PrDLs: A New Kind of Probabilistic Description Logics About Belief*

Jia Tao, Zhao Wen, Wang Hanpin, and Wang Lifu

School of Electronics Engineering and Computer Science, Peking Univ.,Beijing 100871,China
jiat@sei.pku.edu.cn

Abstract. It is generally accepted that knowledge based systems would be smarter if they can deal with uncertainty. Some research has been done to extend Description Logics(DLs) towards the management of uncertainty, most of which concerned the *statistical information* such as "The probability that a randomly chosen bird flies is greater than 0.9". In this paper, we present a new kind of extended DLs to describe degrees of belief such as "The probability that all plastic objects float is 0.3". We also introduce the extended tableau algorithm for Pr$\mathcal{ALC}$ as an example to compute the probability of the implicit knowledge.

Keywords: Reasoning under Uncertainty,Description Logics, Model-based Reasoning, Ontology, Probabilistic, Knowledge Representation.

1 Introduction

It has often been noted that "probability" is a term with dual use: it can be applied to the frequency of occurrence of a specific property in a large sample of objects, and to the degree of belief granted to a proposition[1]. Therefore, the probabilistic extension of Description Logics should also cover these two semantics.

Recently, several probabilistic description logics have been developed to describe uncertainty such as P-CLASSIC[2] and P-$\mathcal{SHOQ}$(D)[3]. Those probabilistic description logics focus on capturing the statistical information about the world, since given some statistical information(say, that 90% of the individuals in concept(or class) C also in concept D), then we can imagine a chance setup in which a randomly chosen individual of C has probability 0.9 of being an individual of D[4].

Such kind of probabilistic description logics inevitably cannot describe *degree of belief*[5,6] such as *the probability that class* C *is a subclass of* D *is 0.9*. In this paper, we will introduce a kind of probabilistic description logics named PrDLs which could describe and reason on such kind of information. We will use $(C \sqsubseteq D)^{\alpha}(0 < \alpha \leq 1)$ to express the degree of belief that class C is a subclass of class D is α and $(a : C)^{\alpha}$ to express the degree of belief that individual a is in class C is α. In many application domains, this kind of probabilistic subsumption semantics is more appropriate. For example,

* Supported by the National Grand Fundamental Research 973 Program of China under Grant No.2002CB312006; the National Natural Science Foundation of China under Grant Nos. 60473058.

H.G. Okuno and M. Ali (Eds.): IEA/AIE 2007, LNAI 4570, pp. 644–654, 2007.

- Dealing with conflict ontologies. On the coming semantic web, there would be a lot of ontologies on the web which are created by different people. Due to the limitation of their knowlegde or other reasons, there must be many conflict axioms and assertions such as one ontology may define the company with no more than 5 employees is *small company* and another ontology defines the company with no more than 10 employees is *small company*. How to make use of such kind of conflict knowledge? One solution is to assign each axiom a probability as degree of belief. For example, we could assign degree of belief 0.5 to the former one and 0.4 to the latter one. The remaining fraction 0.1 is left for the other possible definition about *small company*.
- Making use of knowledge not being proved yet such as Goldbach Conjecture or could not be proved such as the definition of *small company*.
- In the *information integration* domain, using PrDLs to describe the degree of the similarity of the concepts in different ontologies or schemas which is called *ontology mapping*[7] or *schema matching*[8]. For example, if there is a concept *SmallCom* in schema A and a concept *SCompany* in schema B, we may draw a conclusion that the probability $SmallCom \equiv SCompany$ is 0.9. This conclusion should follow the degree of belief semantics but not statistical semantics, because if we apply the statistical semantics, we have to admit $SmallCom \not\equiv SCompany$.

Moreover, PrDLs are different from other probabilistic description logics such as P-CLASSIC not only in semantics, but also in reasoning algorithms. For example, in PrDLs, given $(A \equiv B)^{0.9}$ and $(B \equiv C)^{0.8}$, we can infer $(A \equiv C)^{0.72}$. On the other hand, in P-CLASSIC, the statement that the probability class A(resp. B) euqals class B(resp. C) is 0.9(resp. 0.8) could be written as $Pr(A \sqcap B | A \sqcup B) = 0.9$(resp. $Pr(B \sqcap C | B \sqcup C) = 0.8$). However, we cannot draw any conclusion about $Pr(A \sqcap C | A \sqcup C)$.

In the following we will introduce PrDLs in detail. Section 2 describes the foundations of PrDLs. Section 3 introduces the reasoning algorithm of a simple probabilistic description language Pr$\mathcal{ALC}$ which is extended from the $\mathcal{ALC}$ tableaux algorithm[9]. Section 4 is the related work. Section 5 is the conclusion and future work.

2 Foundations

PrDLs are a family of logic-based knowledge representation formalisms with probabilistic extension. They are based on a common family of languages which provide a set of constructors to build concept (class) and role (property) descriptions.

2.1 Probabilistic Description Languages

PrDLs have the exact same languages with the corresponding DLs. For example, PrDLs also have $\mathcal{ALC}$ and $\mathcal{SHOQ}$, etc. distinguished by the constructors they provide but named Pr$\mathcal{ALC}$ and Pr$\mathcal{SHOQ}$. You can find more introduction in [10].

2.2 Knowledge Base

A PrDL knowledge base is also composed of two distinct part: the TBox and the ABox, but extended by the probabilistic factors.

In this paper, we separate the TBox to the C(oncept)Box and R(ole)Box which contain the axioms about concept and roles, respectively.

CBox. In A PrDL CBox $\mathcal{C}$, an axiom may have the form (Mammal $\equiv$ Animal $\sqcap$ FourLegThing)$^{0.9}$ Which means the probability that the axiom is true is 0.9. Formally, let $\mathcal{L}$ be a probabilistic description language, $C, D \in$ **Cdsc**($\mathcal{L}$)$\mathcal{L}$-concepts[11], a CBox $\mathcal{C}$ is a finite, possibly empty, set of statements of the form $(C \sqsubseteq D)^{\alpha}, 0 < \alpha \leqslant 1$, called *concept inclusion*. **Cdsc**($\mathcal{L}$) and **Rdsc**($\mathcal{L}$) is defined as the set of concepts and roles of description language $\mathcal{L}$. α, the probabilistic weight of the statement, denotes the probability that the statement is true. $(C \equiv D)^{\alpha}$, called *concept equivalence*, is an statement denotes the probability that both $C \sqsubseteq D$ and $D \sqsubseteq C$ is true is α. Statements in $\mathcal{C}$ are called *concept axioms*. If $\alpha = 1$, the statements can be abbreviated to $C \sqsubseteq D$ or $C \equiv D$, called *certain concept axioms*. Otherwise, the statements are called *uncertain concept axioms*.

We can divide the CBox $\mathcal{C}$ into two parts $\mathcal{C}_d$ and $\mathcal{C}_p$. $\mathcal{C}_d$ consists of all the certain concept axioms, while $\mathcal{C}_p$ consists of all the uncertain concept axioms. Here we should not treat $(C \equiv D)^{\alpha}$ as the abbreviation for $(C \sqsubseteq D)^{\alpha}$ and $(D \sqsubseteq C)^{\alpha}$, since $(C \sqsubseteq D)^{\alpha}$ and $(D \sqsubseteq C)^{\alpha}$ imply $(C \equiv D)^{\beta}, max(2\alpha - 1, 0) \leqslant \beta \leqslant \alpha$, but not $(C \equiv D)^{\alpha}$. A concept axiom without its probabilistic weight is called the *certain extension* of the concept axiom. For example, $C \equiv D$ is the certain extension of $(C \equiv D)^{\alpha}$. The certain extension of a CBox $\mathcal{C}$ is a new CBox $\mathcal{C}_e$ whose concept axioms are all coming from the certain extension of the concept axioms in $\mathcal{C}$.

An interpretation $\mathcal{I}$ *satisfies* a concept inclusion $(C \sqsubseteq D)^{\alpha}$, or $\mathcal{I}$ *models* $(C \sqsubseteq D)^{\alpha}$ (written as $\mathcal{I} \models (C \sqsubseteq D)^{\alpha}$), if $\mathcal{I}$ *satisfies* its certain extension(which means $C^{\mathcal{I}} \subseteq D^{\mathcal{I}}$), and it satisfies a concept equivalence $(C \equiv D)^{\alpha}$ (written as $\mathcal{I} \models (C \equiv D)^{\alpha}$), if satisfies its certain extension (which means $C^{\mathcal{I}} = D^{\mathcal{I}}$). An interpretation $\mathcal{I}$ *possibly satisfies* a PrDL CBox $\mathcal{C}$(written as $\mathcal{I} \models \mathcal{C}$), if it satisfies all the certain concept axioms in $\mathcal{C}$. We also say that $\mathcal{I}$ is a *possible model* of $\mathcal{C}$.

RBox. We could give the similar definition about RBox of PrDLs. Let $\mathcal{L}$ be a probabilistic description language, $RN, SN \in \mathbf{R}$ role names, $R_1, R_2 \in$ **Rdsc**($\mathcal{L}$)$\mathcal{L}$-roles, an PrDL RBox $\mathcal{R}$ is a finite, possible empty, set of statements of the form:
- $(RN \in \mathbf{F})^{\alpha}, 0 < \alpha \leqslant 1$, where $\mathbf{F} \subseteq \mathbf{R}$ is a set of *functional roles*, or
- $(SN \in \mathbf{R}_{+})^{\alpha}, 0 < \alpha \leqslant 1$, where $\mathbf{R}_{+} \subseteq \mathbf{R}$ is a set of *transitive roles*, or
- $(R_1 \sqsubseteq R_2)^{\alpha}, 0 < \alpha \leqslant 1$, called *role inclusions*; $(R_1 \equiv R_2)^{\alpha}, 0 < \alpha \leqslant 1$, called *role equivalence*, denotes the probability that both $R_1 \sqsubseteq R_2$ and $R_2 \sqsubseteq R_1$ is true is α.

ABox. Let $\mathcal{L}$ be a probabilistic description language, $a, b \in \mathbf{I}$ individual names, $C \in$ **Cdsc**($\mathcal{L}$) an $\mathcal{L}$-concept and $R \in$ **Rdsc**($\mathcal{L}$) an $\mathcal{L}$-role. An PrDL ABox $\mathcal{A}$ is a finite, possible empty, set of statements of the form $(a : C)^{\alpha}$, called *concept assertions*, or $(< a, b >: R)^{\alpha}$, called *role assertions*. Statements in $\mathcal{A}$ are called *assertions*. If $\alpha = 1$, the assertions can be abbreviated to $a : C$ or $< a, b >: R$ and can be called *certain assertions*. Otherwise, they do not have abbreviated form and is called *uncertain assertions*.

Knowledge Base. A PrDL knowledge base $\mathcal{K}$ is a triple $< \mathcal{C}, \mathcal{R}, \mathcal{A} >$, where $\mathcal{C}$ is a CBox, $\mathcal{R}$ is a RBox, and $\mathcal{A}$ is an ABox. An interpretation $\mathcal{I}$ *satisfies* a knowledge base $\mathcal{K}$, written as $\mathcal{I} \models \mathcal{K}$, iff it satisfies $\mathcal{C}$, $\mathcal{R}$ and $\mathcal{A}$; $\mathcal{K}$ is *satisfiable* (*unsatisfiable*) iff there exists (does not exist) such an interpretation $\mathcal{I}$ that satisfies $\mathcal{K}$.

2.3 Semantics

We use possible worlds[12,13,14] to describe the semantics of the PrDLs. The approach is mapping a PrDL knowledge base onto a set of DL knowledge bases, where the models of each of the latter constitute the set of possible worlds(may be empty set). First, we give the definition of the DL knowledge bases related to a PrDL knowledge base.

Definition 1 (DL KBs Related to the PrDL KB). *Given a PrDL knowledge base* $\mathcal{K} =< C, \mathcal{R}, \mathcal{A} >$. C_d*,* $\mathcal{R}_d$ *and* $\mathcal{A}_d$ *are their certain parts, and* C_p*,* $\mathcal{R}_p$ *and* $\mathcal{A}_p$ *are their uncertain parts. Let* C_{pe}*,* $\mathcal{R}_{pe}$ *and* $\mathcal{A}_{pe}$ *be the certain extension of* C_p*,* $\mathcal{R}_p$ *and* $\mathcal{A}_p$*.The set of DLs* $D_{\mathcal{K}}$ *related to this PrDL is defined as*

$$D_{\mathcal{K}} = \{< C_d \cup C_i, \mathcal{R}_d \cup \mathcal{R}_i, \mathcal{A}_d \cup \mathcal{A}_i > | C_i \subseteq C_{pe} \wedge \mathcal{R}_i \subseteq \mathcal{R}_{pe} \wedge \mathcal{A}_i \subseteq \mathcal{A}_{pe}\}$$

Obviously, a model of the knowledge base of the DL in $D_{\mathcal{K}}$ is also a possible model of $\mathcal{K}$. All the models of the DL knowledge bases in $D_{\mathcal{K}}$ constitute the set of its possible worlds $\mathcal{W}_{\mathcal{K}}$. let $M = (\mathcal{W}_{\mathcal{K}}, \mu)$ denote a probability structure, where μ is a discrete probability distribution on $\mathcal{W}_{\mathcal{K}}$. Then we can define the notion of an extension $[t]_M$ of the term t (could be a concept description, the certain extension of an axiom in $C \cup \mathcal{R}$ or the certain extension of an assertion in $\mathcal{A}$) by means of the following rules. Let w be a world (possible model) of $\mathcal{W}_{\mathcal{K}}$, $\mathcal{K}_d$ a DL knowledge base related to $\mathcal{K}$.

1.If $\exists\alpha \in (0, 1], (t)^{\alpha} \in C \cup \mathcal{R} \cup \mathcal{A}$, then $[t]_M = \alpha = \mu(\cup_{\mathcal{K}_d \in D_{\mathcal{K}} \wedge w \models \mathcal{K}_d \wedge \mathcal{K}_d \models t}\{w\})$

2.else if t is a concept, $[t]_M = 1 - \mu(\cup_{\mathcal{K}_d \in D_{\mathcal{K}} \wedge w \models \mathcal{K}_d \wedge \mathcal{K}_d \not\models t}\{w\})$ $\mathcal{K}_d \not\models t$ denotes concept t is not satisfiable with respect to the knowledge base $\mathcal{K}_d$.

Then, we can define the probability that a concept is satisfiable.

Definition 2 (Concept satisfiability). *Given a PrDL knowledge base* $\mathcal{K}$ *and a concept* C*, the probability that* C *is satisfiable with respect to* $\mathcal{K}$ *is* α *iff* $[C]_M = \alpha$.

Example 1. Given a PrDL knowledge base $\mathcal{K} =< C, \Phi, \Phi >$, where $C = \{(\text{Animal} \sqsubseteq \text{Creature} \sqcap \text{MovableThing})^{0.8}, (\text{Mammal} \sqsubseteq \text{Animal} \sqcap \text{FourLegThing})^{0.9}\}$, we have a possible worlds distribution $M_1 = (\mathcal{W}, \mu)$:

$$\begin{aligned}
P(\mathcal{I}_1) = 0.05 : & \Delta^{\mathcal{I}_1} = \{a, b, c\}, \text{Creature}^{\mathcal{I}_1} = \{a, b, c\}, \text{MovableThing}^{\mathcal{I}_1} = \{a, b\}, \\
& \text{Animal}_1^{\mathcal{I}} = \{a\}, \text{FourLegThing}^{\mathcal{I}_1} = \{c\}, \text{Mammal}^{\mathcal{I}_1} = \{a\} \\
P(\mathcal{I}_2) = 0.75 : & \Delta^{\mathcal{I}_2} = \{a, b, c\}, \text{Creature}^{\mathcal{I}_2} = \{a, b, c\}, \text{MovableThing}^{\mathcal{I}_2} = \{a, b\}, \\
& \text{Animal}^{\mathcal{I}_2} = \{a\}, \text{FourLegThing}^{\mathcal{I}_2} = \{a\}, \text{Mammal}^{\mathcal{I}_2} = \{a\} \\
P(\mathcal{I}_3) = 0.15 : & \Delta^{\mathcal{I}_3} = \{a, b, c\}, \text{Creature}^{\mathcal{I}_3} = \{a, b, c\}, \text{MovableThing}^{\mathcal{I}_3} = \{a, b\}, \\
& \text{Animal}^{\mathcal{I}_3} = \{b, c\}, \text{FourLegThing}^{\mathcal{I}_3} = \{b\}, \text{Mammal}^{\mathcal{I}_3} = \{b\} \\
P(\mathcal{I}_4) = 0.05 : & \Delta^{\mathcal{I}_4} = \{a, b, c\}, \text{Creature}^{\mathcal{I}_4} = \{a, b, c\}, \text{MovableThing}^{\mathcal{I}_4} = \{a, b\}, \\
& \text{Animal}^{\mathcal{I}_4} = \{b, c\}, \text{FourLegThing}^{\mathcal{I}_4} = \{b\}, \text{Mammal}^{\mathcal{I}_4} = \{c\} \\
P(\mathcal{I}_k) = 0.00 : & \mathcal{I}_k \in \mathcal{W} \wedge k \neq 1, 2, 3, 4
\end{aligned}$$

In the possible world $\mathcal{I}_1$, only the first axiom is satisfiable and $\mathcal{I}_3$ only satisfies the second axiom. Both axioms are satisfiable in the possible world $\mathcal{I}_2$. So

$$[\text{Animal} \sqsubseteq \text{Creature} \sqcap \text{MovableThing}]_{M_1} = 0.8$$

$$[\text{Mammal} \sqsubseteq \text{Animal} \sqcap \text{FourLegThing}]_{M_1} = 0.9$$

And the probability that Mammal $\sqsubseteq$ Creature is 0.75 in this model (written as $M_1 \models (\text{Human} \sqsubseteq \text{Animal})^{0.75}$). Actually, the probability will be range from 0.7 to 0.8 with different probability distributions. But if we assume the independence of the terms in the knowledge base, PrDL would only yield a point value 0.72.

2.4 Reasoning Tasks

The following are the main reasoning tasks related to the PrDL knowledge base:

- **Terminology − Satisfiability**: Given a CBox or RBox or both, decide whether their certain part are satisfiable.
- **Assertion − Satisfiability**: Given an ABox $\mathcal{A}$, decide whether the certain part is satisfiable.
- **Concept − Satisfiability**: Given a knowledge base $\mathcal{K}$ and a concept C, compute the probability that C is satisfiable with respect to $\mathcal{K}$.
- **Concept − Subsumption**: Given a knowledge base $\mathcal{K}$ and concepts C, D, compute the probability that C is included in D.
- **Concept − Membership**: Given a knowledge base $\mathcal{K}$, an individual a and a concept C, compute the probability that $a : C$.
- **Role − Subsumption**: Given a knowledge base $\mathcal{K}$ and roles R, S, compute the probability that R is included in S.
- **Role − Membership**: Given a knowledge base $\mathcal{K}$, two individuals a, b and a role R, compute the probability that $< a, b >: R$.

From their definition, we know the first two reasoning tasks are exactly same with description logics.

3 Inference Algorithm

In this section, we will introduce the probabilistic extension of tableaux algorithm for the terminologies that only contain axioms whose certain extensions contain only *unique introductions*[11] and no cycles, called Pr-Tableaux-Algorithm. We only give the Pr-Tableaux-Algorithm for Pr$\mathcal{ALC}$ which is extended from the tableaux algorithm for $\mathcal{ALC}$. First, we will introduce some relative definitions.

Definition 3 (Keys and Their Boolean Algebra). **K** *is a set of identifiers, which also contains the special elements* $\bot$ *and* $\top$ *. Given a Pr$\mathcal{ALC}$ knowledge base* $\mathcal{K} =< C, \mathcal{R}, \mathcal{A} >$*, whose set of possible worlds is* $\mathcal{W}$*. Let* $\varepsilon : C \cup \mathcal{R} \cup \mathcal{A} \longrightarrow \mathbf{K} - \{\bot\}$*, with constrains:1.*$\forall (t)^\alpha \in C \cup \mathcal{R} \cup \mathcal{A}(\alpha = 1 \iff \varepsilon((t)^\alpha) = \top)$*;2.*$\forall (t)^\alpha, (t')^\beta \in C \cup \mathcal{R} \cup \mathcal{A}(\varepsilon((t)^\alpha) = \varepsilon((t')^\beta) \Longrightarrow (t = t' \wedge \alpha = \beta) \vee (\alpha = 1 \wedge \beta = 1))$. *We define* **KE** *to be the extension of* **K** *iff*

1.$\mathbf{K} \subseteq \mathbf{KE}$;

2.*if* $e_1, e_2 \in \mathbf{KE}$, $e_1 \wedge e_2 \in \mathbf{KE}$ *and* $e_1 \vee e_2 \in \mathbf{KE}$.

Let $\mathcal{B}(\mathbf{KE}, \bot, \top, \{\wedge, \vee\})$ *denote the boolean algebra over* **KE**. *Then we can define a mapping* $\omega : \mathbf{KE} \longrightarrow 2^{\mathcal{W}}$. *For any* $e, e_1, e_2 \in \mathbf{KE}$*,we define*

1.$\omega(\top) = \mathcal{W}$
2.$\omega(\bot) = \Phi$
3.$\omega(e_1 \wedge e_2) = \omega(e_1) \cap \omega(e_2)$
4.$\omega(e_1 \vee e_2) = \omega(e_1) \cup \omega(e_2)$
5.*if* $\varepsilon((t)^\alpha) = e$ *and* $(t)^\alpha \in C \cup \mathcal{R} \cup \mathcal{A}$, *then* $\omega(e) = \{w | \ w \models (t)^\alpha\}$

Moreover, we define the probability of an key expression $e \in \mathbf{KE}$ as $P(e) = \mu(\omega(e))$.

Since axioms may have probabilistic factors, the unfolding rules could not just be replaced by the right side of the axioms. The new unfolding rule should add a set Θ as the suffix to each concept name called *weight set*. The elements of Θ are the keys of axioms which contribute to generating the concept related to Θ. The concept with weight set is called *weighted concept*. We consider concept name CN to be the abbreviation of $CN \wr \{\top\}$. So weighted concept is the generalization of concept. Then, we will use the weighted concepts as the basic elements during the inference. Formally

Definition 4 (Weighted Concepts). *Given a $Pr\mathcal{ALC}$ CBox $\mathcal{C}$, let $CN \in \mathbf{C}$ and $\Theta \subseteq$ $\mathbf{KE}$. $CN \wr \Theta$ is called* weighted concept name. Weighted concept descriptions *in $Pr\mathcal{ALC}$ are formed according to following syntax rule:*

$$C, D \longrightarrow CN \wr \Theta | \top \wr \{\top\} | \bot \wr \{\top\} | \neg C | C \sqcap D | C \sqcup D | \exists R.C | \forall R.C$$

Let $\mathbf{WCdsc}(\mathcal{K})$ denotes the set of all weighted concepts of knowledge base $\mathcal{K}$. Then we can define a mapping $\xi : \mathbf{WCdsc}(\mathcal{K}) \rightarrow 2^{\mathbf{K}(\mathcal{K})}$ as

$$\xi(C) = \{e | e \text{ occurs in } C\}$$

Furthermore, we define $P(C) = P(\bigwedge \xi(C)) = \mu(\{e | e \text{ occurs in } C\})$. We can see the possibility structure M will affect the value of $P(C)$. Similarly, we could define *weighted roles* by the exactly same way. So we won't describe it here.

We consider $CN \wr \Theta$ and $CN \wr \Omega$ are different weighted concepts if $\Theta \neq \Omega$ and $RN \wr \Theta$ and $RN \wr \Omega$ are different weighted roles if $\Theta \neq \Omega$.

We should also update the unfolding rule as follows:

Definition 5 (Unfold-Rule). *Given a $Pr\mathcal{ALC}$ CBox $\mathcal{C}$, $CN \wr \Theta$ is a weighted concept names. If there exists an axiom $(CN \equiv C)^{\alpha} \in \mathcal{C}$ whose key is e, we can unfold $CN \wr \Theta$ with following rules:*
1. Replace CN by C;
2. Let $\Theta' = \Theta \cup \{e\}$;
3. Replace each concept name DN appeared in C by $DN \wr \Theta'$.

Example 2. Given A CBox $\mathcal{C} = \{(CN_1 \equiv CN_2 \sqcap CN_3)^{0.9}, (CN_2 \equiv \forall R.CN_4)^{0.3}\}$ $\varepsilon(CN_1 \equiv CN_2 \sqcap CN_3) = e_1, \varepsilon(CN_2 \equiv \forall R.CN_4) = e_2$. Concept $CN_1 \sqcup CN_4$ can be unfolded to

$$(\forall R.CN_4 \wr \{e_1, e_2\} \sqcap CN_3 \wr \{e_1\}) \sqcup CN_4 \wr \{\top\}$$

Definition 6 (Operator +). *Given a weighted concept description C and a weight set Θ, we define an operator "+" between them. The semantic of operation "+" as $C + \Theta = C'$, where C' is a weighted concept description derived from C by replacing each weighted concept name $CN \wr \Omega$ appeared in C by $CN \wr \Theta \cup \Omega$.*

Definition 7 (R-successor). *Given a completion tree, node y is called an R-Successor of node x if there exists some weighted role $R \wr \Theta \in \mathcal{L}(< x, y >)$.*

Given a concept D in NNF and a CBox $\mathcal{C}$, We assign a key e_i to each axiom in $\mathcal{C}$ following Definition 3. We can expansion D by the expansion rules shown in Table 1(C, C_1 and C_2 are weighted concepts). *The expansion procedure won't stop until there is no rule could be applied.*Actually, we could get a set of completion tree $\mathcal{T} = \{T_1, ..., T_n\}$

by $\sqcup$-rule. Next, we are able to compute the probability that concept description D is satisfiable according to these completion tree. First, we should redefine the *clash*.

Definition 8 (Possible Clash). *Let T be a completion tree for concept D. If for some $CN \in \mathbf{C}$ and x of T, $\{CN \wr \Theta, \neg CN \wr \Omega\} \subseteq \mathcal{L}(x)$, T is said to contain a* possible clash *c(written as $T \models_p c$).*

Let $\xi(c) = \Theta \cup \Omega$, then the probability clash c is true is $P(c) = P(\bigwedge \xi(c))$.

Given a weighted concept C of knowledge base $\mathcal{K}$, let $\mathcal{T}$ is the set of all the completion trees generated by the Table 1. If there exists a completion tree $T \in \mathcal{T}$ which has no possible clash, the probability C is satisfiable is 1(written as $[D]_M = 1$). Otherwise, 2. $[D]_M = 1-$the probability that every $T \in \mathcal{T}$ has a clash. So the equation can be rewritten to $[D]_M = \begin{cases} 1 & \exists T \in \mathcal{T}.T \text{ has no possible clash} \\ 1 - P(\bigwedge_{T \in \mathcal{T}}(\bigvee_{T \models_p c} c)) & else \end{cases}$

According to definition 8, $\bigwedge_{T \in \mathcal{T}}(\bigvee_{T \models_p c} c)$ could be translated to equivalent boolean expression of $\mathbf{KE}(\mathcal{K})$. So we could transform it to disjunction normal form(DNF) of keys. Let it be $K_1 \vee K_2 \vee \ldots \vee K_n$. If we assume the axioms are independent with each other, $P(\bigwedge_{T \in \mathcal{T}}(\bigvee\{c|T \models_p c\}))$ can be computed by the following formula

$$P(\bigwedge_{T \in \mathcal{T}}(\bigvee_{T \models_p c} c)) = P(K_1 \vee K_2 \vee \ldots \vee K_n) = \sum_{i=1}^{n}(-1)^{i-1}(\sum_{1 \leqslant j_1 < \ldots < j_i \leqslant n} P(e_{j_1} \wedge \ldots \wedge e_{j_i}))$$

We could change the form of $\bigwedge_{T \in \mathcal{T}}(\bigvee_{T \models_p c} c$ to $\bigvee_{i=1}^{n} \bigwedge\{c_{i1}, ..., c_{im}\}, m = |\mathcal{T}|, n = \prod_{T \in \mathcal{T}} |\{c|c \in T\}|$. $\{c_{i1}, ..., c_{im}\}$ is a set of the possible clashes and each of them comes from different completion trees in $\mathcal{T}$. The meaning of such kind of set is a clash composition that could make the concept unsatisfiable. The probability that any one of such kind of composition is true is the probability that the concept is unsatisfiable. We call such composition *possible clash composition*. Given a possible clash composition ψ, We define

$$\xi(\psi) = \bigcup_{c \in \psi} \xi(c).$$

Finally, we can prove that the probability computed by the Pr-Tableaux-Algorithm is equal to the concept extensions introduced in 2.3.

Table 1. The tableaux expansion rules for Pr$\mathcal{ALC}$

Name	Condition	Action
$\sqcap$-rule	1.$C_1 \sqcap C_2 \in \mathcal{L}(x)$, 2.$\{C_1, C_2\} \nsubseteq \mathcal{L}(x)$	$\mathcal{L}(x) = \mathcal{L}(x) \cup \{C_1, C_2\}$
$\sqcup$-rule	1.$C_1 \sqcup C_2 \in \mathcal{L}(x)$, 2.$\{C_1, C_2\} \cap \mathcal{L}(x) = \Phi$	$\mathcal{L}(x) = \mathcal{L}(x) \cup \{C\}$ for some $C \in \{C_1, C_2\}$
$\exists$-rule	1. $\exists R.C \in \mathcal{L}(x)$, 2. x has no R$\wr\xi(C)$-successor with $C \in \mathcal{L}(y)$	create a new node y with $\mathcal{L}(< x, y >) = R \wr \xi(C)$ and $\mathcal{L}(y) = \{C\}$
$\forall$-rule	1. $\forall R.C \in \mathcal{L}(x)$, 2. there is an R$\wr\Theta$-successor y of x with $C \notin \mathcal{L}(y)$	$\mathcal{L}(y) = \mathcal{L}(y) \cup \{C\}$
Unfold-rule	1.$CN \wr \Theta \in \mathcal{L}(x)$(resp. $\neg CN \wr \Theta \in \mathcal{L}(x)$), 2.$(CN \equiv C)^{\alpha} \in C$, $\varepsilon((CN \equiv C)^{\alpha}) = e$, and $C + \{\Theta \cup \{e\}\} \notin \mathcal{L}(x)$ (resp. $\sim C + \{\Theta \cup \{e\}\} \notin \mathcal{L}(x)$)	$\mathcal{L}(x) = \mathcal{L}(x) \cup \{C + \{\Theta \cup \{e\}\}\}$ (resp. $\mathcal{L}(x) = \mathcal{L}(x) \cup \{\sim C + \{\Theta \cup \{e\}\}\}$)

Lemma 1. *Given a Pr$\mathcal{ALC}$ knowledge base $\mathcal{K} =< C, \Phi, \Phi >$ with the probability structure $M =< \mathcal{W}, \mu >$ and the set of related $\mathcal{ALC}$ knowledge bases $D_{\mathcal{K}} = \{\mathcal{K}_1^D, ..., \mathcal{K}_n^D\}$, let $D_{\mathcal{K}}^C = \{\mathcal{K}_d =< C_d, \Phi, \Phi > |\mathcal{K}_d \nvDash C \wedge \mathcal{K}_d \in D_{\mathcal{K}}\}$. Then we have*

$$[C]_M = 1 - P\left(\bigvee_{\mathcal{K}_d \in D_{\mathcal{K}}^C} \left(\bigwedge_{t \in C_d \wedge (t)^\alpha \in C} \varepsilon((t)^\alpha)\right)\right).$$

Proof. According to Definition 3, For any $\mathcal{K}_d \in D_{\mathcal{K}}^C$, we have $\omega\left(\bigwedge_{t \in C_d \wedge (t)^\alpha \in C} \varepsilon((t)^\alpha)\right) = \{w | w \vDash \mathcal{K}_d\}$ then,

$$\begin{aligned} [C]_M &= 1 - \mu(\cup_{\mathcal{K}_d \in D_{\mathcal{K}}^C} \{w | w \vDash \mathcal{K}_d\}) \\ &= 1 - \mu\left(\cup_{\mathcal{K}_d \in D_{\mathcal{K}}^C} \omega\left(\bigwedge_{t \in C_d \wedge (t)^\alpha \in C} \varepsilon((t)^\alpha)\right)\right) \\ &= 1 - \mu\left(\omega\left(\bigvee_{\mathcal{K}_d \in D_{\mathcal{K}}^C} \left(\bigwedge_{t \in C_d \wedge (t)^\alpha \in C} \varepsilon((t)^\alpha)\right)\right)\right) \\ &= 1 - P\left(\bigvee_{\mathcal{K}_d \in D_{\mathcal{K}}^C} \left(\bigwedge_{t \in C_d \wedge (t)^\alpha \in C} \varepsilon((t)^\alpha)\right)\right) \end{aligned}$$

For simplicity, we define $\xi(\mathcal{K}_d) = \{\varepsilon((t)^\alpha) | t \in C_d \wedge (t)^\alpha \in C\}$, then $[C]_M = 1 - P(\bigvee_{\mathcal{K}_d \in D_{\mathcal{K}}^C} (\bigwedge \xi(\mathcal{K}_d))$

Lemma 2. *Given a Pr$\mathcal{ALC}$ knowledge base $\mathcal{K} =< C, \Phi, \Phi >$ with the probability structure $M =< \mathcal{W}, \mu >$ and the set of related $\mathcal{ALC}$ knowledge bases $D_{\mathcal{K}} = \{\mathcal{K}_1^D, ..., \mathcal{K}_n^D\}$, let $D_{\mathcal{K}}^C = \{\mathcal{K}_d =< C_d, \Phi, \Phi > |\mathcal{K}_d \nvDash C \wedge \mathcal{K}_d \in D_{\mathcal{K}}\}$. For any $\mathcal{K}_d \in D_{\mathcal{K}}^C$, its clash composition ψ_d generated by the $\mathcal{ALC}$ tableau algorithm has corresponding possible clash composition ψ generated by the Pr$\mathcal{ALC}$ tableau algorithm (which means each possible clash c in ψ has a corresponding clash c_d in ψ_d only without weight set) and $\xi(\psi) \subseteq \xi(\mathcal{K}_d)$*

Proof. First, we could prove that each completion tree T_d of C with respect to $\mathcal{K}_d \in D_{\mathcal{K}}^C$ which is generated by the $\mathcal{ALC}$ tableau algorithm is a "sub-tree"(with same root node) of some T generated by Pr$\mathcal{ALC}$ tableau algorithm without considering the weight set, since Pr$\mathcal{ALC}$ tableau algorithm is the extension of the $\mathcal{ALC}$ tableau algorithm. Similarly, each T generated by Pr$\mathcal{ALC}$ tableau algorithm must have a corresponding T_d of C with respect to $\mathcal{K}_d$ which is a "sub-tree" of T. So each concept D occurred in each node of T_d could be found its weighted version $D \wr \Theta$ in the corresponding node of T and $\Theta \subseteq \xi(\mathcal{K}_d)$. Then each clash c_d occurred in T_d has a corresponding possible clash c in T and $\xi(c) \subseteq \xi(\mathcal{K}_d)$. Consequently, the clash composition ψ_d of T_d has a corresponding possible clash composition ψ of T and $\xi(\psi) \subseteq \xi(\mathcal{K}_d)$.

Lemma 3. *Given a Pr$\mathcal{ALC}$ knowledge base $\mathcal{K} =< C, \Phi, \Phi >$ with the probability structure $M =< \mathcal{W}, \mu >$, a concept C and the set of related $\mathcal{ALC}$ knowledge bases $D_{\mathcal{K}} = \{\mathcal{K}_1^D, ..., \mathcal{K}_n^D\}$, let $D_{\mathcal{K}}^C = \{\mathcal{K}_d =< C_d, \Phi, \Phi > |\mathcal{K}_d \nvDash C \wedge \mathcal{K}_d \in D_{\mathcal{K}}\}$. If ψ is a possible clash composition generated by Pr$\mathcal{ALC}$, then there is a related $\mathcal{ALC}$ knowledge base $\mathcal{K}_d$ with $\xi(\mathcal{K}_d) = \xi(\psi)$ and $\mathcal{K}_d \in D_{\mathcal{K}}^C$.*

Proof. According to the tableau algorithm of Pr$\mathcal{ALC}$ we have introduced, the set of the axioms $\rho = \{(t)^\alpha | (t)^\alpha \in C \wedge \varepsilon((t)^\alpha) \in \xi(\psi)\}$ is sufficient for generating the possible clash composition ψ. So the related $\mathcal{ALC}$ knowledge base $\mathcal{K}_d$ whose axioms all come from the certain extension of the axioms in ρ must be able to generate a corresponding clash composition by the $\mathcal{ALC}$ tableau algorithm since it is the specialization of Pr$\mathcal{ALC}$ tableau algorithm. Then, $\mathcal{K}_d \in D^C_{\mathcal{K}}$.

Theorem 1 (Correctness of the Algorithm). *The probability computed by the Pr-Tableaux-Algorithm introduced in this section is equal to the concept extensions defined in section 2.3*

Proof. According to *Lemma 1*, we only need to prove that $P(\bigvee_{\mathcal{K}_d \in D^C_{\mathcal{K}}}(\bigwedge \xi(\mathcal{K}_d))) = P(\bigvee_{\text{each } \psi \text{ of } \mathcal{T}} \bigwedge \psi)$ where ψ is a possible clash composition.

According to *Lemma 2*, we could obtain $P(\bigvee_{\mathcal{K}_d \in D^C_{\mathcal{K}}}(\bigwedge \xi(\mathcal{K}_d))) \leq P(\bigvee_{\text{each } \psi \text{ of } \mathcal{T}} \bigwedge \psi)$

According to *Lemma 3*, we could get $P(\bigvee_{\mathcal{K}_d \in D^C_{\mathcal{K}}}(\bigwedge \xi(\mathcal{K}_d))) \geq P(\bigvee_{\text{each } \psi \text{ of } \mathcal{T}} \bigwedge \psi)$

4 Related Work

Halpern et al. have done much research on degree of belief(subject probability) and statistical information(object probability)[15,16]. They mainly focus on the relation ship between these two kind of uncertainty[16,17,18], belief change[19,20] and probabilistic reasoning[13,4]. For description logics, Heinsohn[21] presents a probabilistic extension of the description logic $\mathcal{ALC}$, which allows to represent generic statistical information about concepts and roles, and which is essentially based on probabilistic reasoning in probabilistic logics, similar to[14,22]. Jaeger[23] gives a probabilistic extension of the description logic, which allows for generic (resp., assertional) statistical information about concepts and roles (resp., concept instances), but does not support statistical information about role instances. The uncertain reasoning formalism in [23] is essentially based on probabilistic reasoning in probabilistic logics, as the one in [21]. The work by Koller et al. [2] gives a probabilistic generalization of the CLASSIC description logic. Like Heinsohn's work [21], it is based on inference in Bayesian networks as underlying probabilistic reasoning formalism. Giugno presents a probabilistic extension of $\mathcal{SHOQ}$(D)[3], which allows to represent generic statistical knowledge about concept and roles and the assertional statistical knowledge about concept and role instance. Baader extends Description Logics with modal operators in [24]to describe belief but not degree of belief.

5 Conclusion and Future Work

We have presented a probabilistic version of description logics–PrDLs which are used to represent the degree of belief of the axioms and assertions in the knowledge base, which are very useful in many application area. We also introduce an inference algorithm for Pr$\mathcal{ALC}$ to discover the possible implicit knowledge.

In future, we will improve our work in two aspects. First, develop the inference algorithms which are suitable for more expressive probabilistic description languages such as the knowledge bases with none empty RBox, the knowledge bases with general inclusions and so on; Second, combine PrDLs with other probabilistic description logics describing statistical information.

References

1. Jaeger, M.: A logic for default reasoning about probabilities. Uncertainty in Artificial Intelligence, 352–359 (1994)
2. Koller, D., Levy, A.Y., Pfeffer, A.: P-CLASSIC: A tractable probablistic description logic. In: AAAI/IAAI, pp. 390–397 (1997)
3. Giugno, R., Lukasiewicz, T.: P-SHOQ(D): A probabilistic extension of SHOQ(D) for probabilistic ontologies in the semantic web. In: Flesca, S., Greco, S., Leone, N., Ianni, G. (eds.) JELIA 2002. LNCS (LNAI), vol. 2424, pp. 86–97. Springer, Heidelberg (2002)
4. Halpern, J.Y.: An analysis of first-order logics of probability. In: IJCAI, pp. 1375–1381 (1989)
5. Poole, D.: Representing diagnostic knowledge for probabilistic horn abduction. In: IJCAI, pp. 1129–1137 (1991)
6. Kyburg Jr., H.E.: Higher order probabilities and intervals. Int. J. Approx. Reasoning 2(3), 195–209 (1988)
7. Noy, N.F.: Semantic integration: A survey of ontology-based approaches. SIGMOD Record 33(4), 65–70 (2004)
8. Rahm, E., Bernstein, P.A.: A survey of approaches to automatic schema matching. VLDB J 10(4), 334–350 (2001)
9. Baader, F., Nutt, W.: Basic description logics. In: Baader, F., Calvanese, D., McGuinness, D.L., Nardi, D., Patel-Schneider, P.F. (eds.) Description Logic Handbook, pp. 43–95. Cambridge University Press, Cambridge (2003)
10. Schmidt-Schauß, M., Smolka, G.: Attributive concept descriptions with complements. Artif. Intell. 48(1), 1–26 (1991)
11. Pan, J.Z.: Description Logics: Reasoning Support for the Semantic Web. PhD thesis, School of Computer Science, The University of Manchester, Oxford Rd, Manchester M13 9PL, UK (2004)
12. Fagin, R., Halpern, J.Y.: Reasoning about knowledge and probability. J. ACM 41(2), 340–367 (1994)
13. Fagin, R., Halpern, J.Y., Megiddo, N.: A logic for reasoning about probabilities. Inf. Comput. 87(1/2), 78–128 (1990)
14. Nilsson, N.: Probabilistic logic. ai 28, 71–87 (1986)
15. Halpern, J.Y.: Reasoning about knowledge: A survey. In: Gabbay, D.M., Hogger, C.J., Robinson, J.A. (eds.) Handbook of Logic in Artificial Intelligence and Logic Programming, Epistemic and Temporal Reasoning, vol. 4, pp. 1–34. Clarendon Press, Oxford (1995)
16. Bacchus, F., Grove, A.J., Halpern, J.Y., Koller, D.: From statistical knowledge bases to degrees of belief. Artif. Intell. 87(1-2), 75–143 (1996)
17. Fagin, R., Halpern, J.Y.: Uncertainty, belief, and probability. In: IJCAI, pp. 1161–1167 (1989)
18. Bacchus, F., Grove, A.J., Halpern, J.Y., Koller, D.: Generating degrees of belief from statistical information: An overview. In: Shyamasundar, R.K. (ed.) Foundations of Software Technology and Theoretical Computer Science. LNCS, vol. 761, pp. 318–325. Springer, Heidelberg (1993)
19. Friedman, N., Halpern, J.Y.: Modeling belief in dynamic systems, part I: Foundations. Artif. Intell. 95(2), 257–316 (1997)
20. Friedman, N., Halpern, J.Y.: Modeling belief in dynamic systems, part II: Revision and update. J. Artif. Intell. Res. (JAIR) 10, 117–167 (1999)

21. Heinsohn, J.: Probabilistic description logics. In: de Mantaras, R.L., Poole, D. (eds.) Proceedings of the 10th Conference on Uncertainty in Artificial Intelligence, pp. 311–318. Morgan Kaufmann, San Francisco (1994)
22. Lukasiewicz, T.: Probabilistic deduction with conditional constraints over basic events. In: KR, pp. 380–393 (1998)
23. Jaeger, M.: Probabilistic reasoning in terminological logics. In: KR, pp. 305–316 (1994)
24. Baader, F., Laux, A.: Terminological logics with modal operators. IJCAI (1), 808–815 (1995)