



**TECHNISCHE  
UNIVERSITÄT  
DRESDEN**

**Technische Universität Dresden  
Institute for Theoretical Computer Science  
Chair for Automata Theory**

# **LTCS–Report**

## **Learning General Concept Inclusions in Probabilistic Description Logics**

Francesco Kriegel

LTCS-Report 15-14

Postal Address:  
Lehrstuhl für Automatentheorie  
Institut für Theoretische Informatik  
TU Dresden  
01062 Dresden

<http://lat.inf.tu-dresden.de>

Visiting Address:  
Nöthnitzer Str. 46  
Dresden

# Learning General Concept Inclusions in Probabilistic Description Logics

Francesco Kriegel  
Institute of Theoretical Computer Science  
Technische Universität Dresden, Germany  
francesco@tcs.inf.tu-dresden.de

June 6, 2015

## Abstract

Probabilistic interpretations consist of a set of interpretations with a shared domain and a measure assigning a probability to each interpretation. Such structures can be obtained as results of repeated experiments, e.g., in biology, psychology, medicine, etc. A translation between probabilistic and crisp description logics is introduced and then utilised to reduce the construction of a base of general concept inclusions of a probabilistic interpretation to the crisp case for which a method for the axiomatisation of a base of GCIs is well-known.

**Keywords:** probabilistic description logic, terminological axioms, automatic learning, base of gcis, general concept inclusions, probabilistic interpretation

## Contents

|    |   |    |
|----|---|----|
| 1  | Introduction  | 3  |
| 2  | The Description Logics $\mathbb{P}_{01}\mathcal{FL}\mathcal{E}^\perp$ and $\mathbb{P}_{\geq}\mathcal{FL}\mathcal{E}^\perp Q^{\geq}$ | 3  |
| 3  | Translation between $\mathbb{P}_{01}\mathcal{FL}\mathcal{E}^\perp$ and $\mathcal{FL}\mathcal{E}^\perp$                              | 4  |
| 4  | Construction of a Base of GCIs in $\mathbb{P}_{01}\mathcal{FL}\mathcal{E}^\perp$  | 6  |
| 5  | Translation between $\mathbb{P}_{\geq}\mathcal{FL}\mathcal{E}^\perp Q^{\geq}$ and $\mathcal{FL}\mathcal{E}^\perp Q^{\geq}$          | 7  |
| 6  | Construction of a Base of GCIs in $\mathbb{P}_{\geq}\mathcal{FL}\mathcal{E}^\perp Q^{\geq}$   | 9  |
| 7  | Probabilistic Model-Based Most-Specific Concept Descriptions  | 10 |
| 8  | Choice of Semantics   | 11 |
| 9  | Complexity of Base Construction   | 12 |
| 10 | Conclusion  | 13 |
| A  | Unused Results  | 15 |

# 1 Introduction

This document proposes a method for axiomatising a base of general concept inclusions for probabilistic interpretations. It is obtained by means of a translation between probabilistic description logics and crisp description logics, and well-known results for the construction of a base of GCIs for interpretations in crisp description logics. However, this approach does not add further complexity as the translations may be computed in polynomial time. There are several approaches for an integration of probabilities into description logics. This document follows the basic definitions of Lutz and Schröder in [12] where a probabilistic interpretation is defined as a family of standard interpretations over the same domain such that each interpretation has a specific probability. These structures naturally arise from experiments, e.g., in biology, psychology, or medicine, respectively, that are repeated several times. If for example the experiments may produce results with errors, or some effects may not always be observed, then repetition is advantageous. For all such sequences of interpretations over a shared domain the probability measure can easily be defined as a uniform discrete probability measure over all observed interpretations. We call such probabilistic interpretations (quasi-)uniform.

At first we introduce the probabilistic description logics  $\mathbb{P}_{01}\mathcal{FL}\mathcal{E}^\perp$  and  $\mathbb{P}_{\geq}\mathcal{FL}\mathcal{E}^\perp Q^{\geq}$ . Then we present a translation between  $\mathbb{P}_{01}\mathcal{FL}\mathcal{E}^\perp$  and  $\mathcal{FL}\mathcal{E}^\perp$  that satisfies certain consistency properties w.r.t. the underlying probabilistic interpretation. By means of the translation we utilise previous results for the construction of a base of general concept inclusions. In particular, the notion of a (canonical) base of GCIs is used here that has been found by Baader and Distel in [3, 2, 8] for the description logic  $\mathcal{EL}^\perp$  w.r.t. greatest fixpoint semantics. Furthermore, it has been adapted by Borchmann, Distel, and Kriegel in [6] for  $\mathcal{EL}^\perp$  w.r.t. role-depth bounds, and has been extended towards the more expressive description logic  $\mathcal{AL}\mathcal{E}Q^{\geq}\mathcal{N}^{\leq}(\text{Self})$  in [11] (hence, may also be applied to the smaller description logics  $\mathcal{FL}\mathcal{E}^\perp$  and  $\mathcal{FL}\mathcal{E}^\perp Q^{\geq}$ ). The construction of a base of probabilistic GCIs is also generalised towards the more expressive description logic  $\mathbb{P}_{\geq}\mathcal{FL}\mathcal{E}^\perp Q^{\geq}$  but only in the case of quasi-uniform probabilistic interpretations.

Most-specific generalisations in probabilistic description logics have been subject of previous research. In [15, 13, 14] Peñaloza and Turhan investigated methods for the construction of most-specific concept description (w.r.t. a knowledge base) and least common subsumers in probabilistic  $\mathcal{EL}$ . Later, in [10, 9] Ecke, Peñaloza, and Turhan, extended their results towards nominals and complex role inclusion axioms. This document also provides a method for the construction of probabilistic model-based most-specific concept descriptions (w.r.t. probabilistic interpretations).

## 2 The Description Logics $\mathbb{P}_{01}\mathcal{FL}\mathcal{E}^\perp$ and $\mathbb{P}_{\geq}\mathcal{FL}\mathcal{E}^\perp Q^{\geq}$

At first, we introduce the probabilistic description logic  $\mathbb{P}\mathcal{FL}\mathcal{E}^\perp Q^{\geq}$  that extends the well-known description logic  $\mathcal{FL}\mathcal{E}$ . A role is either a role name  $r \in N_R$  or of the form  $\mathbb{P}_{\bowtie p} r$  for a  $\bowtie \in \{<, \leq, =, \geq, >\}$ , a role name  $r \in N_R$ , and a probability threshold  $p \in (0, 1]$ . Furthermore, concept descriptions may be built according to the following syntax rule where  $s$  denotes a role,  $A \in N_C$  a concept name,  $n \geq 2$  an integer,  $\bowtie \in \{<, \leq, =, \geq, >\}$ , and  $p \in [0, 1]$ :

$$C ::= \perp \mid \top \mid A \mid C \sqcap C \mid \exists s. C \mid \forall s. C \mid \geq n. s. C \mid \mathbb{P}_{\bowtie p} C$$

The description logic  $\mathbb{P}_{01}\mathcal{FL}\mathcal{E}^\perp$  does not allow for qualified  $\geq$ -restrictions  $\geq n.s.C$ , and only allows for probabilistic concept and role constructors  $\mathbb{P}_{>0}$  and  $\mathbb{P}_{=0}$ . Furthermore, the description logic  $\mathbb{P}_{\geq}\mathcal{FL}\mathcal{E}^\perp\mathcal{Q}^\geq$  only allows for probabilistic constructors  $\mathbb{P}_{\geq p}$ .

A detailed overview on probabilistic extensions of the description logics  $\mathcal{ALC}$  and  $\mathcal{EL}$  and several complexity results for reasoning in probabilistic description logics have been given by Lutz and Schröder in [12].

A *probability measure* on a countable set  $W$  is a mapping  $\mathbb{P}: 2^W \rightarrow [0, 1]$  such that  $\mathbb{P}(\emptyset) = 0$  and  $\mathbb{P}(W) = 1$  hold, and furthermore for all countable pairwise disjoint sequences  $(U_n)_{n \in \mathbb{N}}$  of subsets  $U_n \subseteq W$  it is true that  $\mathbb{P}(\bigsqcup_{n \in \mathbb{N}} U_n) = \sum_{n \in \mathbb{N}} \mathbb{P}(U_n)$ , i.e.,  $\mathbb{P}$  is  $\sigma$ -additive. For a subset  $U \subseteq W$  the value  $\mathbb{P}(U)$  is the *probability* of  $U$  w.r.t.  $\mathbb{P}$ .

Let  $(N_C, N_R)$  be a signature. A *probabilistic interpretation* over  $(N_C, N_R)$  is a tuple  $\mathcal{I} = (\Delta^\mathcal{I}, W, (\cdot^{\mathcal{I}_w})_{w \in W}, \mathbb{P})$  that consists of a set  $\Delta^\mathcal{I}$ , called *domain*, a countable set  $W$  of *worlds*, an *extension function*  $\cdot^{\mathcal{I}_w}$  for each world  $w \in W$ , and a probability measure  $\mathbb{P}$  on  $W$ . For each world  $w \in W$  the tuple  $(\Delta^\mathcal{I}, \cdot^{\mathcal{I}_w})$  is an interpretation over  $(N_C, N_R)$  that may be extended to all  $\mathcal{FL}\mathcal{E}^\perp\mathcal{Q}^\geq$ -concept descriptions in the canonical way. Furthermore, for the probabilistic constructors  $\mathbb{P}_{\bowtie p}$  with  $\bowtie \in \{<, \leq, =, \geq, >\}$  and  $p \in [0, 1]$  their extensions are defined as follows:

$$\begin{aligned} (\mathbb{P}_{\bowtie p} C)^{\mathcal{I}_w} &:= \{d \in \Delta^\mathcal{I} \mid \mathbb{P}\{w \in W \mid d \in C^{\mathcal{I}_w}\} \bowtie p\}, \\ (\mathbb{P}_{\bowtie p} r)^{\mathcal{I}_w} &:= \{(d, e) \in \Delta^\mathcal{I} \times \Delta^\mathcal{I} \mid \mathbb{P}\{w \in W \mid (d, e) \in r^{\mathcal{I}_w}\} \bowtie p\}. \end{aligned}$$

Note that the extensions are independent of the world  $w$ , i.e., they coincide in all worlds of the probabilistic interpretation. An individual  $d \in \Delta^\mathcal{I}$  is in the extension of  $\mathbb{P}_{>0} C$  if and only if  $d$  is *possibly* in the extension of  $C$ , and is in the extension of  $\mathbb{P}_{=1} C$  iff  $d$  is *almost surely* in the extension of  $C$ .

A world  $w \in W$  is called *possible* if its probability is not 0, i.e., if  $\mathbb{P}\{w\} > 0$  holds; otherwise we call  $w$  *impossible*. For a probabilistic interpretation we denote the set of all possible worlds by  $W_\varepsilon$ , and the set of all impossible worlds by  $W_0$ . Of course,  $W_\varepsilon \uplus W_0$  is a partition of  $W$ , and  $\mathbb{P}(W_\varepsilon) = 1$  and  $\mathbb{P}(W_0) = 0$  hold.

A *general concept inclusion* (GCI) is of the form  $C \sqsubseteq D$  where  $C$  and  $D$  are concept descriptions. It holds in a probabilistic interpretation  $\mathcal{I}$  if and only if  $C^{\mathcal{I}_w} \subseteq D^{\mathcal{I}_w}$  is satisfied for all worlds  $w \in W$ , and we shall denote this by  $\mathcal{I} \models C \sqsubseteq D$ .

Let  $\mathcal{I}$  be a probabilistic interpretation. Then a TBox  $\mathcal{B}$  is called *base of GCIs* for  $\mathcal{I}$  if  $\mathcal{I}$  models all GCIs in  $\mathcal{B}$ , i.e.,  $\mathcal{B}$  is *sound*, and whenever a GCI holds in  $\mathcal{I}$  then it follows from  $\mathcal{B}$ , i.e.,  $\mathcal{B}$  is *complete*.

### 3 Translation between $\mathbb{P}_{01}\mathcal{FL}\mathcal{E}^\perp$ and $\mathcal{FL}\mathcal{E}^\perp$

It is readily verified that  $d \in (\mathbb{P}_{>0} C)^{\mathcal{I}_w}$  holds if and only if there is a possible world  $v \in W_\varepsilon$  such that  $d \in C^{\mathcal{I}_v}$  hold. Analogously,  $d \in (\mathbb{P}_{=1} C)^{\mathcal{I}_w}$  is equivalent to the statement that  $d \in C^{\mathcal{I}_v}$  is true for all possible worlds  $v \in W_\varepsilon$ . Similar statements hold for the probabilistic role constructors  $\mathbb{P}_{>0} r$  and  $\mathbb{P}_{=1} r$ . Hence, it is possible to translate  $\mathbb{P}_{01}\mathcal{FL}\mathcal{E}^\perp$ -concept descriptions into  $\mathcal{FL}\mathcal{E}^\perp$ -concept descriptions and vice versa.

For this purpose a new role name  $\omega_{\mathbb{P}}$ , and role names  $r_{>0}, r_{=1}$  for each existing role name  $r \in N_R$ , are introduced into the signature, and we shall denote the extended signature by:

$$(N_C, N_R)_{01}^{\mathbb{P}} := (N_C, N_R \uplus \{ \omega_{\mathbb{P}} \} \uplus \{ r_{>0}, r_{=1} \mid r \in N_R \}).$$

Then the translation function  $\tau: \mathbb{P}_{01}\mathcal{FL}\mathcal{E}^{\perp}(N_C, N_R) \rightarrow \mathcal{FL}\mathcal{E}^{\perp}(N_C, N_R)_{01}^{\mathbb{P}}$  and its inverse  $\tau^{-1}$  are inductively defined as follows:

$$\begin{array}{ll} \tau(r) := r & \tau^{-1}(r) := r \\ \tau(\mathbb{P}_{>0} r) := r_{>0} & \tau^{-1}(r_{>0}) := \mathbb{P}_{>0} r \\ \tau(\mathbb{P}_{=1} r) := r_{=1} & \tau^{-1}(r_{=1}) := \mathbb{P}_{=1} r \\ \tau(A) := A & \tau^{-1}(A) := A \\ \tau(C \sqcap D) := \tau(C) \sqcap \tau(D) & \tau^{-1}(C \sqcap D) := \tau^{-1}(C) \sqcap \tau^{-1}(D) \\ \tau(\exists s. C) := \exists \tau(s). \tau(C) & \tau^{-1}(\exists s. C) := \exists \tau^{-1}(s). \tau^{-1}(C) \\ \tau(\forall s. C) := \forall \tau(s). \tau(C) & \tau^{-1}(\forall s. C) := \forall \tau^{-1}(s). \tau^{-1}(C) \\ \tau(\mathbb{P}_{>0} C) := \exists \omega_{\mathbb{P}}. \tau(C) & \tau^{-1}(\exists \omega_{\mathbb{P}}. C) := \mathbb{P}_{>0} \tau^{-1}(C) \\ \tau(\mathbb{P}_{=1} C) := \forall \omega_{\mathbb{P}}. \tau(C) & \tau^{-1}(\forall \omega_{\mathbb{P}}. C) := \mathbb{P}_{=1} \tau^{-1}(C) \end{array}$$

For each probabilistic interpretation  $\mathcal{I} = (\Delta^{\mathcal{I}}, W, (\cdot^{\mathcal{I}_w})_{w \in W}, \mathbb{P})$  over  $(N_C, N_R)$  we define the interpretation  $\mathcal{I}^{\times} := (\Delta^{\mathcal{I}} \times W, \cdot^{\mathcal{I}^{\times}})$  over  $(N_C, N_R)_{01}^{\mathbb{P}}$  whose extension function is given as follows:

$$\begin{array}{ll} A^{\mathcal{I}^{\times}} := \{ (d, w) \mid d \in A^{\mathcal{I}_w} \} & (A \in N_C) \\ r^{\mathcal{I}^{\times}} := \{ ((d, w), (e, w)) \mid (d, e) \in r^{\mathcal{I}_w} \} & (r \in N_R) \\ \omega_{\mathbb{P}}^{\mathcal{I}^{\times}} := \{ ((d, v), (d, w)) \mid \mathbb{P} \{ w \} > 0 \} \\ r_{>0}^{\mathcal{I}^{\times}} := \{ ((d, w), (e, w)) \mid (d, e) \in (\mathbb{P}_{>0} r)^{\mathcal{I}_w} \} \\ r_{=1}^{\mathcal{I}^{\times}} := \{ ((d, w), (e, w)) \mid (d, e) \in (\mathbb{P}_{=1} r)^{\mathcal{I}_w} \} \end{array}$$

The special role  $\omega_{\mathbb{P}}$  connects each individual  $d$  in an arbitrary world to itself in a possible world. Then the following lemma shows the connection between the given translation functions.

**Lemma 1.** *Let  $\mathcal{I} = (\Delta^{\mathcal{I}}, W, (\cdot^{\mathcal{I}_w})_{w \in W}, \mathbb{P})$  be a probabilistic interpretation,  $d \in \Delta^{\mathcal{I}}$  an individual,  $w \in W$  a world, and  $C$  a  $\mathbb{P}_{01}\mathcal{FL}\mathcal{E}^{\perp}$ -concept description. Then the following equivalence holds:*

$$d \in C^{\mathcal{I}_w} \text{ if and only if } (d, w) \in \tau(C)^{\mathcal{I}^{\times}}.$$

*Proof.* by structural induction on  $C$ .

induction base:  $C = A$

Of course, it holds that  $\tau(A) = A$ . Thus, the equivalence follows by definition of  $\mathcal{I}^{\times}$ .

inductive step:  $C = D \sqcap E$

$$\begin{aligned} d \in (D \sqcap E)^{\mathcal{I}_w} &\Leftrightarrow d \in D^{\mathcal{I}_w} \text{ and } d \in E^{\mathcal{I}_w} \\ &\stackrel{\text{IH}}{\Leftrightarrow} (d, w) \in \tau(D)^{\mathcal{I}^{\times}} \text{ and } (d, w) \in \tau(E)^{\mathcal{I}^{\times}} \\ &\Leftrightarrow (d, w) \in (\tau(D) \sqcap \tau(E))^{\mathcal{I}^{\times}} = \tau(D \sqcap E)^{\mathcal{I}^{\times}} \end{aligned}$$

inductive step:  $C = \exists r. D$

$$\begin{aligned} d \in (\exists r. D)^{\mathcal{I}_w} &\Leftrightarrow \exists e \in \Delta^{\mathcal{I}}: (d, e) \in r^{\mathcal{I}_w} \text{ and } e \in D^{\mathcal{I}_w} \\ &\stackrel{\text{I.H.}}{\Leftrightarrow} \exists e \in \Delta^{\mathcal{I}}: ((d, w), (e, w)) \in \tau(r)^{\mathcal{I}^\times} \text{ and } (e, w) \in \tau(D)^{\mathcal{I}^\times} \\ &\Leftrightarrow (d, w) \in (\exists r. \tau(D))^{\mathcal{I}^\times} = \tau(\exists r. D)^{\mathcal{I}^\times} \end{aligned}$$

The equivalences are also satisfied for probabilistic roles  $\mathbb{P}_{>0} r$ , since

$$(d, e) \in (\mathbb{P}_{>0} r)^{\mathcal{I}_w} \Leftrightarrow ((d, w), (e, w)) \in r_{>0}^{\mathcal{I}^\times}$$

and  $\tau(\mathbb{P}_{>0} r) = r_{>0}$  hold by definition. Analogously for  $\mathbb{P}_{=1} r$ .

inductive step:  $C = \forall r. D$

$$\begin{aligned} d \in (\forall r. D)^{\mathcal{I}_w} &\Leftrightarrow \forall e \in \Delta^{\mathcal{I}}: (d, e) \in r^{\mathcal{I}_w} \text{ implies } e \in D^{\mathcal{I}_w} \\ &\stackrel{\text{I.H.}}{\Leftrightarrow} \forall e \in \Delta^{\mathcal{I}}: ((d, w), (e, w)) \in \tau(r)^{\mathcal{I}^\times} \text{ implies } (e, w) \in \tau(D)^{\mathcal{I}^\times} \\ &\Leftrightarrow (d, w) \in (\forall r. \tau(D))^{\mathcal{I}^\times} = \tau(\forall r. D)^{\mathcal{I}^\times} \end{aligned}$$

With the same arguments as for existential restrictions, the statements also hold for probabilistic roles.

inductive step:  $C = \mathbb{P}_{>0} D$

$$\begin{aligned} d \in (\mathbb{P}_{>0} C)^{\mathcal{I}_w} &\Leftrightarrow \exists v \in W: \mathbb{P} \{ v \} > 0 \text{ and } d \in C^{\mathcal{I}_v} \\ &\stackrel{\text{I.H.}}{\Leftrightarrow} \exists v \in W: ((d, w), (d, v)) \in \omega_{\mathbb{P}}^{\mathcal{I}^\times} \text{ and } (d, v) \in \tau(C)^{\mathcal{I}^\times} \\ &\Leftrightarrow (d, w) \in (\exists \omega_{\mathbb{P}}. \tau(C))^{\mathcal{I}^\times} = \tau(\mathbb{P}_{>0} C)^{\mathcal{I}^\times} \end{aligned}$$

inductive step:  $C = \mathbb{P}_{=1} D$

$$\begin{aligned} d \in (\mathbb{P}_{=1} C)^{\mathcal{I}_w} &\Leftrightarrow \forall v \in W: \mathbb{P} \{ v \} > 0 \text{ implies } d \in C^{\mathcal{I}_v} \\ &\stackrel{\text{I.H.}}{\Leftrightarrow} \forall v \in W: ((d, w), (d, v)) \in \omega_{\mathbb{P}}^{\mathcal{I}^\times} \text{ implies } (d, v) \in \tau(C)^{\mathcal{I}^\times} \\ &\Leftrightarrow (d, w) \in (\forall \omega_{\mathbb{P}}. \tau(C))^{\mathcal{I}^\times} = \tau(\mathbb{P}_{=1} C)^{\mathcal{I}^\times} \end{aligned}$$

□

As a corollary it follows that  $C^{\mathcal{I}_w} \times \{ w \} = \tau(C)^{\mathcal{I}^\times} \cap (\Delta^{\mathcal{I}} \times \{ w \})$  and hence also

$$\tau(C)^{\mathcal{I}^\times} = \bigsqcup_{w \in W} C^{\mathcal{I}_w} \times \{ w \}$$

hold for all  $\mathbb{P}_{01}\mathcal{FL}\mathcal{E}^\perp$ -concept descriptions  $C$  and all probabilistic interpretations  $\mathcal{I}$ .

## 4 Construction of a Base of GCIs in $\mathbb{P}_{01}\mathcal{FL}\mathcal{E}^\perp$

The translation  $\tau$  can additionally be used to translate valid general concept inclusions of  $\mathcal{I}$  into valid general concept inclusions of  $\mathcal{I}^\times$ . Since  $\tau$  has an inverse we may also translate GCIs in the opposite direction. A more sophisticated characterisation is given in the next lemma.

**Lemma 2.** *Let  $\mathcal{I}$  be a probabilistic interpretation and  $C, D$  be  $\mathbb{P}_{01}\mathcal{FL}\mathcal{E}^\perp$ -concept descriptions. Then the general concept inclusion  $C \sqsubseteq D$  holds in  $\mathcal{I}$  if and only if the translated GCI  $\tau(C) \sqsubseteq \tau(D)$  holds in  $\mathcal{I}^\times$ .*

*Proof.* Consider an arbitrary individual  $d \in \Delta^{\mathcal{I}}$  and an arbitrary world  $w \in W$ . Then the following equivalences hold:

$$\begin{aligned} \mathcal{I} \models C \sqsubseteq D &\Leftrightarrow \forall w \in W \forall d \in \Delta^{\mathcal{I}} : d \in C^{\mathcal{I}_w} \Rightarrow d \in D^{\mathcal{I}_w} \\ &\stackrel{\text{Lem. 1}}{\Leftrightarrow} \forall (d, w) \in \Delta^{\mathcal{I}} \times W : (d, w) \in \tau(C)^{\mathcal{I}^\times} \Rightarrow (d, w) \in \tau(D)^{\mathcal{I}^\times} \\ &\Leftrightarrow \mathcal{I}^\times \models \tau(C) \sqsubseteq \tau(D). \end{aligned}$$

□

Having all necessary notions and lemmata at hand, we are now ready to formulate the main proposition for the construction of a base of GCIs in  $\mathbb{P}_{01}\mathcal{FL}\mathcal{E}^\perp$ . We have seen that we may translate between valid GCIs of  $\mathcal{I}$  and  $\mathcal{I}^\times$ , and the following proposition shows that it is possible to translate a base for  $\mathcal{I}^\times$  into a base for  $\mathcal{I}$ .

**Proposition 3.** *Let  $\mathcal{I}$  be a probabilistic interpretation. Every base of GCIs for the interpretation  $\mathcal{I}^\times$  can be translated into a base of GCIs for  $\mathcal{I}$ ; in particular, if  $\mathcal{B}$  is a base of GCIs for  $\mathcal{I}^\times$ , then the set  $\tau^{-1}(\mathcal{B}) := \{ \tau^{-1}(C) \sqsubseteq \tau^{-1}(D) \mid C \sqsubseteq D \in \mathcal{B} \}$  is a base of GCIs for  $\mathcal{I}$ .*

*Proof.* Firstly, we show soundness of the translation  $\tau^{-1}(\mathcal{B})$ . For this purpose consider a GCI  $C \sqsubseteq D \in \mathcal{B}$ . Since  $\mathcal{B}$  is a base for  $\mathcal{I}^\times$ , it follows that  $C \sqsubseteq D$  holds in  $\mathcal{I}^\times$ . By Lemma 6, we may conclude that  $\tau^{-1}(C) \sqsubseteq \tau^{-1}(D)$  holds in  $\mathcal{I}$ .

Secondly, we prove completeness of  $\tau^{-1}(\mathcal{B})$ . Let  $C \sqsubseteq D$  be a GCI holding in  $\mathcal{I}$ . Lemma 6 then states that  $\tau(C) \sqsubseteq \tau(D)$  holds in  $\mathcal{I}^\times$  and thus follows from  $\mathcal{B}$ . It remains to show that  $\tau^{-1}(\mathcal{B})$  entails  $C \sqsubseteq D$ . Consider an arbitrary model  $\mathcal{J}$  of the translation  $\tau^{-1}(\mathcal{B})$ . Using Lemma 6 it follows that  $\mathcal{J}^\times$  must be a model of  $\mathcal{B}$ . By completeness of  $\mathcal{B}$ , we conclude that  $\mathcal{J}^\times \models \tau(C) \sqsubseteq \tau(D)$ , and finally Lemma 6 yields that  $\mathcal{J} \models C \sqsubseteq D$ . Consequently,  $\tau^{-1}(\mathcal{B})$  is complete for  $\mathcal{I}$ . □

However, the converse direction cannot be shown as not every interpretation over  $(N_C, N_R)_{01}^{\mathbb{P}}$  is induced by a probabilistic interpretation and hence may have different extensions for the additional role names. So we are not able to prove that minimality of the base is preserved.

## 5 Translation between $\mathbb{P}_{\geq}\mathcal{FL}\mathcal{E}^\perp Q^{\geq}$ and $\mathcal{FL}\mathcal{E}^\perp Q^{\geq}$

A probabilistic interpretation  $\mathcal{I}$  is called *quasi-uniform* if all possible worlds have the same probability, i.e., if  $\mathbb{P}\{v\} = \mathbb{P}\{w\}$  holds for all  $v, w \in W_\varepsilon$ . Then  $\mathbb{P}\{w\} = \varepsilon$  holds for all possible worlds  $w \in W_\varepsilon$  where  $\varepsilon := \frac{1}{|W_\varepsilon|}$ , and in particular it follows that only finitely many possible worlds exist. A quasi-uniform probabilistic interpretation is *uniform* if it does not contain impossible worlds.

In this section we only consider quasi-uniform probabilistic interpretations. Hence, let  $\mathcal{I}$  be quasi-uniform with probability  $\varepsilon$  for each possible world. We will extend the translation function  $\tau$  as introduced in the previous section to a translation  $\tau_\varepsilon$  from  $\mathbb{P}_{\geq}\mathcal{FL}\mathcal{E}^\perp Q^{\geq}$  to



$\mathcal{FL}\mathcal{E}^\perp\mathcal{Q}^\geq$ . For this purpose we have to extend the signature  $(N_C, N_R)$  by adding new roles  $r_{\geq k}$  for each role name  $r \in N_R$ , in particular we define:

$$(N_C, N_R)_{\geq}^{\mathbb{P}} := (N_C, N_R \uplus \{ \omega_{\mathbb{P}} \} \uplus \{ r_{\geq k} \mid k \in \{1, \dots, |W_\varepsilon|\} \}).$$

Then the mapping  $\tau_\varepsilon$  extends  $\tau$  as follows:

$$\begin{aligned} \tau_\varepsilon(\mathbb{P}_{\geq p} r) &:= r_{\geq \lceil \frac{p}{\varepsilon} \rceil} & \tau_\varepsilon^{-1}(r_{\geq k}) &:= \mathbb{P}_{\geq k \cdot \varepsilon} r \\ \tau_\varepsilon(\mathbb{P}_{\geq p} C) &:= \exists \omega_{\mathbb{P}}. \tau_\varepsilon(C) \quad (p \in (0, \varepsilon]) & \tau_\varepsilon^{-1}(\exists \omega_{\mathbb{P}}. C) &:= \mathbb{P}_{\geq \varepsilon} \tau_\varepsilon^{-1}(C) \\ \tau_\varepsilon(\mathbb{P}_{\geq p} C) &:= \geq \lceil \frac{p}{\varepsilon} \rceil. \omega_{\mathbb{P}}. \tau_\varepsilon(C) \quad (p \in (\varepsilon, 1)) & \tau_\varepsilon^{-1}(\geq n. \omega_{\mathbb{P}}. C) &:= \mathbb{P}_{\geq n \cdot \varepsilon} \tau_\varepsilon^{-1}(C) \\ \tau_\varepsilon(\mathbb{P}_{\geq 1} C) &:= \forall \omega_{\mathbb{P}}. \tau_\varepsilon(C) & \tau_\varepsilon^{-1}(\forall \omega_{\mathbb{P}}. C) &:= \mathbb{P}_{\geq 1} \tau_\varepsilon^{-1}(C) \end{aligned}$$

Of course, the induced interpretation  $\mathcal{I}^\times$  must also interpret the new role names  $r_{\geq k}$ . Hence, we define the following extensions for them:

$$\begin{aligned} r_{\geq k}^{\mathcal{I}^\times} &:= \{ ((d, w), (e, w)) \mid (d, e) \in (\mathbb{P}_{\geq k \cdot \varepsilon} r)^{\mathcal{I}_w} \} \\ &= \{ ((d, w), (e, w)) \mid \mathbb{P}\{ w \in W \mid (d, e) \in r^{\mathcal{I}_w} \} \geq k \cdot \varepsilon \}, \end{aligned}$$

i.e.,  $((d, w), (e, w)) \in r_{\geq k}^{\mathcal{I}^\times}$  holds iff there are  $k$  possible worlds  $w$  that satisfy  $(d, e) \in r^{\mathcal{I}_w}$ .

Unfortunately, then the mappings  $\tau_\varepsilon$  and  $\tau_\varepsilon^{-1}$  are not mutually inverse. It only holds that  $\tau_\varepsilon(\tau_\varepsilon^{-1}(C)) = C$ . For the concept description  $C = \mathbb{P}_{\geq p} A$  it holds that  $\tau_\varepsilon(\mathbb{P}_{\geq p} A) = \geq \lceil \frac{p}{\varepsilon} \rceil. \omega_{\mathbb{P}}. A$ , and hence  $\tau_\varepsilon^{-1}(\tau_\varepsilon(\mathbb{P}_{\geq p} A)) = \mathbb{P}_{\geq \varepsilon \cdot \lceil \frac{p}{\varepsilon} \rceil} A$ . Obviously, if  $p$  is not of the form  $k \cdot \varepsilon$  for a  $k \in \mathbb{N}$ , then both concept descriptions are not the same. However, we may show that the concepts  $\tau_\varepsilon^{-1}(\tau_\varepsilon(C))$  and  $C$  have the same extensions w.r.t. the interpretation  $\mathcal{I}$ .

**Lemma 4.** *Let  $\mathcal{I}$  be a quasi-uniform probabilistic interpretation with  $\frac{1}{\varepsilon}$  possible worlds, i.e., the probability of each possible world is  $\varepsilon$ . Then for each  $\mathbb{P}_{\geq} \mathcal{FL}\mathcal{E}^\perp\mathcal{Q}^\geq$ -concept description  $C$  and all worlds  $w \in W$  the following equation holds:*

$$C^{\mathcal{I}_w} = (\tau_\varepsilon^{-1}(\tau_\varepsilon(C)))^{\mathcal{I}_w}.$$

*Proof.* by structural induction on  $C$ .

(induction base) Let  $C = A$  be a concept name. Then it holds that  $\tau_\varepsilon^{-1}(\tau_\varepsilon(A)) = A$ , and hence the claim is trivial.

(induction step) At first consider a probabilistic concept description  $C = \mathbb{P}_{\geq p} D$ . Then we have the following equivalences:

$$\begin{aligned} d \in (\mathbb{P}_{\geq p} D)^{\mathcal{I}_w} &\Leftrightarrow \mathbb{P}\{ w \in W \mid d \in D^{\mathcal{I}_w} \} \geq p \\ &\stackrel{*}{\Leftrightarrow} \mathbb{P}\{ w \in W \mid d \in D^{\mathcal{I}_w} \} \geq \varepsilon \cdot \lceil \frac{p}{\varepsilon} \rceil \\ &\stackrel{\text{IH}}{\Leftrightarrow} \mathbb{P}\{ w \in W \mid d \in (\tau_\varepsilon^{-1}(\tau_\varepsilon(D)))^{\mathcal{I}_w} \} \geq \varepsilon \cdot \lceil \frac{p}{\varepsilon} \rceil \\ &\Leftrightarrow d \in (\mathbb{P}_{\geq \varepsilon \cdot \lceil \frac{p}{\varepsilon} \rceil} \tau_\varepsilon^{-1}(\tau_\varepsilon(D)))^{\mathcal{I}_w} \\ &\Leftrightarrow d \in (\tau_\varepsilon^{-1}(\tau_\varepsilon(\mathbb{P}_{\geq p} D)))^{\mathcal{I}_w}. \end{aligned}$$

For the equivalence \* note that  $p \leq \varepsilon \cdot \lceil \frac{p}{\varepsilon} \rceil$  always holds. The other direction follows from the fact that for each individual  $d$  which satisfies  $\mathbb{P}\{w \in W \mid d \in D^{\mathcal{I}_w}\} \geq p$  there must be at least  $\lceil \frac{p}{\varepsilon} \rceil$  possible worlds  $w$  with  $d \in D^{\mathcal{I}_w}$  since all possible worlds have probability  $\varepsilon$ . Hence, it suffices to enforce a probability  $\geq \varepsilon \cdot \lceil \frac{p}{\varepsilon} \rceil$ .

Analogously, we can prove that  $s^{\mathcal{I}_w} = (\tau_\varepsilon^{-1}(\tau_\varepsilon(s)))^{\mathcal{I}_w}$  holds for all (probabilistic) roles  $s$ .

Consider a conjunction  $C = D \sqcap E$ . Then we infer the following equations:

$$\begin{aligned} (D \sqcap E)^{\mathcal{I}_w} &= D^{\mathcal{I}_w} \cap E^{\mathcal{I}_w} \\ &\stackrel{\text{I.H.}}{=} (\tau_\varepsilon^{-1}(\tau_\varepsilon(D)))^{\mathcal{I}_w} \cap (\tau_\varepsilon^{-1}(\tau_\varepsilon(E)))^{\mathcal{I}_w} \\ &= (\tau_\varepsilon^{-1}(\tau_\varepsilon(D)) \sqcap \tau_\varepsilon^{-1}(\tau_\varepsilon(E)))^{\mathcal{I}_w} \\ &= (\tau_\varepsilon^{-1}(\tau_\varepsilon(D \sqcap E)))^{\mathcal{I}_w}. \end{aligned}$$

Finally, let  $C = \exists s. D$  be an existential restriction. Then we can make the following observations:

$$\begin{aligned} d \in (\exists s. D)^{\mathcal{I}_w} &\Leftrightarrow \exists e \in \Delta^{\mathcal{I}}: (d, e) \in s^{\mathcal{I}_w} \text{ and } e \in D^{\mathcal{I}_w} \\ &\stackrel{\text{I.H.}}{\Leftrightarrow} \exists e \in \Delta^{\mathcal{I}}: (d, e) \in (\tau_\varepsilon^{-1}(\tau_\varepsilon(s)))^{\mathcal{I}_w} \text{ and } e \in (\tau_\varepsilon^{-1}(\tau_\varepsilon(D)))^{\mathcal{I}_w} \\ &\Leftrightarrow d \in (\exists \tau_\varepsilon^{-1}(\tau_\varepsilon(s)). \tau_\varepsilon^{-1}(\tau_\varepsilon(D)))^{\mathcal{I}_w} \\ &\Leftrightarrow d \in (\tau_\varepsilon^{-1}(\tau_\varepsilon(\exists s. D)))^{\mathcal{I}_w}. \end{aligned}$$

The case of  $C$  being a value restriction or a qualified  $\geq$ -restriction can be treated analogously.  $\square$

**Lemma 5.** *Let  $\mathcal{I}$  be a quasi-uniform probabilistic interpretation with  $\frac{1}{\varepsilon}$  possible worlds. Then for all individuals  $d \in \Delta^{\mathcal{I}}$ , all worlds  $w \in W$ , and all  $\mathbb{P}_{\geq p} \mathcal{FLE}^\perp \mathcal{Q}^{\geq}$ -concept descriptions  $C$ , the following equivalence holds:*

$$d \in C^{\mathcal{I}_w} \Leftrightarrow (d, w) \in \tau_\varepsilon(C)^{\mathcal{I}^\times}.$$

*Proof.* analogously to Lemma 1. We only show the induction step for a concept description  $\mathbb{P}_{\geq p} C$  where  $p \in (\varepsilon, 1)$ . According to the definition of  $\tau_\varepsilon$ , we have that  $\tau_\varepsilon(\mathbb{P}_{\geq p} C) = \geq \lceil \frac{p}{\varepsilon} \rceil \cdot \omega_{\mathbb{P}} \cdot \tau_\varepsilon(C)$  holds. Furthermore, the following equivalences hold:

$$\begin{aligned} d \in (\mathbb{P}_{\geq p} C)^{\mathcal{I}_w} &\Leftrightarrow \mathbb{P}\{w \in W \mid d \in C^{\mathcal{I}_w}\} \geq p \\ &\Leftrightarrow \exists \geq \lceil \frac{p}{\varepsilon} \rceil w \in W_\varepsilon: d \in C^{\mathcal{I}_w} \\ &\stackrel{\text{I.H.}}{\Leftrightarrow} \exists \geq \lceil \frac{p}{\varepsilon} \rceil v \in W: ((d, w), (d, v)) \in \omega_{\mathbb{P}}^{\mathcal{I}^\times} \text{ and } (d, v) \in \tau_\varepsilon(C)^{\mathcal{I}^\times} \\ &\Leftrightarrow (d, w) \in (\geq \lceil \frac{p}{\varepsilon} \rceil \cdot \omega_{\mathbb{P}} \cdot \tau_\varepsilon(C))^{\mathcal{I}^\times}. \end{aligned}$$

$\square$

## 6 Construction of a Base of GCIs in $\mathbb{P}_{\geq p} \mathcal{FLE}^\perp \mathcal{Q}^{\geq}$

In a previous section we have seen how a base of  $\mathbb{P}_{01} \mathcal{FLE}^\perp$ -GCIs holding in a probabilistic interpretation  $\mathcal{I}$  can be constructed by means of a base of  $\mathcal{FLE}^\perp$ -GCIs holding in the induced

interpretation  $\mathcal{I}^\times$  over the extended signature  $(N_C, N_R)_{01}^{\mathbb{P}}$ . A similar result can be obtained in the case of a uniform probabilistic interpretation in the description logic  $\mathbb{P}_{\geq} \mathcal{FL}\mathcal{E}^\perp \mathcal{Q}^\geq$ . A more sophisticated answer is given in the next proposition.

**Lemma 6.** *Let  $\mathcal{I}$  be a quasi-uniform probabilistic interpretation with  $\frac{1}{\varepsilon}$  possible worlds, and  $C, D$  let be  $\mathbb{P}_{\geq} \mathcal{FL}\mathcal{E}^\perp \mathcal{Q}^\geq$ -concept descriptions. Then the general concept inclusion  $C \sqsubseteq D$  holds in  $\mathcal{I}$  if and only if the translated GCI  $\tau_\varepsilon(C) \sqsubseteq \tau_\varepsilon(D)$  holds in  $\mathcal{I}^\times$ .*

*Proof.* analogously to Lemma 2. □

**Proposition 7.** *Let  $\mathcal{I}$  be a quasi-uniform probabilistic interpretation with  $\frac{1}{\varepsilon}$  possible worlds. If  $\mathcal{B}$  is a base of  $\mathcal{FL}\mathcal{E}^\perp \mathcal{Q}^\geq$ -GCIs for the induced interpretation  $\mathcal{I}^\times$ , then the translation*

$$\tau_\varepsilon^{-1}(\mathcal{B}) := \{ \tau_\varepsilon^{-1}(C) \sqsubseteq \tau_\varepsilon^{-1}(D) \mid C \sqsubseteq D \in \mathcal{B} \}$$

*is a base of  $\mathbb{P}_{\geq} \mathcal{FL}\mathcal{E}^\perp \mathcal{Q}^\geq$ -GCIs for  $\mathcal{I}$ .*

*Proof.* analogously to Proposition 3. □

## 7 Probabilistic Model-Based Most-Specific Concept Descriptions

Model-based most-specific concept descriptions (w.r.t. interpretations) have been introduced by Baader and Distel in [3, 2] as an adaption of the well-known notion of most-specific concept descriptions (w.r.t. knowledge bases). Ecke, Peñaloza, and Turhan, investigated those most-specific concept descriptions and also least common subsumers in probabilistic extensions of the light-weight description logic  $\mathcal{EL}$ , cf. [13, 14, 15, 9, 10]. However, they gave constructions for those generalisations w.r.t. knowledge bases (w.r.t. open-world assumption). In the following text the notion of a probabilistic mmsc w.r.t. interpretations (w.r.t. closed-world assumption) is introduced. Furthermore, we present a proposition that reduces their computation to crisp description logics.

**Definition 8.** *Let  $\mathcal{I} = (\Delta^{\mathcal{I}}, W, (\cdot^{\mathcal{I}_w})_{w \in W}, \mathbb{P})$  be a probabilistic interpretation and  $X \subseteq \Delta^{\mathcal{I}}$  a set of individuals. Then a  $\mathbb{P}_{01} \mathcal{FL}\mathcal{E}^\perp$ -concept description  $C$  is called probabilistic model-based most-specific concept description (pmmsc) of  $X$  in  $\mathcal{I}$  if it satisfies the following conditions:*

(PM1) *For all worlds  $w \in W$  it holds that  $X \subseteq C^{\mathcal{I}_w}$ .*

(PM2) *If  $D$  is a  $\mathbb{P}_{01} \mathcal{FL}\mathcal{E}^\perp$ -concept description such that  $X \subseteq \bigcap_{w \in W} D^{\mathcal{I}_w}$  holds, then  $C \sqsubseteq D$ .*

*In the same way we may define the pmmsc in  $\mathbb{P}_{\geq} \mathcal{FL}\mathcal{E}^\perp \mathcal{Q}^\geq$ .*

All pmmscs for a subset  $X$  in  $\mathcal{I}$  are equivalent, and hence we shall denote *the* pmmsc by  $X^{\mathcal{I}}$ .

**Proposition 9.** *Let  $\mathcal{I}$  be a probabilistic interpretation and  $X \subseteq \Delta^{\mathcal{I}}$  a set of individuals. Then the following statements hold:*

1. *The  $\mathbb{P}_{01} \mathcal{FL}\mathcal{E}^\perp$ -mmsc  $X^{\mathcal{I}}$  is equivalent to the translation  $\tau^{-1}((X \times W)^{\mathcal{I}^\times})$ .*

2. If  $\mathcal{I}$  is quasi-uniform with  $\frac{1}{\epsilon}$  possible worlds, then the  $\mathbb{P}_{\geq} \mathcal{FLE}^{\perp} Q^{\geq}$ -mmsc  $X^{\mathcal{I}}$  is equivalent to the translation  $\tau_{\epsilon}^{-1}((X \times W)^{\mathcal{I}^{\times}})$ .

*Proof.* Both statements can be proven analogously. We show the two conditions of a pmmsc according to Definition 8.

Firstly, the definition of the mmsc in the default setting yields that  $X \times W \subseteq (X \times W)^{\mathcal{I}^{\times} \mathcal{I}^{\times}}$ . From Lemmata 1 and 5 it follows that  $X \subseteq (\tau^{-1}((X \times W)^{\mathcal{I}^{\times}}))^{\mathcal{I}_w}$  for all worlds  $w \in W$ , i.e.,  $\tau^{-1}((X \times W)^{\mathcal{I}^{\times}})$  satisfies (PM1).

Secondly, consider a concept description  $D$  such that  $X \subseteq D^{\mathcal{I}_w}$  holds for all worlds  $w \in W$ . Consequently, by Lemmata 1 and 5 it follows that  $X \times \{w\} \subseteq \tau(D)^{\mathcal{I}^{\times}}$  holds for all  $w \in W$ , i.e.,  $X \times W \subseteq \tau(D)^{\mathcal{I}^{\times}}$  is true. By definition of mmscs, we conclude that  $(X \times W)^{\mathcal{I}^{\times}} \sqsubseteq \tau(D)$  holds. Then Lemmata 2 and 6 yield that  $\tau^{-1}(X \times W)^{\mathcal{I}^{\times}} \sqsubseteq D$  holds, and hence  $\tau^{-1}(X \times W)^{\mathcal{I}^{\times}}$  satisfies (PM2).  $\square$

## 8 Choice of Semantics

Upon translation of the probabilistic interpretation  $\mathcal{I}$  to the crisp interpretation  $\mathcal{I}^{\times}$ , we have to introduce the additional role  $\omega_{\mathbb{P}}$  to encode the possibility of worlds. However, this leads to cyclic interpretations as then every pair  $(p, w)$  is connected to all pairs  $(p, v)$  where  $v \in W_{\epsilon}$  is a possible world. Of course,  $\mathcal{I}$  must contain at least one possible world to ensure that  $\mathbb{P}(W_{\epsilon}) = 1$  holds. However, in cyclic interpretations like  $\mathcal{I}^{\times}$  all model-based most-specific concept descriptions only exist w.r.t. a role-depth bound (cf. [6]), or in gfp-semantics (cf. [8]). The limitation of the role-depth is a practical means to ensure the existence of mmscs and is used here.

Usually, mmscs are computed from description graphs induced by interpretations. It turns out that in the case of interpretations constructed from probabilistic interpretations we do not have to consider all paths in the graph. In particular, the following lemma shows that we may ignore paths with two subsequent  $Q$   $\omega_{\mathbb{P}}$ -edges.

**Lemma 10.** For arbitrary  $\mathbb{P}_{\geq} \mathcal{FLE}^{\perp} Q^{\geq}$ -concept descriptions  $C, D$  and probability thresholds  $p, q \in (0, 1]$  the following equivalence holds:

$$\mathbb{P}_{\geq p} (C \sqcap \mathbb{P}_{\geq q} D) \equiv \mathbb{P}_{\geq p} C \sqcap \mathbb{P}_{\geq q} D.$$

*Proof.* The statement easily follows from the following observations:

$$\begin{aligned} & d \in (\mathbb{P}_{\geq p} (C \sqcap \mathbb{P}_{\geq q} D))^{\mathcal{I}_w} \\ \Leftrightarrow & \mathbb{P}\{w \in W \mid d \in C^{\mathcal{I}_w} \text{ and } \mathbb{P}\{w \in W \mid d \in D^{\mathcal{I}_w}\} \geq q\} \geq p \\ \Leftrightarrow & \mathbb{P}\{w \in W \mid d \in C^{\mathcal{I}_w}\} \geq p \text{ and } \mathbb{P}\{w \in W \mid d \in D^{\mathcal{I}_w}\} \geq q \\ \Leftrightarrow & d \in (\mathbb{P}_{\geq p} C \sqcap \mathbb{P}_{\geq q} D)^{\mathcal{I}_w}. \end{aligned}$$

As the equivalences hold for arbitrary probabilistic interpretations  $\mathcal{I}$  and worlds  $w \in W$ , the concept equivalence is true in general.  $\square$

The lemma above yields (after translation w.r.t.  $\tau$ , or  $\tau_e$ , respectively) that in order to compute model-based most-specific concept descriptions we do not have to consider any paths in the description graph of  $\mathcal{I}^\times$  that have two subsequent  $Q \omega_P$ -edges.

However, we cannot interchange  $\exists r$  and  $\mathbb{P}_{\geq p}$  restrictions as even in the simplest case the concept descriptions  $\exists r. \mathbb{P}_{\geq \frac{1}{2}} A$  and  $\mathbb{P}_{\geq \frac{1}{2}} \exists r. A$  may have different extensions in a probabilistic interpretation. Consider for example the uniform probabilistic interpretation  $\mathcal{I} = (\{d, e\}, \{v, w\}, \mathcal{I}_v, \mathcal{I}_w, \{\{v\} \mapsto \frac{1}{2}, \{w\} \mapsto \frac{1}{2}\})$ . The extension functions are given by the two graphs below:



Then it holds that  $(\exists r. \mathbb{P}_{\geq \frac{1}{2}} A)^{\mathcal{I}_{w_1}} = \{d\}$ , but  $(\mathbb{P}_{\geq \frac{1}{2}} \exists r. A)^{\mathcal{I}_{w_1}} = \emptyset$ .

For the computation of the induced interpretation  $\mathcal{I}^\times$  all vertices in the two graphs above are equipped with  $\omega_P$ -loops and furthermore there are  $\omega_P$ -edges between vertices for the same individual in different worlds. Then the following mmscs can be obtained:

$$\begin{aligned} (d, w_1)^{\mathcal{I}^\times} &= \exists r. \mathbb{P}_{>0} A \sqcap \mathbb{P}_{>0} \exists r. \mathbb{P}_{>0} A \\ (d, w_2)^{\mathcal{I}^\times} &= \mathbb{P}_{>0} \exists r. \mathbb{P}_{>0} A \\ (e, w_1)^{\mathcal{I}^\times} &= \mathbb{P}_{>0} A \\ (e, w_2)^{\mathcal{I}^\times} &= A \sqcap \mathbb{P}_{>0} A \\ (\Delta^{\mathcal{I}} \times W)^{\mathcal{I}^\times} &= \top \end{aligned}$$

If  $\mathcal{I}$  is a probabilistic interpretation such that all mmscs exist in the interpretations  $\mathcal{I}_w$  for all worlds  $w \in W$ , i.e., all interpretations  $\mathcal{I}_w$  are acyclic, then also all mmscs exist in the induced interpretation  $\mathcal{I}^\times$ . They can be computed by means of restricted unravellings as follows: Consider the description graph  $\mathcal{G}_{\mathcal{I}^\times}^{(d,w)}$  of  $\mathcal{I}^\times$  that is rooted at  $(d, w)$ . Then we consider the restricted unravelling  $\mathcal{G}_{\mathcal{I}^\times}^{(d,w)} \upharpoonright_{\infty}^{\omega_P}$  such that only paths in  $\mathcal{G}_{\mathcal{I}^\times}^{(d,w)}$  that do not have two subsequent  $\omega_P$ -edges are allowed as vertices in the restricted unravelling. Since the mmscs in all interpretations  $\mathcal{I}_w$  exist, there are no infinite paths from each  $(e, v)$ . As a consequence, we obtain that all model-based most-specific concept descriptions exist in  $\mathcal{I}^\times$ .

## 9 Complexity of Base Construction

The complexity of the construction of a probabilistic base can be double-exponential in the size of the input interpretation. The translation of the probabilistic interpretation  $\mathcal{I}$  to the crisp interpretation  $\mathcal{I}^\times$  can be obtained in polynomial time. The same holds for the translation of concept descriptions, i.e., they may be translated in polynomial time. Furthermore, the computation of a base of GCIs for a crisp interpretation has double-exponential time complexity in the worst case. This is due to the fact that the necessary induced context  $\mathbb{K}_{\mathcal{I}}$  of an interpretation may have exponential size in  $\mathcal{I}$  (since there may be exponentially many model-based most-specific concept descriptions for  $\mathcal{I}$ ), and furthermore the canonical

implicational base of a formal context may have an exponential size w.r.t. the size of the formal context. Hence, the construction of bases of GCIs for probabilistic interpretations also has a double-exponential time complexity in the worst case.

## 10 Conclusion

We have defined translations between probabilistic description logics and crisp description logics that preserve entailment of general concept inclusions. They have been used to reduce the problem of construction of a base of GCIs for a probabilistic interpretation to the same problem in crisp description logics for which a well-known and practical solution exists. For this purpose we used the description  $\mathcal{FL}\mathcal{E}^\perp\mathcal{Q}^\geq$  that was equipped with probabilistic role and concept constructors, in the first case only allowing probabilities  $> 0$  and  $= 1$  to express *possibility* and *certainty almost everywhere*, and in the other case only allowing for lower-bound probabilities  $\geq p$  where  $p \in (0, 1]$ .

Furthermore, we have shown how most-specific concept descriptions can be constructed for probabilistic interpretations – again by a reduction to the crisp case.

## References

- [1] Franz Baader, Diego Calvanese, Deborah L. McGuinness, Daniele Nardi, and Peter F. Patel-Schneider, editors. *The Description Logic Handbook: Theory, Implementation, and Applications*. Cambridge University Press, New York, NY, USA, 2003.
- [2] Franz Baader and Felix Distel. A finite basis for the set of EL-implications holding in a finite model. Technical Report 07-02, Inst. für Theoretische Informatik, TU Dresden, Dresden, Germany, 2007.
- [3] Franz Baader and Felix Distel. A finite basis for the set of EL-implications holding in a finite model. In Raoul Medina and Sergei Obiedkov, editors, *Proceedings of the 6th International Conference on Formal Concept Analysis, (ICFCA 2008)*, volume 4933 of *Lecture Notes in Artificial Intelligence*, pages 46–61. Springer, 2008.
- [4] Daniel Borchmann. Towards an error-tolerant construction of  $\mathcal{EL}^\perp$ -ontologies from data using formal concept analysis. In *Formal Concept Analysis, 11th International Conference, ICFCA 2013, Dresden, Germany, May 21-24, 2013. Proceedings*, pages 60–75, 2013.
- [5] Daniel Borchmann. *Learning Terminological Knowledge with High Confidence from Erroneous Data*. PhD thesis, Dresden University of Technology, Dresden, Germany, 2014.
- [6] Daniel Borchmann, Felix Distel, and Francesco Kriegel. Axiomatization of General Concept Inclusions from Finite Interpretations. LTCS-Report 15-13, Chair for Automata Theory, Institute for Theoretical Computer Science, Technische Universität Dresden, Dresden, Germany, 2015.
- [7] Felix Distel. Model-based most specific concepts in description logics with value restrictions. Technical Report 08-04, Institute for theoretical computer science, TU Dresden, Dresden, Germany, 2008.

- [8] Felix Distel. *Learning Description Logic Knowledge Bases from Data using Methods from Formal Concept Analysis*. PhD thesis, Dresden University of Technology, Dresden, Germany, 2011.
- [9] Andreas Ecke, Rafael Peñaloza, and Anni-Yasmin Turhan. Role-depth bounded least common subsumer in prob- $\mathcal{EL}$  with nominals. In Thomas Eiter, Birte Glimm, Yevgeny Kazakov, and Markus Krötzsch, editors, *Proceedings of the 26th International Workshop on Description Logics (DL-2013)*, volume 1014 of *CEUR-WS*, pages 670–688, Ulm, Germany, July 2013.
- [10] Andreas Ecke, Rafael Peñaloza, and Anni-Yasmin Turhan. Completion-based generalization inferences for the description logic  $\mathcal{ELOR}$  with subjective probabilities. *International Journal of Approximate Reasoning*, 55(9):1939–1970, 2014.
- [11] Francesco Kriegel. Extracting ALEQR(Self)-knowledge bases from graphs. In *Proceedings of the International Workshop on Social Network Analysis using Formal Concept Analysis (SNAFCA-2015)*, CEUR Workshop Proceedings. CEUR-WS.org, 2015.
- [12] Carsten Lutz and Lutz Schröder. Probabilistic description logics for subjective uncertainty. In *Principles of Knowledge Representation and Reasoning: Proceedings of the Twelfth International Conference, KR 2010, Toronto, Ontario, Canada, May 9-13, 2010*, 2010.
- [13] Rafael Peñaloza and Anni-Yasmin Turhan. Role-depth bounded least common subsumers by completion for  $\mathcal{EL}$ - and Prob- $\mathcal{EL}$ -TBoxes. In V. Haarslev, D. Toman, and G. Weddell, editors, *Proc. of the 2010 Description Logic Workshop (DL'10)*, volume 573 of *CEUR-WS*, 2010.
- [14] Rafael Peñaloza and Anni-Yasmin Turhan. Towards approximative most specific concepts by completion for  $\mathcal{EL}$  with subjective probabilities. In Thomas Lukasiewicz, Rafael Peñaloza, and Anni-Yasmin Turhan, editors, *Proceedings of the First International Workshop on Uncertainty in Description Logics (UniDL'10)*, volume 613 of *CEUR-WS*, 2010.
- [15] Rafael Peñaloza and Anni-Yasmin Turhan. Instance-based non-standard inferences in  $\mathcal{EL}$  with subjective probabilities. In Fernando Bobillo, Paulo C. G. Costa, Claudia d'Amato, Nicola Fanizzi, Kathryn B. Laskey, Kenneth J. Laskey, Thomas Lukasiewicz, Matthias Nickles, and Michael Pool, editors, *Uncertainty Reasoning for the Semantic Web II, International Workshops URSW 2008-2010 Held at ISWC and UniDL 2010 Held at FLoC, Revised Selected Papers*, number 7123 in *Lecture Notes in Computer Science*, pages 80–98. Springer-Verlag, 2013.

## A Unused Results

**Lemma 11.** *Let  $C$  and  $D$  be two  $\mathbb{P}_{01}\mathcal{FL}\mathcal{E}^\perp$ -concept descriptions. If  $\tau(C) \sqsubseteq \tau(D)$  holds, then also  $C \sqsubseteq D$  is satisfied.*

*Proof.* Let  $\mathcal{I}$  be an arbitrary probabilistic interpretation over  $(N_C, N_R)$ , and consider its induced interpretation  $\mathcal{I}^\times$  over  $(N_C, N_R \cup \{\omega_P\})$ . By presumption, it follows that  $\tau(C)^{\mathcal{I}^\times} \subseteq \tau(D)^{\mathcal{I}^\times}$ . Now consider an arbitrary world  $w \in W$  and an individual  $d \in \Delta^{\mathcal{I}}$ . Using the previous Lemma 1 we get the following:

$$\begin{aligned} d \in C^{\mathcal{I}_w} &\Leftrightarrow (d, w) \in \tau(C)^{\mathcal{I}^\times} \\ &\Rightarrow (d, w) \in \tau(D)^{\mathcal{I}^\times} \\ &\Leftrightarrow d \in D^{\mathcal{I}_w}. \end{aligned}$$

As a consequence, we get that  $C \sqsubseteq D$  holds. □