Epistemic and Statistical Probabilistic Ontologies

Fabrizio Riguzzi, Elena Bellodi, Evelina Lamma, and Riccardo Zese

ENDIF – University of Ferrara, Via Saragat 1, I-44122, Ferrara, Italy
{fabrizio.riguzzi,elena.bellodi,evelina.lamma}@unife.it
riccardo.zese@student.unife.it

Abstract. We present DISPONTE, a semantics for probabilistic ontologies that is based on the distribution semantics for probabilistic logic programs. In DISPONTE the axioms of a probabilistic ontology can be annotated with an epistemic or a statistical probability. The epistemic probability represents a degree of confidence in the axiom, while the statistical probability considers the populations to which the axiom is applied.

1 Introduction

Uncertainty has been recognized as an important feature for the Semantic Web [36, 24]. In order to be able to represent and reason with probabilistic knowledge, various authors have advocated the use of probabilistic ontologies, see e.g. [26], and many proposals have been put forward for allowing ontology languages, and OWL in particular, to represent uncertainty [30, 10, 22, 23].

The integration of probability into logic has been much studied lately, with many different proposals. In the field of logic programming, the distribution semantics [33] has emerged as one of the most effective approaches [33, 28, 8].

In [3] we applied this approach to ontological languages and, in particular, to the OWL DL fragment, that is based on the description logic $\mathcal{SHOIN}(\mathbf{D})$. We called the approach DISPONTE for "DIstribution Semantics for Probabilistic ONTologiEs" (Spanish for "get ready"). The idea is to annotate axioms of a theory with a probability and assume that each axiom is independent of the others. In this paper we extend DISPONTE by considering two types of probabilistic annotations, an epistemic type, that represents a degree of belief in the axiom as a whole as in [3], and a new statistical type, that considers the populations to which the axiom is applied. Statistical probabilities allow to represent partial concept overlapping and knowledge on random individuals of populations. The two types of probability can be used separately or jointly in the same OWL DL knowledge base. The probability of a query is computed from a covering set of explanations by solving the disjoint sum problem.

The paper is organized as follows. Section 2 introduces Description Logics. Section 3 presents DISPONTE and explains the principles on which it is based, while Section 4 illustrates approaches for performing reasoning. Section 5 describes related work and, finally, Section 6 concludes the paper.

2 Description Logics

Description Logics (DLs) are knowledge representation formalisms that possess nice computational properties such as decidability and/or low complexity, see [1, 2] for excellent introductions. DLs are particularly useful for representing ontologies and have been adopted as the basis of the Semantic Web. For example, the OWL DL sublanguage of OWL is based on the $\mathcal{SHOIN}(\mathbf{D})$ DL.

While DLs can be translated into predicate logic, they are usually represented using a syntax based on concepts and roles. A concept corresponds to a set of individuals of the domain while a role corresponds to a set of couples of individuals of the domain. In order to illustrate DLs, we now describe $\mathcal{SHOIN}$ following [24].

Let $\mathbf{A}$, $\mathbf{R}$ and $\mathbf{I}$ be sets of *atomic concepts*, *roles* and *individuals*, respectively. A *role* is either an atomic role $R \in \mathbf{R}$ or the inverse R^- of an atomic role $R \in \mathbf{R}$. We use $\mathbf{R}^-$ to denote the set of all inverses of roles in $\mathbf{R}$. An RBox $\mathcal{R}$ consists of a finite set of *transitivity axioms $Trans(R)$*, where $R \in \mathbf{R}$, and *role inclusion axioms $R \sqsubseteq S$*, where $R, S \in \mathbf{R} \cup \mathbf{R}^-$.

Concepts are defined by induction as follows. Each $A \in \mathbf{A}$ is a concept, $\bot$ and $\top$ are concepts, and if $a \in \mathbf{I}$, then $\{a\}$ is a concept. If C, $C1$ and $C2$ are concepts and $R \in \mathbf{R} \cup \mathbf{R}^-$, then $(C_1 \sqcap C_2)$, $(C_1 \sqcup C_2)$, and $\neg C$ are concepts, as well as $\exists R.C$, $\forall R.C$, $n \geq R$ and $n \leq R$ for an integer $n \geq 0$.

A *TBox $\mathcal{T}$* is a finite set of *concept inclusion axioms $C \sqsubseteq D$*, where C and D are concepts. We use $C \equiv D$ to abbreviate $C \sqsubseteq D$ and $D \sqsubseteq C$. An *ABox $\mathcal{A}$* is a finite set of *concept membership axioms $a : C$*, *role membership axioms $(a, b) : R$*, *equality axioms $a = b$*, and *inequality axioms $a \neq b$*, where C is a concept, $R \in \mathbf{R}$ and $a, b \in \mathbf{I}$. A *knowledge base $\mathcal{K} = (\mathcal{T}, \mathcal{R}, \mathcal{A})$* consists of a TBox $\mathcal{T}$, an RBox $\mathcal{R}$ and an ABox $\mathcal{A}$.

$\mathcal{SHOIN}$ is decidable iff there are no number restrictions on non-simple roles. A role is non-simple iff it is transitive or it has transitive subroles.

A knowledge base $\mathcal{K}$ is usually assigned a semantics in terms of set-theoretic interpretations and models of the form $\mathcal{I} = (\Delta^{\mathcal{I}}, \cdot^{\mathcal{I}})$. The semantics of DLs can be given equivalently by transforming a DL knowledge base into a predicate logic theory and then using the model-theoretic semantics of the resulting theory. A translation of $\mathcal{SHOIN}$ into First Order Logic with Counting Quantifiers is given in the following as an extension of the one given in [34]. The translation uses two functions π_x and π_y that map concept expressions to logical formulas, where π_x is given by

$$\pi_x(A) = A(x)$$
$$\pi_x(\neg C) = \neg \pi_x(C)$$
$$\pi_x(\{a\}) = (x = a)$$
$$\pi_x(C \sqcap D) = \pi_x(C) \wedge \pi_x(D)$$
$$\pi_x(C \sqcup D) = \pi_x(C) \wedge \pi_x(D)$$

$$\pi_x(\exists R.C) = \exists y.R(x,y) \land \pi_y(C)$$
$$\pi_x(\exists R^-.C) = \exists y.R(y,x) \land \pi_y(C)$$
$$\pi_x(\forall R.C) = \forall y.R(x,y) \rightarrow \pi_y(C)$$
$$\pi_x(\forall R^-.C) = \forall y.R(y,x) \rightarrow \pi_y(C)$$
$$\pi_x(\geq nR) = \exists^{\geq n}y.R(x,y)$$
$$\pi_x(\geq nR^-) = \exists^{\geq n}y.R(y,x)$$
$$\pi_x(\leq nR) = \exists^{\leq n}y.R(x,y)$$
$$\pi_x(\leq nR^-) = \exists^{\leq n}y.R(y,x)$$

and π_y is obtained from π_x by replacing x with y and vice-versa.

Table 1 shows the translation of each axiom of $\mathcal{SHOIN}$ knowledge bases.

Axiom	Translation
$C \sqsubseteq D$	$\forall x.\pi_x(C) \rightarrow \pi_x(D)$
$R \sqsubseteq S$	$\forall x,y.R(x,y) \rightarrow S(x,y)$
$Trans(R)$	$\forall x,y,z.R(x,z) \land R(z,y) \rightarrow S(x,y)$
$a : C$	$C(a)$
$(a,b) : R$	$R(a,b)$
$a = b$	$a = b$
$a \neq b$	$a \neq b$

Table 1. Translation of $\mathcal{SHOIN}$ axioms into predicate logic.

$\mathcal{SHOIN}(\mathbf{D})$ adds to $\mathcal{SHOIN}$ datatype roles, i.e., roles that map an individual to an element of a datatype such as integers, floats, etc. Then new concept definitions, involving datatype roles, are added, that mirror those involving roles introduced above. We also assume that we have predicates over the datatypes.

A query over a knowledge base is usually an axiom for which we want to test the entailment from the knowledge base. The entailment test may be reduced to checking the satisfiability of a concept in the knowledge base, i.e., the non-emptiness of the concept. For example, the entailment of the axiom $C \sqsubseteq D$ may be tested by checking the satisfiability of the concept $C \sqcap \neg D$.

3   The DISPONTE Semantics for Probabilistic Ontologies

A *probabilistic knowledge base* is a set of *certain axioms*, that take the form of DL axioms, of *epistemic probabilistic axioms* of the form

$$p ::_e E \tag{1}$$

where p is a real number in $[0,1]$ and E is a TBox, RBox or ABox axiom, and of *statistical probabilistic axioms* of the form

$$p ::_s E \tag{2}$$

where p is a real number in $[0, 1]$ and E is a TBox or RBox axiom.

In axioms of the form (1), the notation $::_e$ introduces probability p as an epistemic probability, i.e., as the degree of our belief in axiom E, while in axioms of the form (2) the notation $::_s$ interprets p as a statistical probability, i.e., as information regarding random individuals from certain populations. For example, an epistemic probabilistic concept inclusion axiom of the form

$$p ::_e C \sqsubseteq D \tag{3}$$

represents the fact that we believe in the truth of $C \sqsubseteq D$ with probability p. A statistical probabilistic concept inclusion axiom of the form

$$p ::_s C \sqsubseteq D \tag{4}$$

instead means that a random individual of class C has probability p of belonging to D, thus representing the statistical information that a fraction p of the individuals of C belong to D. In this way, the overlap between C and D is quantified by the statistical probability p. The difference between the two axioms is that, if two individuals belong to class C, the probability that they both belong to D according to (3) is p, since p represents the truth of the formula as a whole, while according to (4) is $p \cdot p$, since each individual has probability p of belonging to class D and the two events are independent. Therefore, statistical probability regards the knowledge we have about population of a given domain.

In order to give a semantics to such probabilistic knowledge bases, we consider their translation into predicate logic. The idea of DISPONTE is to associate independent Boolean random variables to (instantiations of) the formulas in predicate logic that are obtained from the translation of the axioms. By assigning values to every random variable we obtain a *world*, the set of predicate logic formulas whose random variable is assigned to 1.

To clarify what we mean by instantiations, we assume a fixed interpretation domain $\Delta^{\mathcal{I}}$ that is non-empty and possibly infinite. Given a predicate logic formula F and a domain $\Delta^{\mathcal{I}}$, we define a *substitution* θ as a set of couples x/i where x is a variable universally quantified in the outermost quantifier in F and $i \in \Delta^{\mathcal{I}}$. The application of θ to F, indicated by $F\theta$, is called an *instantiation of F* and is obtained by replacing x with i in F and by removing x from the external quantification for every couple x/i in θ. By instantiating the universally qualified variables we are able to separately represent each individual to which the axiom is applied.

To obtain a world w of a probabilistic knowledge base T, we translate every axiom into a predicate logic formula and we replace each individual a appearing in the knowledge base $\mathcal{K}$ with $a^{\mathcal{I}}$. Every formula obtained from an axiom without a probability annotation is included in w. For each axiom of the form (1), we decide whether or not to include it in w. For each axiom of the form (2), we generate all the substitutions for the variables of the equivalent predicate logic formula indicated in Table 2.

There may be an infinite number of instantiations. For each instantiated formula we decide whether or not to include it in w. In this way we obtain

Axiom	Variables to be instantiated
$p ::_s C \sqsubseteq D$	x
$p ::_s R \sqsubseteq S$	x, y
$p ::_s Trans(R)$	x, y, z

Table 2. Axiom variables to be instantiated.

a predicate logic theory which can be assigned a model-theoretic semantics. A query is entailed by a world if it is true in every model of the world.

To formally define the semantics we follow the approach of [28]. An *atomic choice* in this context is a triple (F_i, θ_j, k) where F_i is the formula obtained by translating the ith axiom, θ_j is a substitution and $k \in \{0, 1\}$. k indicates whether (F_i, θ_j, k) is chosen to be included in a world ($k = 1$) or not ($k = 0$). If F_i is obtained from an unannotated axiom, then $\theta_j = \emptyset$ and $k = 1$. If F_i is obtained from an axiom of the form (1), then $\theta_j = \emptyset$. If F_i is obtained from an axiom of the form (2), then θ_j instantiates the variables indicated in Table 2. Note that, differently from [28], substitutions do not ground formulas but this is not a core requirement of [28].

A *composite choice* κ is a consistent set of atomic choices, i.e., $(F_i, \theta_j, k) \in \kappa, (F_i, \theta_j, m) \in \kappa \Rightarrow k = m$ (only one decision for each formula). The probability of composite choice κ is $P(\kappa) = \prod_{(F_i, \theta_j, 1) \in \kappa} p_i \prod_{(F_i, \theta_j, 0) \in \kappa} (1 - p_i)$. A *selection* σ is a total composite choice, i.e., it contains an atomic choice (F_i, θ_j, k) for every instantiation $F_i \theta_j$ of every formula of the theory. Since the domain may be infinite, selections may, too. Let us indicate with $\mathcal{S}_T$ the set of all selections. A selection σ identifies a theory w_σ called a *world* in this way: $w_\sigma = \{F_i \theta_j | (F_i, \theta_j, 1) \in \sigma\}$. Let us indicate with $\mathcal{W}_T$ the set of all worlds. A composite choice κ identifies a set of worlds $\omega_\kappa = \{w_\sigma | \sigma \in \mathcal{S}_T, \sigma \supseteq \kappa\}$. We define the set of worlds identified by a set of composite choices K as $\omega_K = \bigcup_{\kappa \in K} \omega_\kappa$.

A composite choice κ is an *explanation* for a query Q if Q is entailed by every world of ω_κ. A set of composite choices K is *covering* with respect to Q if every world w_σ in which Q is entailed is such that $w_\sigma \in \omega_K$. Two composite choices κ_1 and κ_2 are *incompatible* if their union is inconsistent. A set K of composite choices is *mutually incompatible* if for all $\kappa_1 \in K, \kappa_2 \in K, \kappa_1 \neq \kappa_2 \Rightarrow \kappa_1$ and κ_2 are incompatible. We define *the probability of a mutually incompatible set of composite choices* K as $P(K) = \sum_{\kappa \in K} P(\kappa)$. Two sets of composite choices K_1 and K_2 are *equivalent* if $\omega_{K_1} = \omega_{K_2}$, i.e., if they identify the same set of worlds.

Kolmogorov defined probability functions (or measures) as real-valued functions over an algebra Ω of subsets of a set $\mathcal{W}$ called the *sample space*. The set Ω is an algebra of $\mathcal{W}$ iff (1) $\mathcal{W} \in \Omega$, (2) Ω is closed under complementation, i.e., $\omega \in \Omega \rightarrow (\mathcal{W} \setminus \omega) \in \Omega$ and (3) Ω is closed under finite union, i.e., $\omega_1 \in \Omega, \omega_2 \in \Omega \rightarrow (\omega_1 \cup \omega_2) \in \Omega$. The elements of Ω are called *measurable sets*. Not every subset of $\mathcal{W}$ need be present in Ω.

Given a sample space $\mathcal{W}$ and an algebra Ω of subsets of $\mathcal{W}$, a probability measure is a function $\mu : \Omega \rightarrow R$ that satisfies the following axioms: (1) $\mu(\omega) \geq 0$

for all $\omega \in \Omega$, (2) $\mu(\mathcal{W}) = 1$, (3) $\omega_1 \cap \omega_2 = \emptyset \rightarrow \mu(\omega_1 \cup \omega_2) = \mu(\omega_1) + \mu(\omega_2)$ for all $\omega_1 \in \Omega, \omega_2 \in \Omega$.

If $F\theta$ is an instantiated formula and κ is a composite choice such that $\kappa \cap \{(F,\theta,0),(F,\theta,1)\} = \emptyset$, the *split* of κ on $F\theta$ is the set of composite choices $S_{F\theta,\kappa} = \{\kappa \cup \{(F,\theta,0)\}, \kappa \cup \{(F,\theta,1)\}\}$. It is easy to see that κ and $S_{F\theta,\kappa}$ identify the same set of possible worlds, i.e., that $\omega_\kappa = \omega_{S_{F\theta,\kappa}}$.

Following [28], we can prove the following results.

Theorem 1 (Existence of a mutually incompatible set of composite choices). *Given a finite set K of finite composite choices, there exists a finite set K' of mutually incompatible finite composite choices such that $\omega_K = \omega_{K'}$.*

Proof. Given a finite set of finite composite choices K, there are two possibilities to form a new set K' of composite choices so that K and K' describe the same set of possible worlds:

1. **removing dominated elements**: if $\kappa_1, \kappa_2 \in K$ and $\kappa_1 \subset \kappa_2$, let $K' = K \setminus \{\kappa_2\}$.
2. **splitting elements**: if $\kappa_1, \kappa_2 \in K$ and $\kappa_1 \cup \kappa_2$ is consistent (and neither is a superset of the other), there is an $(F,\theta,k) \in \kappa_1 \setminus \kappa_2$. We replace κ_2 by the split of κ_2 on $F\theta$. Let $K' = K \setminus \{\kappa_2\} \cup S_{F\theta,\kappa_2}$.

In both cases $\omega_K = \omega_{K'}$. If we repeat this two operations until neither is applicable we obtain a splitting algorithm (see Figure 1) that terminates because K is a finite set of finite composite choices. The resulting set K' is mutually incompatible and is equivalent to the original set. $\qquad\square$

Fig. 1. Splitting Algorithm.

```
 1: procedure SPLIT(K)
 2:     Input: set of composite choices K
 3:     Output: mutually incompatible set of composite choices equivalent to K
 4:     loop
 5:         if ∃κ₁, κ₂ ∈ K and κ₁ ⊂ κ₂ then
 6:             K := K \ {κ₂}
 7:         else
 8:             if ∃κ₁, κ₂ ∈ K, such that κ₁ ∪ κ₂ is inconsistent then
 9:                 choose (F, θ, k) ∈ κ₁ \ κ₂
10:                 K := K \ {κ₂} ∪ S_{Fθ,κ₂}
11:             else
12:                 exit and return K
13:             end if
14:         end if
15:     end loop
16: end procedure
```

Theorem 2 (Equivalence of the probability of two equivalent mutually incompatible finite set of finite composite choices). *If K_1 and K_2 are both mutually incompatible finite sets of finite composite choices such that they are equivalent then $P(K_1) = P(K_2)$.*

Proof. The theorem is the same as Lemma A.8 in [29]. We report here the proof for the sake of clarity.

Consider the set D of all instantiated formulas $F\theta$ that appear in an atomic choice in either K_1 and K_2. This set is finite. Each composite choice in K_1 and K_2 has atomic choices for a subset of D. For both K_1 and K_2, we repeatedly replace each composite choice κ of K_1 and K_2 with its split K' on an $F_i\theta_j$ from D that does not appear in κ. This procedure does not change the total probability as the probabilities of $(F_i, \theta_j, 0)$ and $(F_i, \theta_j, 1)$ sum to 1.

At the end of this procedure the two sets of composite choices will be identical. In fact, any difference can be extended into a possible world belonging to ω_{K_1} but not to ω_{K_2} or vice versa. $\square$

We can thus define a unique probability measure $\mu : \Omega_T \to [0, 1]$ where Ω_T is defined as the set of sets of worlds identified by finite sets of finite composite choices: $\Omega_T = \{\omega_K | K$ is a finite set of finite composite choices$\}$. It is easy to see that Ω_T is an algebra over $\mathcal{W}_T$.

Then μ is defined by $\mu(\omega_K) = P(K')$ where K' is a finite mutually incompatible set of finite composite choices equivalent to K. $\langle \mathcal{W}_T, \Omega_T, \mu \rangle$ is a probability space according to Kolmogorov's definition.

The probability of a query Q is given by $P(Q) = \mu(\{w | w \in \mathcal{W}_T \wedge w \models Q\})$. If Q has a finite set K of finite explanations such that K is covering then $\{w | w \in \mathcal{W}_T \wedge w \models Q\} = \omega_K \in \Omega_T$ and $P(Q)$ is well-defined.

Example 1. Let us consider the following ontology, inspired by the `people+pets` ontology proposed in [27]:

$$\exists hasAnimal.Pet \sqsubseteq PetOwner$$
$$(kevin, fluffy) : hasAnimal$$
$$(kevin, tom) : hasAnimal$$
$$0.4 ::_e fluffy : Cat$$
$$0.3 ::_e tom : Cat$$
$$0.6 ::_e Cat \sqsubseteq Pet$$

The predicate logic formulas (without external quantifiers) equivalent to the probabilistic axioms are

$$F_1 = Cat(fluffy)$$
$$F_2 = Cat(tom)$$
$$F_3 = Cat(x) \to Pet(x)$$

A covering set of explanations for the query axiom $Q = kevin : PetOwner$ is $K = \{\kappa_1, \kappa_2\}$ where $\kappa_1 = \{(F_1, \emptyset, 1), (F_3, \emptyset, 1)\}$ and $\kappa_2 = \{(F_2, \emptyset, 1), (F_3, \emptyset, 1)\}$.

An equivalent mutually exclusive set K' of explanations can be obtained by applying the splitting algorithm. In this case $K' = \{\kappa'_1, \kappa'_2\}$ where $\kappa'_1 = \{(F_1, \emptyset, 1), (F_3, \emptyset, 1), (F_2, \emptyset, 0)\}$ and $\kappa'_2 = \{(F_2, \emptyset, 1), (F_3, \emptyset, 1)\}$. So $P(Q) = 0.4 \cdot 0.6 \cdot 0.7 + 0.3 \cdot 0.6 = 0.348$.

Example 2. If the axiom $0.6 ::_e Cat \sqsubseteq Pet$ in Example 1 is replaced by $0.6 ::_s Cat \sqsubseteq Pet$ then the query would have the explanations $K = \{\kappa_1, \kappa_2\}$ where $\kappa_1 = \{(F_1, \emptyset, 1), (F_3, \{x/fluffy\}, 1)\}$ and $\kappa_2 = \{(F_2, \emptyset, 1), (F_3, \{x/tom\}, 1)\}$.

An equivalent mutually exclusive set K' of explanations obtained by applying the splitting algorithm is $K' = \{\kappa'_1, \kappa'_2, \kappa'_3\}$ with $\kappa'_1 = \{(F_1, \emptyset, 1), (F_3, \{x/fluffy\}, 1), (F_2, \emptyset, 0)\}$, $\kappa'_2 = \{(F_1, \emptyset, 1), (F_3, \{x/fluffy\}, 1), (F_2, \emptyset, 1), (F_3, \{x/tom\}, 0)\}$ and $\kappa'_3 = \{(F_2, \emptyset, 1), (F_3, \{x/tom\}, 1)\}$. So $P(Q) = 0.4 \cdot 0.6 \cdot 0.7 + 0.4 \cdot 0.6 \cdot 0.3 \cdot 0.4 + 0.3 \cdot 0.6 = 0.3768$.

Example 3. Let us consider a slightly different ontology:

$$0.5 ::_s \exists hasAnimal.Pet \sqsubseteq PetOwner$$
$$(kevin, fluffy) : hasAnimal$$
$$(kevin, tom) : hasAnimal$$
$$fluffy : Cat$$
$$tom : Cat$$
$$0.6 ::_s Cat \sqsubseteq Pet$$

The predicate logic formulas without external quantifiers equivalent to the probabilistic axioms are

$$F_1 = \exists y.hasAnimal(x, y) \wedge Pet(y) \rightarrow PetOwner(x)$$
$$F_2 = Cat(x) \rightarrow Pet(x)$$

A covering set of explanations for the query axiom $Q = kevin : PetOwner$ is $K = \{\kappa_1, \kappa_2\}$ where $\kappa_1 = \{(F_1, \{x/kevin\}, 1), (F_2, \{x/fluffy\}, 1)\}$ and $\kappa_2 = \{(F_1, \{x/kevin\}, 1), (F_2, \{x/tom\}, 1)\}$.

An equivalent mutually exclusive set K' of explanations obtained by applying the splitting algorithm is $K' = \{\kappa'_1, \kappa'_2\}$ where $\kappa'_1 = \{(F_1, \{x/kevin\}, 1), (F_2, \{x/fluffy\}, 1), (F_2, \{x/tom\}, 0)\}$ and $\kappa'_2 = \{(F_1, \{x/kevin\}, 1), (F_2, \{x/tom\}, 1)\}$. So $P(Q) = 0.5 \cdot 0.6 \cdot 0.4 + 0.5 \cdot 0.6 = 0.42$.

Example 4. Let us consider the ontology:

$$0.7 ::_s Schoolchild \sqsubseteq European$$
$$0.4 ::_s Schoolchild \sqsubseteq OnlyChild$$
$$0.6 ::_s European \sqsubseteq GoodInMath$$
$$0.5 ::_s OnlyChild \sqsubseteq GoodInMath$$

The predicate logic formulas without the external quantifiers equivalent to the probabilistic axioms are:

$$F_1 = Schoolchild(x) \rightarrow European(x)$$

$$F_2 = Schoolchild(x) \rightarrow OnlyChild(x)$$
$$F_3 = European(x) \rightarrow GoodInMath(x)$$
$$F_4 = OnlyChild(x) \rightarrow GoodInMath(x)$$

A covering set of explanations for the query axiom $Q = Schoolchild \sqsubseteq GoodInMath$ is $K = \{\kappa_1, \kappa_2\}$ where $\kappa_1 = \{(F_1, \{x/i\}, 1), (F_3, \{x/i\}, 1)\}$ and $\kappa_2 = \{(F_2, \{x/i\}, 1), (F_4, \{x/i\}, 1)\}$, where i is an anonymous member of $\Delta^{\mathcal{I}}$. After splitting we get $K' = \{\kappa_1', \kappa_2', \kappa_3'\}$ where $\kappa_1' = \{(F_1, \{x/i\}, 1), (F_3, \{x/i\}, 1)\}$, $\kappa_2' = \{(F_1, \{x/i\}, 0), (F_2, \{x/i\}, 1), (F_4, \{x/i\}, 1)\}$ and $\kappa_3' = \{(F_1, \{x/i\}, 1), (F_3, \{x/i\}, 0), (F_2, \{x/i\}, 1), (F_4, \{x/i\}, 1)\}$. So $P(Q) = 0.7 \cdot 0.6 + 0.3 \cdot 0.4 \cdot 0.5 + 0.7 \cdot 0.4 \cdot 0.4 \cdot 0.5 = 0.536$.

4   Reasoning under the DISPONTE semantics

The BUNDLE algorithm presented in [3] computes the probability of queries from a probabilistic ontology that follows the DISPONTE semantics with only epistemic probabilities. BUNDLE uses an underlying DL reasoner that is able to return explanations for queries such as Pellet [35]. The *explain* function of Pellet is used for this purpose [15, 17, 12, 16]. BUNDLE makes the explanations mutually incompatible by using Binary Decision Diagrams (BDD)[4] because this approach has been demonstrated to be faster than using the splitting algorithm [18, 31]. BDDs are used to compute the probability using the dynamic programming algorithm of [8].

In order to also deal with statistical probabilities, Pellet needs to be modified so that it records, besides the axioms that have been used to answer the query, also the individuals to which they are applied. We are currently working on modifying the tableau expansion rules of Pellet so that the individuals on which they operate are recorded together with the axioms. The first results on this activity are reported in [32], which presents a performance evaluation of inference over the real probabilistic ontology for breast cancer risk assesment both for BUNDLE and PRONTO.

5   Related Work

[13] proposed an extension of the description logic $\mathcal{ALC}$ that is able to express statistical information on the terminological knowledge such as partial concept overlapping. Similarly, [20] presented a probabilistic description logic based on Bayesian networks that deals with statistical terminological knowledge. [13, 20] do not allow probabilistic assertional knowledge about concept and role instances. [14] allows assertional knowledge about concept and role instances together with statistical terminological knowledge and combines the resulting probability distributions using cross-entropy minimization but does not allow epistemic statements.

[9] proposed a probabilistic extension of OWL that admits a translation into Bayesian networks. The semantics that is proposed assigns a probability distribution $P(i)$ over individuals, i.e. $\sum_i P(i) = 1$, and assigns a probability to a class C as $P(C) = \sum_{i \in C} P(i)$, while we assign a probability measure to sets of worlds. PR-OWL [6, 5] is an upper ontology that provides a framework for building probabilistic ontologies. It allows to use the first-order probabilistic logic MEBN [21] for representing uncertainty in ontologies. The use of a full fledged first-order probabilistic logic distinguishes this work from ours, where we tried to provide a minimal extension to description logics.

A different approach to the combination of description logic with probability is taken by [10, 22, 23] where the authors use probabilistic lexicographic entailment from probabilistic default reasoning. The logics proposed in these papers allow both terminological probabilistic knowledge as well as assertional probabilistic knowledge about instances of concepts and roles. PRONTO [19] is one of the systems that allows to perform inference in this semantics. These works are based on Nilsson's probabilistic logic [25], where a probabilistic interpretation Pr defines a probability distribution over the set of interpretations Int. The probability of a logical formula F according to Pr, denoted $Pr(F)$, is the sum of all $Pr(I)$ such that $I \in Int$ and $I \models F$. A probabilistic knowledge base K is a set of probabilistic formulas of the form $F \geq p$. A probabilistic interpretation Pr satisfies $F \geq p$ iff $Pr(F) \geq p$. Pr satisfies K, or Pr is a model of K, iff Pr satisfies all $F \geq p \in K$. $Pr(F) \geq p$ is a tight logical consequence of K iff p is the infimum of $Pr(F)$ subject to all models Pr of K.

Nilsson's probabilistic logic differs from the distribution semantics: while the first computes the lowest p such that $Pr(F) \geq p$ holds for all Pr, the latter computes p such that $P(F) = p$. Thus Nilsson's logic allows weaker conclusions: consider a probabilistic ontology composed of the axioms $0.4 ::_e a : C.$ and $0.5 ::_e b : C.$ and a probabilistic knowledge base composed of $C(a) \geq 0.4$ and $C(b) \geq 0.5$. The distribution semantics allows to say that $P(a : C \vee b : C) = 0.7$, while with Nilsson's logic the lowest p such that $Pr(C(a) \vee C(b)) \geq p$ holds is 0.5. This is due to the fact that in the distribution semantics the probabilistic axioms are considered as independent, which allows to make stronger conclusions.

Other approaches, such as [7, 11], combine a liteweight ontology language, *DL-Lite* and Datalog+/- respectively, with graphical models, Bayesian networks and Markov networks respectively. In both cases, an ontology is composed of a set of annotated axioms and a graphical model and the annotations are sets of assignments of random variables from the graphical model. The semantics is assigned by considering the possible worlds of the graphical model and by stating that an axiom holds in a possible world if the assignments in its annotation hold. The probability of a conclusion is then the sum of the probabilities of the possible worlds where the conclusion holds. Our approach provides a tighter integration of probability in ontologies as we do not rely on an additional graphical model.

In summary, we allow to extend Description Logics languages with more complete treatment of probabilistic knowledge that includes subjective and statistical statement.

6 Conclusions

We extended the DISPONTE semantics by allowing the representation of statistical information on concepts and populations. In this way both epistemic and partial overlap information can be expressed in a seamless way.

References

1. Baader, F., Calvanese, D., McGuinness, D.L., Nardi, D., Patel-Schneider, P.F. (eds.): The Description Logic Handbook: Theory, Implementation, and Applications. Cambridge University Press (2003)
2. Baader, F., Horrocks, I., Sattler, U.: Description logics. In: Handbook of knowledge representation, chap. 3, pp. 135–179. Elsevier (2008)
3. Bellodi, E., Lamma, E., Riguzzi, F., Albani, S.: A distribution semantics for probabilistic ontologies. In: International Workshop on Uncertainty Reasoning for the Semantic Web. No. 778 in CEUR Workshop Proceedings, Sun SITE Central Europe (2011)
4. Bryant, R.E.: Graph-based algorithms for boolean function manipulation. IEEE Trans. on Comput. 35(8), 677–691 (1986)
5. Carvalho, R.N., Laskey, K.B., Costa, P.C.G.: PR-OWL 2.0 - bridging the gap to OWL semantics. In: International Workshop on Uncertainty Reasoning for the Semantic Web (2010)
6. Costa, P.C.G., Laskey, K.B., Laskey, K.J.: PR-OWL: A Bayesian ontology language for the semantic web. In: International Workshop on Uncertainty Reasoning for the Semantic Web. LNCS, vol. 5327, pp. 88–107. Springer (2008)
7. d'Amato, C., Fanizzi, N., Lukasiewicz, T.: Tractable reasoning with Bayesian description logics. In: International Conference on Scalable Uncertainty Management. LNCS, vol. 5291, pp. 146–159. Springer (2008)
8. De Raedt, L., Kimmig, A., Toivonen, H.: ProbLog: A probabilistic Prolog and its application in link discovery. In: International Joint Conference on Artificial Intelligence. pp. 2462–2467 (2007)
9. Ding, Z., Peng, Y.: A probabilistic extension to ontology language OWL. In: Hawaii International Conference On System Sciences. IEEE (2004)
10. Giugno, R., Lukasiewicz, T.: P-SHOQ(D): A probabilistic extension of SHOQ(D) for probabilistic ontologies in the semantic web. In: European Conference on Logics in Artificial Intelligence. LNCS, vol. 2424, pp. 86–97. Springer (2002)
11. Gottlob, G., Lukasiewicz, T., Simari, G.I.: Conjunctive query answering in probabilistic Datalog+/- ontologies. In: International Conference on Web Reasoning and Rule Systems. LNCS, vol. 6902, pp. 77–92. Springer (2011)
12. Halaschek-Wiener, C., Kalyanpur, A., Parsia, B.: Extending tableau tracing for ABox updates. Tech. rep., University of Maryland (2006)
13. Heinsohn, J.: Probabilistic description logics. In: Conference on Uncertainty in Artificial Intelligence. pp. 311–318. Morgan Kaufmann (1994)
14. Jaeger, M.: Probabilistic reasoning in terminological logics. In: International Conference on Principles of Knowledge Representation and Reasoning. pp. 305–316 (1994)
15. Kalyanpur, A., Parsia, B., Cuenca-Grau, B., Sirin, E.: Tableaux tracing in SHOIN. Tech. Rep. 2005-66, University of Maryland (2005)

16. Kalyanpur, A., Parsia, B., Horridge, M., Sirin, E.: Finding all justifications of OWL DL entailments. In: International Semantic Web Conference. LNCS, vol. 4825, pp. 267–280. Springer (2007)
17. Kalyanpur, A., Parsia, B., Sirin, E., Hendler, J.A.: Debugging unsatisfiable classes in OWL ontologies. J. Web Sem. 3(4), 268–293 (2005)
18. Kimmig, A., Demoen, B., Raedt, L.D., Costa, V.S., Rocha, R.: On the implementation of the probabilistic logic programming language problog. Theor. Prac. Log. Prog. 11(2-3), 235–262 (2011)
19. Klinov, P.: Pronto: A non-monotonic probabilistic description logic reasoner. In: European Semantic Web Conference. LNCS, vol. 5021, pp. 822–826. Springer (2008)
20. Koller, D., Levy, A.Y., Pfeffer, A.: P-CLASSIC: A tractable probablistic description logic. In: National Conference on Artificial Intelligence. pp. 390–397 (1997)
21. Laskey, K.B., Costa, P.C.G.: Of starships and Klingons: Bayesian logic for the 23rd century. In: Conference in Uncertainty in Artificial Intelligence. pp. 346–353. AUAI Press (2005)
22. Lukasiewicz, T.: Probabilistic default reasoning with conditional constraints. Ann. Math. Artif. Int. 34(1-3), 35–88 (2002)
23. Lukasiewicz, T.: Expressive probabilistic description logics. Artif. Int. 172(6-7), 852–883 (2008)
24. Lukasiewicz, T., Straccia, U.: Managing uncertainty and vagueness in description logics for the semantic web. J. Web Sem. 6(4), 291–308 (2008)
25. Nilsson, N.J.: Probabilistic logic. Artif. Intell. 28(1), 71–87 (1986)
26. Obrst, L., McCandless, D., Stoutenburg, S., Fox, K., Nichols, D., Prausa, M., Sward, R.: Evolving use of distributed semantics to achieve net-centricity. In: AAAI Fall Symposium (2007)
27. Patel-Schneider, P, F., Horrocks, I., Bechhofer, S.: Tutorial on OWL (2003)
28. Poole, D.: Abducing through negation as failure: stable models within the independent choice logic. J. Log. Program. 44(1-3), 5–35 (2000)
29. Poole, D.: Probabilistic horn abduction and Bayesian networks. Artif. Intell. 64(1) (1993)
30. Predoiu, L., Stuckenschmidt, H.: Probabilistic extensions of semantic web languages - a survey. In: The Semantic Web for Knowledge and Data Management: Technologies and Practices. Idea Group Inc (2008)
31. Riguzzi, F.: Extended semantics and inference for the Independent Choice Logic. Log. J. IGPL 17(6), 589–629 (2009)
32. Riguzzi, F., Lamma, E., Bellodi, E., Zese, R.: Semantics and inference for probabilistic ontologies. In: Popularize Artificial Intelligence Workshop. pp. 41–46. No. 860 in CEUR Workshop Proceedings, Sun SITE Central Europe (2012)
33. Sato, T.: A statistical learning method for logic programs with distribution semantics. In: International Conference on Logic Programming. pp. 715–729. MIT Press (1995)
34. Sattler, U., Calvanese, D., Molitor, R.: Relationships with other formalisms. In: Description Logic Handbook. pp. 137–177 (2003)
35. Sirin, E., Parsia, B., Cuenca-Grau, B., Kalyanpur, A., Katz, Y.: Pellet: A practical OWL-DL reasoner. J. Web Sem. 5(2), 51–53 (2007)
36. URW3-XG: Uncertainty reasoning for the World Wide Web, final report (2005)