**
$\kappa(A, var)$	$A(var)$
$\kappa(\neg C, var)$	$\neg(\kappa(C, var))$
$\kappa(R, var, var')$	$R(var, var')$
$\kappa(C \sqcap D, var)$	$\kappa(C, var) \wedge \kappa(D, var)$
$\kappa(C \sqcup D, var)$	$\kappa(C, var) \vee \kappa(D, var)$
$\kappa(\forall R.C, var)$	$\forall(var')(R(var, var') \rightarrow \kappa(C, var'))$
$\kappa(\exists R.C, var)$	$\exists(var')(R(var, var') \wedge \kappa(C, var'))$
$\kappa(a : C)$	$\kappa(C, x)[a/x]$
$\kappa((a, b) : R)$	$R[a/x, b/y]$
$\kappa(C \sqsubseteq D, x)$	$\forall(x)(\kappa(C, x) \rightarrow \kappa(D, x))$
$\kappa((B\|A)[l, u], x)$	$l \leq w(B(r)\|A(r)) \leq u$

FOL [5] (in the Table 3 A, B stand for concept names, R for a role name, C, D for concepts, r for a fresh constant, $var \in \{x, y\}; var' = x$ if $var = y$ and y if $var = x$).

This function transforms a P-$\mathcal{ALC}$ PTBox into a FOPL_2 theory. The most important thing is that it translates *generic* PTBox constraints into *ground* probabilistic formulas for a fresh constant r. The implications of this will be discussed in Section 4.

Faithfulness We next show that this translation is faithful by establishing correspondence between models in P-$\mathcal{ALC}$ and FOPL_2. Observe, that in contrast to [6], here we consider the natural choice of states in Type 2 model structure in which they correspond to first-order models of the knowledge base.

Theorem 1. *Let $PT = (\mathcal{O}, \mathcal{P})$ be a PTBox in P-$\mathcal{ALC}$ and $F = \{\kappa(\phi)|\phi \in \mathcal{O} \cup \mathcal{P}\}$ be the corresponding $FOPL_2$ theory. Then for every P-$\mathcal{ALC}$ model Pr of PT there exists a corresponding Type 2 structure $M = (D, S, \mu)$ such that 1) $M \models \kappa(\phi)$ for all $\phi \in \mathcal{O}$ and 2) $M \models l \leq w(B(r)|A(r)) \leq u$ for all conditional constraints $(B|A)[l, u]$ in $\mathcal{P}$, and vice versa, where κ is defined according to Table 1.*

Proof. We prove only ($\Rightarrow$). Let $Pr : \mathcal{I}_\Phi \rightarrow [0, 1]$ be a model of PT. Pr satisfies classical ontology $\mathcal{O}$ so there exists a classical model $\mathcal{I} = (\Delta^\mathcal{I}, \cdot^\mathcal{I})$ of $\mathcal{O}$. We first extend $\Delta^\mathcal{I}$ to ensure that all possible worlds are realizable over it (one possibility is to take the disjoint union of all realizations)[3]. Then we construct a Type 2 structure (D, S, μ) as follows: let $D = \Delta^\mathcal{I}$ and S be the set of all interpretations of predicate names (translations of concept and roles names) and r over $\Delta^\mathcal{I}$ that satisfy classical formulas in F. S must be non-empty: let $s_\mathcal{I}$ be such that $s_\mathcal{I}(P) = \kappa^{-1}(P)^\mathcal{I}$ for all predicate names and $s_\mathcal{I}(r)$ is an arbitrary domain element. Since r is a fresh constant and κ encompasses a standard and faithful translation from $\mathcal{ALC}$ to FOL, $s_\mathcal{I}$ is a model of all classical formulas in F and therefore *1)* holds.

[2] For a possible world $I = \{C_i\}$ we use the notation $\kappa(I)$ to denote the set $\{\kappa(C_i)\}$
[3] This is only possible if the DL does not allow for nominals.

The rest is to define a probability distribution μ that satisfies probabilistic formulas in F. Recall that Pr is probability distribution over the set of (possible worlds). We define a function σ which maps each world $I = \{C_i\}$ to a set of states $\sigma(I) \subseteq S$ as follows: $\sigma(I) = \{s | s \models \kappa(I)(r)\}$. Then let $\mu(\sigma(I)) = Pr(I)$ for all possible worlds. It is not hard to see that μ is a probability distribution as it mimics the probability distribution Pr. Finally, (D, S, μ) satisfies all formulas of the kind $l \leq w(B(r)|A(r)) \leq u$ in F because $\mu(B(r) \wedge A(r)) = Pr(B \sqcap A), \mu(A(r)) = Pr(A))$ (by construction, e.g., $\mu(A(r)) = \mu(\{s|s \models A(r)\}) = \sum_{Ct_i \models A} \mu\{\sigma(Ct_i)\} = \sum_{Ct_i \models A} Pr(Ct_i) = Pr(A))$ and $Pr \models (B|A)[l, u]$ and therefore *2)* holds. □

Theorem 1 implies that the translation preserves satisfiability and entailments.

Translation of PABoxes One particularly odd restriction of P-$\mathcal{SHIQ}$ is that PABoxes cannot be combined into a single set of formulas. This is so because PABox constraints are modeled as generic constraints and the information about the individual is present only on a meta-level (as a label of the PABox). Therefore, to extend our translation to PABoxes we either have to translate them into a corresponding disjoint set of labeled FOPL_2 theories or make special arrangements to faithfully translate them into a combined FOPL_2 theory. We opt for the latter because it will let us get rid of any meta-logical aspects and help analyze a P-$\mathcal{SHIQ}$ ontology as a standard FOPL_2 theory.

Since PABoxes in P-$\mathcal{SHIQ}$ are isolated from each other, the translation should preserve that isolation. The most obvious way to prevent any interaction between sets of formulas in a single logical theory is to make their signatures disjoint. However, the translation should not only respect disjointness of PABoxes but also preserve their interaction with PTBox and the classical part of the ontology (see Example 1).

Example 1. Consider the following PTBox: $\mathcal{P} = \{(FlyingObject|Bird)[0.9, 1],$ $(FlyingObject|\neg Bird)[0, 0.5]\}$ and two PABoxes: $\mathcal{P}_{Tweety} = \{(Bird|\top)[1, 1]\}$, $\mathcal{P}_{Sam} = \{(\neg Bird|\top)[1, 1]\}$. Obviously, if these sets of axioms are translated and combined into a single FOPL_2 theory then it will contain a conflicting pair of formulas $\{w(Bird(r)) \geq 0.9, w(Bird(r)) \leq 0.5\} \subseteq F$. This inconsistency can be avoided by introducing fresh first-order predicates for every PABox: $\{w(Bird_{Tweety}(r)) \geq 0.9,$ $w(Bird_{Sam}(r)) \leq 0.5\}$. However, this would break any connection between PTBox and PABox axioms, for example, prevent the following entailments:
$\{w(FlyingObject_{Tweety}(r)) \geq 0.9, w(FlyingObject_{Sam}(r)) \leq 0.5\}$.

One can faithfully extend the translation to PABoxes by introducing fresh concept names to *relativize* each TBox and PTBox axiom for every probabilistic individual to avoid inconsistencies. The transformation will consist of the following steps[4]:

- Firstly, we transform a P-$\mathcal{ALC}$ ontology $PO = (\mathcal{O}, \mathcal{P}, (\mathcal{P}_o))$ into a set of PTBoxes $\{(\mathcal{O}, \mathcal{P} \cup \mathcal{P}_o)\} \cup \{(\mathcal{O}, \mathcal{P})\}$. Informally, we create a copy PTBox for every probabilistic individual (PT_o) and make them isolated from each other. Now, instead of one PTBox and a set of PABoxes we have just a set of PTBoxes. This step preserves probabilistic entailments in the following sense: $PO \models (B|A)[l, u]$ iff $(\mathcal{O}, \mathcal{P}) \models (B|A)[l, u]$ and $PO \models (B|A)[l, u]$ for o iff $PT_o \models (B|A)[l, u]$.

[4] Full example is available at http://www.cs.man.ac.uk/~klinovp/research/pshiq/example.pdf.

- Secondly, we transform every PTBox PT_o into PT'_o by renaming every concept name C into C_o in all TBox axioms and conditional constraints. It is easy to see that $PT_o \models C \sqsubseteq D$ iff $PT'_o \models C_o \sqsubseteq D_o$ and $PT_o \models (B|A)[l,u]$ iff $PT'_o \models (B_o|A_o)[l,u]$. Intuitively, we have created a fresh copy of each PTBox to guard against possible conflicts between PABox constraints for different probabilistic individuals. Signatures of PT'_o are pairwise disjoint and denoted as Σ_o.
- Next, we union all PT'_o with disjoint signatures (including the original $PT = (\mathcal{O}, \mathcal{P})$) into a single unified PTBox $PT_U = \bigcup_{o \in I_p} PT_o \cup PT$ with signature $\Sigma_U = \bigcup_{o \in I_p} \Sigma_o \cup \Sigma$.
- Finally we can apply the previously presented faithful translation to PT_U and obtain a single FOPL$_2$ theory which corresponds to the original P-$\mathcal{ALC}$ ontology.

A necessary condition for faithfulness of this transformation is that the original isolation of PABoxes is preserved by creating fresh copies of PTBoxes. In particular, this means that the unified PTBox cannot entail any subsumption relation between concept expressions C_{o_1} and C_{o_2} defined over disjoint signatures except of the case when one of them is either $\top$ or $\bot$. If this is false, for example, if $PT_U \models C_{o_1} \sqsubseteq C_{o_2}$ then the following PABox constraints represented as $(C_{o_1}|\top)[1,1]$ and $(C_{o_2}|\top)[0,0]$ will be mutually inconsistent in PT_U (but they were consistent in the original P-$\mathcal{ALC}$ because they belonged to different PABoxes isolated from each other). This condition is formalized in the following lemma (whose proof is omitted for brevity):

Lemma 1. *Let $\mathcal{O}_1$ and $\mathcal{O}_2$ be copies of a satisfiable $\mathcal{ALC}$ ontology $\mathcal{O}$ with disjoint signatures Σ_1 and Σ_2, and $\mathcal{O}_U$ be the union of $\mathcal{O}_1$ and $\mathcal{O}_2$. Then for any concept expressions C_1, C_2 over Σ_1 and Σ_2 respectively such that $\mathcal{O}_1 \not\models C_1 \sqsubseteq \bot$ and $\mathcal{O}_1 \not\models \top \sqsubseteq C_2$, $\mathcal{O}_U \not\models C_1 \sqsubseteq C_2$.*

Now we can obtain the main result:

Theorem 2. *Let $PO = (\mathcal{O}, \mathcal{P}, (\mathcal{P}_o))$ be a P-$\mathcal{ALC}$ ontology and F be a FOPL$_2$ theory obtained by combining PABoxes and translating the resulting PTBox into FOPL. Then for every P-$\mathcal{ALC}$ model Pr_o of $PT_o = (\mathcal{O}, \mathcal{P} \cup \mathcal{P}_o)$ for every probabilistic individual o there exists a corresponding Type 2 structure $M = (D, S, \mu)$ such that:*

1. *$M \models \kappa(\phi)$ for all $\phi \in \mathcal{O}$,*
2. *$M \models l \leq w(B(r)|A(r)) \leq u$ for all conditional constraints $(B|A)[l,u]$ in $\mathcal{P}$,*
3. *$M \models l \leq w(B_o(r)|A_o(r)) \leq u$ for all conditional constraints $(B|A)[l,u]$ in $\mathcal{P}_o$,*

and vice versa, where κ is defined according to Table 1.

Proof. Due to Theorem 1 it suffices to show that the steps 1-3 of the transformation preserve probabilistic models. This can be done by establishing a correspondence between possible worlds of each PT_o and PT_U. Since there are no subsumptions between concept expressions over signatures of different PTBoxes (see Lemma 1), each possible world I_o in PT_o corresponds to a finite set of possible worlds of PT_U defined as: $\sigma(I_o) = \{I_U | C_{i_o} \in I_U \text{ iff } C_i \in I_o\}$ (each C_{i_o} is a new concept name for C_i introduced on step 2). Then, a probability distribution over all possible worlds in PT_U can be defined as $Pr_U(I_U) = Pr_o(I_o)/|\sigma(I_o)|$. It follows that for any concept C over Σ_o,

$Pr_o(C)$ is equal to $Pr_U(C_o)$ where C_o is the correspondingly renamed concept. Therefore, $Pr_U \models (B_o|A_o)[l,u]$ if $Pr_o \models (B|A)[l,u]$. The reverse direction can be proved along the same lines (i.e., $Pr_o(I_o)$ can be defined as $\sum_{I_U \in \sigma(I_o)} Pr_U(I_U)$).

4 Discussion

The main conclusion following from the presented translation is that P-$\mathcal{SHIQ}$ all PTBox statements express *degrees of belief* (i.e. subjective probabilities) about a single, yet unnamed, individual. This is not an easily expected outcome because the variable-free syntax of P-$\mathcal{SHIQ}$ may give a misleading impression that PTBox constraints correspond to universally quantified formulas of some sort. The fact that probabilistic individuals are not translated to corresponding constants in FOPL$_2$ (in contrast to classical individuals) is also not a trivial outcome. Both these features of P-$\mathcal{SHIQ}$ have important implications, but before moving to them, let us consider another, perhaps more naturally looking translation and explain why it is not faithful.

It may well appear that conditional constraints in P-$\mathcal{SHIQ}$ should be interpreted as implicitly universally quantified formulas analogously to probabilistic logic programming. That way, $(B|A)[l,u]$ corresponds to $\forall x(l \leq w(B(x)|A(x)) \leq u)$. However, the standard behavior of the universal quantifier is incompatible with the P-$\mathcal{SHIQ}$ semantics in which classical and probabilistic individuals are separated. For example, the PTBox $(\{a : \neg A\}, \{(A|\top)[1,1]\})$ is satisfiable although the corresponding FOPL theory $\{\neg A(a), \forall x(w(A(x)) = 1)\}$ is not.

There is a possibility to interpret conditional constraints in P-$\mathcal{SHIQ}$ as closed quantified formulas, but this requires a non-standard quantifier which makes the variable act as a random designator. This idea dates back to Cheeseman who originally proposed to use formulas of the form $\forall x.pr[B(x)|A(x)][l,u]$ to capture statistical knowledge [7]. In fact, the fresh constant r used in our translation plays the role of such non-standard quantifier. However, as pointed out by several authors (see especially [3] [8] [9]), such formulas *cannot* serve as representations of statistical assertions because their interpretations are not based on proportions of domain elements[5].

Unfortunately, using Type 2 semantics to interpret different kinds of probabilities complicates not only the representation of statistics but also the combination of statistical assertions with probabilistic statements about specific individuals (degrees of belief). In particular, this requires modeling of PABox constraints in P-$\mathcal{SHIQ}$ as generic PTBox statements with information about individuals presenting only on a meta-level. This is the reason why PABox statements do not correspond to ground probabilistic formulas in FOPL$_2$. If they did, then there would be no connection between a "statistical" statement $(FlyingObject|Bird)[0.9,1]$ (represented in FOPL$_2$ as $(0.9 \leq flyingobject(r)|bird(r)) \leq 1)$ and a belief statement $(tweety : Bird)[1,1]$ (represented as $1 \leq w(bird(tweety) \leq 1)$ since beliefs about r cannot affect beliefs about $tweety$. Therefore $(tweety : Bird)[1,1]$ is effectively modeled as $(Bird|\top)[1,1]$

[5] We must mention that P-$\mathcal{SHIQ}$ could, in principle, be translated to FOPL with domain-based semantics by employing a known translation between domain-based probability and possible-world-based probability (see [10] for details). However, this will solve no issues with P-$\mathcal{SHIQ}$ as it will still behave as FOPL$_2$ with single probabilistic individual.

(or as $1 \leq w(bird(r)) \leq 1$ in FOPL$_2$) with the individual name $tweety$ lifted at the meta-level to serve as a label for the corresponding PABox.

However, this introduces other problems which are responsible for the limitations of P-$\mathcal{SHIQ}$. Since PABox constraints expressing probabilistic knowledge about different probabilistic individuals *must* be isolated from each other, there appears to be no straightforward way of combining them. In particular, this prohibits representation of classical or probabilistic role assertions between different probabilistic individuals or, in other words, the logic does not support probabilistic relational structures[6]. Thus, it can be concluded that, in essence, P-$\mathcal{SHIQ}$ is closer to a propositional probabilistic logic rather than to a full-fledged probabilistic first-order formalism.

The problems mentioned above cannot be solved simply by adopting an appropriate semantics for representing statistics, such as Type 1 semantics in which probability distributions are defined over the interpretation domain. Such an attempt has been made by Giugno and Lukasiewicz in the early paper on P-$\mathcal{SHOQ}$ [1]. In that logic probabilistic concept membership assertions were represented using nominals, for example, $(C|\{a\})[0.5, 1]$. Unfortunately, as proved by Halpern, closed first-order formulas can only have probability 0 or 1 in any Type 1 probabilistic model (see Lemma 2.3 in [3]) so the representation is unsatisfactory. It is not hard to see that the probability of $(C|\{a\})$, equivalent to $\frac{Pr(C \sqcap \{a\})}{Pr(\{a\})}$, is 0 if $a^{\mathcal{I}} \notin C^{\mathcal{I}}$ or 1 if $a^{\mathcal{I}} \in C^{\mathcal{I}}$ if Pr is defined over $\Delta^{\mathcal{I}}$.

All the features and limitations explained above are by no means unique to P-$\mathcal{SHIQ}$. They have been discovered and studied for first-order logics by a number of authors who claimed that neither domain-based nor possible world-based semantics *by itself* is suitable for representation and reasoning about different kinds of probabilities. However, their proper combination (called Type 3 semantics [3]) has the required potential. The corresponding logic (FOPL$_3$) is free of any limitations described above, is completely axiomatizable for a range of interesting fragments (e.g., logics with bounded model property such as $\mathcal{ALC}$), and can be used for defining probabilistic DLs.

5 Probabilistic Description Logic with Combined Semantics

In this section we briefly outline the syntax and semantics of the extended probabilistic DL for representation and reasoning about different kinds of probabilities. The language corresponds to the DL fragment of FOPL$_3$ with the principle of direct inference [9]. We loosely call it P-$\mathcal{DL}$ (where $\mathcal{DL}$ stands for a subset of $\mathcal{SROIQ}$).

Syntax Analogously to P-$\mathcal{SHIQ}$ the syntax of P-$\mathcal{DL}$ is based on conditional constraints. However, we distinguish between statistical constraints and belief constraints by providing different syntactic constructs for each. *Statistical conditional constraints* are expressions of the form $(D|C)_{stat}[l, u]$ where D, C are concept expressions. *Belief constraints* are expressions of the form $(\phi)_{belief}[l, u]$ or $(\psi|\phi)_{belief}[l, u]$ where ψ, ϕ are ABox assertions. We define PTBox to be a set of statistical constraints, and PABox to

[6] Allowing nominals in the classical part of the language lets us express probabilistic roles $R(a, b)[l, u]$ as $(\exists R.\{b\}|\top)[l, u]$ for a [2]. However, this is still very restrictive because there cannot be a PABox for b (in other words, b cannot be a probabilistic individual).

be a set of belief constraints. An ontology in P-$\mathcal{DL}$ is a triple $(\mathcal{O}, \mathcal{P}_{stat}, \mathcal{P}_{belief})$ where $\mathcal{O}$ is a $\mathcal{DL}$ ontology, $\mathcal{P}_{stat}$ is a PTBox and $\mathcal{P}_{belief}$ is a PABox.

Semantics Both types of conditional constraints are interpreted using the Type 3 structure $M = (\Delta, S, Pr_{stat}, Pr_{belief})$ [3]. Here Δ is a non-empty domain, S is a set of states that correspond to interpretations of concept, role and individual names over Δ, Pr_{stat} is a probability distribution over Δ, and Pr_{belief} is a probability distribution over S. For a state $s \in S$ we use $s(C)$ (resp. $s(R), s(a)$) to express the interpretation of a concept C (resp. role R and individual a) in s. For an axiom η we write $s \models \eta$ if η is satisfied by the corresponding interpretation. Such combined structure is used to interpret both statistical and belief statements respectively in the following way:

- Statistical probability of a concept C in M in a state s (denoted as $C^{(M,s)}$) is equal to $Pr_{stat}\{d \in \Delta | d \in s(C)\}$. $(D|C)^{M,s}$ is an abbreviation of $\frac{(D \sqcap C)^{(M,s)}}{C^{(M,s)}}$.
- Subjective probability of an ABox assertion ϕ (denoted as ϕ^M is equal to $Pr_{belief}\{s \in S | s \models \phi\}$. $(\psi|\phi)^M$ is an abbreviation of $\frac{(\psi \sqcap \phi)^M}{\phi^M}$.
- M satisfies a statistical constraint $(D|C)_{stat}[l, u]$ if $\forall s \in S, (D|C)^{(M,s)} \in [l, u]$.
- M satisfies a belief constraint $(\psi|\phi)_{belief}[l, u]$ if $(\psi|\phi)^M \in [l, u]$.

Direct Inference FOPL$_3$ provides means for representing and reasoning about different kinds of probabilities but, as it stands, it does not support any relationship between them. However, in most scenarios, e.g., in actuarial reasoning, it is desirable to infer subjective beliefs from available classical and statistical knowledge. Such reasoning is often called *direct inference* and it can be supported in FOPL$_3$ and its fragments.

The main idea behind direct inference, that goes back to Reichenbach's reference class reasoning [11], is to consider every individual to be a *typical* representative of the *smallest* class of objects which it belongs to and for which *reliable* statistics is available. For example, the probability that Tweety flies should be equal to the probability that a randomly taken object, having the same set of properties as Tweety, flies. There are a few proposed ways to implement this idea, one of which we sketch below.

One can capture the notion of typicality directly by equating the degree of belief in a ground formula to the *expectation* of the statistical probability of its *randomized* version given the rest of classical and statistical formulas, as proposed in [9]. Randomization is replacement of all constants in ground formulas by fresh variables. Expectation of a field term f is a rigid (i.e. not depending on a state) term defined as $E(f)^M = \sum_{s \in S} Pr_{belief}(s) \times [f]^{(M,s)}$. The expectation operator and conditioning on statistical formulae are only used on the semantic, not syntactic, level of P-$\mathcal{DL}$.

Consider the example. Let $\{(Fly|Bird)[0.9, 1]\}$ be PTBox and $Bird(tweety)$ be an ABox axiom. Then the degree of belief in $Bird(tweety)$ is within the bounds of $E(bird(v)|0.9 \leq w(fly(v)|bird(v)) \geq 1)$, where v is a fresh constant introduced by randomization. The resulting interval is $[0.9, 1]$, as expected. Note that P-$\mathcal{DL}$ will probably require a non-monotonic mechanism similar to P-$\mathcal{SHIQ}$ to handle situations when an explicitly specified subjective probability is different from the computed via direct inference (e.g., when the individuals in question are *not* typical).

Direct inference via randomization serves the same purpose as P-$\mathcal{SHIQ}$'s way of combining PTBox and PABox constraints (in that sense P-$\mathcal{SHIQ}$ can be thought of as an implementable, non-monotonic *approximation* of FOPL_3). However, it is considerably less restrictive because it does not require representing PABox statements as universal PTBox constraints. Since all belief statements about particular individual are ground formulas with proper constants (like $tweety$), they can be combined in a single theory. Thus the representation supports arbitrary relational structures involving different probabilistic individuals and does not force unnatural separation of PABoxes. It is also possible to make the assumption that a pair of individuals are typical thus enabling the inference of probabilistic role assertions. Finally, it supports smooth integration of classical knowledge (i.e. ABox axioms) and beliefs about the same individual while P-$\mathcal{SHIQ}$ requires separation between classical and probabilistic individuals.

6 Conclusion

In this paper we have presented a new look at the probabilistic DL P-$\mathcal{SHIQ}$ as a fragment of probabilistic first-order logic. We gave a translation of P-$\mathcal{SHIQ}$ knowledge bases into FOPL_2 theories and proved its faithfulness. This brought an extra insight into P-$\mathcal{SHIQ}$, most importantly, into its limitations. It appears that the major restriction, namely the lack of support of relational structure for probabilistic individuals, is caused by attempt to use the possible world based semantics for different kinds of probabilities. This makes the probabilistic component of P-$\mathcal{SHIQ}$ essentially propositional (i.e. all probabilistic statements relate to a single constant r). We sketched how a more direct fragment of FOPL, which we called P-$\mathcal{DL}$, could overcome these limitations while still retaining the ability to combine probabilities of different sorts. Future investigations include decidability, implementability, and modelling applicability of P-$\mathcal{DL}$.

References

1. Giugno, R., Lukasiewicz, T.: P-SHOQ(D): A probabilistic extension of SHOQ(D) for probabilistic ontologies in the semantic web. In: JELIA. (2002) 86–97
2. Lukasiewicz, T.: Expressive probabilistic description logics. Artif. Intell. **172(6-7)** (2008) 852–883
3. Halpern, J.Y.: An analysis of first-order logics of probability. Artif Intell. **46** (1990) 311–350
4. Koller, D., Halpern, J.Y.: Irrelevance and conditioning in first-order probabilistic logic. In: Advances in Artificial Intelligence Conference. (1996) 569–576
5. Baader, F., Calvanese, D., McGuiness, D., Nardi, D., Patel-Schneider, P.F.: Description Logic Handbook. Cambridge University Press (2003)
6. Klinov, P., Parsia, B., Sattler, U.: On correspondences between probabilistic first-order and description logics. In: Description Logic Workshop. (2009)
7. Cheeseman, P.: An inquiry into computer understanding. Comp. Intell. **4** (1988) 58–66
8. Bacchus, F.: Lp, a logic for representing and reasoning with statistical knowledge. Comp. Intell. **6** (1990) 209–231
9. Bacchus, F.: Representing and reasoning with probabilistic knowledge. MIT Press (1990)
10. Abadi, M., Halpern, J.Y.: Decidability and expressiveness for first-order logics of probability. Inform. and Comp. **112**(1) (1994) 1–36
11. Reichenbach, H.: Theory of Probability. University of California Press, Berkeley (1949)