Journal of Artificial Intelligence Research 58 (2017) 1-66 Submitted 05/16; published 01/17

Probabilistic Description Logics for Subjective Uncertainty

Víctor Gutiérrez-Basulto VICTOR@INFORMATIK.UNI-BREMEN.DE
Jean Christoph Jung JEANJUNG@INFORMATIK.UNI-BREMEN.DE
Carsten Lutz CLU@INFORMATIK.UNI-BREMEN.DE
University of Bremen
Bibliothekstraße 1, 28359 Bremen
Germany

Lutz Schröder LUTZ.SCHROEDER@FAU.DE
Friedrich-Alexander-University Erlangen-Nürnberg
Schlossplatz 4, 91054 Erlangen
Germany

Abstract

We propose a family of probabilistic description logics (DLs) that are derived in a principled way from Halpern's probabilistic first-order logic. The resulting probabilistic DLs have a two-dimensional semantics similar to temporal DLs and are well-suited for representing subjective probabilities. We carry out a detailed study of reasoning in the new family of logics, concentrating on probabilistic extensions of the DLs $\mathcal{ALC}$ and $\mathcal{EL}$, and showing that the complexity ranges from PTime via ExpTime and 2ExpTime to undecidable.

1. Introduction

Description logics (DLs) are a popular family of knowledge representation formalisms that underlie ontology languages such as the W3C standards OWL and OWL 2 (W3C, 2012). Since most traditional DLs are fragments of first-order logic (FO), they support only the representation of crisp and definite knowledge, and do not provide any built-in means to represent uncertainty of any kind. This is a shortcoming in application areas in which uncertain aspects of domain concepts need to be modelled. As a guiding example, we consider bio-medical applications where ontologies are being used with particular success. Almost every bio-medical ontology contains uncertain concepts of some sort, although typically modelled in an inappropriate way. Take for example the well-known and widely used medical ontology SNOMED CT (Price & Spackman, 2000), which comprises a variety of classes with names such as '*Probable tubo-ovarian abscess*', '*Natural death with probable cause suspected*', '*Probable diagnosis*', '*Probably present*', '*Basal cell tumour, uncertain whether benign or malignant*', etc. Similar naming schemes can be found in other bio-medical ontologies such as GALEN (Rector & Horrocks, 1997). Since traditional DLs are used to represent these ontologies, the aspect of uncertainty indicated in the names of these concepts is not reflected in their modelling. For example, nothing is said in SNOMED CT about the class '*Natural death with probable cause suspected*' other than that it is a subclass of '*Natural death*'.

The need to represent this and other kinds of uncertainty has led to various proposals for *probabilistic DLs*; for an overview, see the survey by Lukasiewicz and Straccia (2008) and Section 5 later on. The proposals differ considerably in the way in which probabilities are used, in the syntax, in the chosen semantics, and in the intended application. In this paper, we follow a principled approach to defining probabilistic DLs by viewing them as fragments of probabilistic FO (Halpern, 1990) in the same way as classical DLs are fragments of classical FO. Our aim is to obtain a family of probabilistic DLs that are suitable for representing uncertain aspects of domain concepts. We note that an FO-based approach to probabilistic DLs has been advocated earlier (Sebastiani, 1994), but was never developed in serious detail. To some extent, it can be seen as complementary to approaches that obtain probabilistic DLs by extending probabilistic propositional logic, see for example (Lukasiewicz, 2008).

In probabilistic FO as defined by Halpern (1990), a distinction is made between *statistical probabilities* as formalized by *Type 1* probabilistic FO and *subjective probabilities* as formalized by *Type 2* probabilistic FO. The statistical view is concerned with probability distributions on the domain, and the interest is typically in statements such as "80% of all patients with positive serological blood tests actually do have Lyme disease". Contrastingly, the subjective view regards probabilities as degrees of belief and is concerned with probability distributions on a set of possible worlds, each one described by a classical FO structure. In this paper, we focus on subjective probabilisties and define probabilistic DLs as fragments of probabilistic FO of Type 2. Our main reason for doing so is that Type 2 provides an appropriate semantics for probabilistic statements about *individuals*, which is crucial for the intended application, while Type 1 does not. To illustrate, consider the statement "patient John has Lyme disease with at least 80% probability". This cannot be adequately modelled in probabilistic FO of Type 1 since this logic is interpreted in single FO structures (endowed with a probability distribution), where John will either have or not have Lyme disease. In Type 2, we can include in our set of possible worlds some structures in which John has Lyme disease and others in which he does not, and then state that worlds of the former kind have a total probability of at least .8. Note that probabilistic statements about individuals are crucial for capturing uncertain aspects of domain concepts. For example, mentioned SNOMED CT class '*Basal cell tumour, uncertain whether benign or malignant*' contains those (individual!) basal cell tumours that are benign in some worlds and malignant in others, reasonably with both types of worlds exceeding a certain threshold probability.

Just like in Type 2 probabilistic FO, our semantics does not force the user to specify or axiomatize one single probability distribution. Instead, she only specifies constraints on the set of admissible such distributions; in other words, we follow a deductive approach rather than a model-based one (the probably most prevalent example of the latter kind of approach being Bayesian networks). We also emphasize that, in contrast to several other proposals for probabilistic DLs, the logics in our family are monotonic. While this precludes some typical non-monotonic probabilistic inference patterns (such as changing our degree of belief about the flight abilities of a given bird once we learn that it is a penguin), it results in more transparent semantics. In fact, we follow Halpern in believing that it is important to first understand the monotonic aspects of probability before mixing in non-monotonic ones, which are typically much more controversial.

We consider several probabilistic extensions of the basic description logics $\mathcal{ALC}$ and $\mathcal{EL}$. While $\mathcal{ALC}$ is a well-behaved logical core of the OWL 2 DL ontology language in which standard reasoning tasks are ExpTime-complete (Baader, Calvanese, McGuinness, Nardi, & Patel-Schneider, 2003), $\mathcal{EL}$ is the basic member of a family of DLs that underly the OWL

	base case	+ prob. roles	$P_{>0}C, P_{=1}C$	+ prob. roles
Prob-$\mathcal{ALC}$	ExpTime	≥2ExpTime	ExpTime	2ExpTime
Prob-$\mathcal{EL}$			PTime	PSpace

Table 1: Main results.

2 EL profile of OWL 2 and in which standard reasoning is tractable (Baader, Brandt, & Lutz, 2005). Large-scale medical ontologies such as Snomed CT are often formulated in $\mathcal{EL}$ or slight extensions thereof. For both DLs, we start with a probabilistic extension that adds a concept constructor $P_{\sim n}C$, with $\sim \in \{\leq, <, =, >, \geq\}$, denoting the set of all elements which we believe with probability $\sim n$ to be an instance of the concept C. For example, we can express a probable tubo-ovarian abscess simply as $P_{\geq n}\mathsf{TuboOvarianAbscess}$, for some suitable n. We also consider further extensions with probabilistic role constructors, systems of linear or even polynomial inequalities over concept expressions, and certain forms of independence constraints. For example, in the concept $\exists P_{\geq 0.9}\mathsf{hasDiagnosis}.\,\mathsf{LymeDisease}$ the probabilistic constructor $P_{\geq 0.9}$ is applied to the role name $\mathsf{hasDiagnosis}$. The concept thus denotes the set of all patients that have Lyme disease as a probable diagnosis. As a very simple application of linear inequalities, we can express the class of all lymphomas that are considered more likely to be Mantle cell lymphomas than follicular lymphomas as $\mathsf{Lymphoma} \wedge P(\mathsf{MantleCellLymphoma}) > P(\mathsf{FollicularLymphoma})$.[1] An example for an independence statement is $\mathsf{indep}(\mathsf{AB0}, \mathsf{Male})$ expressing that having blood type AB0 is independent of gender. Full semantic details and more modelling examples will be given in Section 2. Our probabilistic DLs support a rather general form of ABoxes, which also include probabilistic features. For a simple example, the probabilistic ABox assertion $P_{\geq 0.8}\exists\mathsf{hasDisease}.\mathsf{LymeDisease}(\mathsf{john})$ expresses that we believe with probability at least 0.8 that John has Lyme disease, as in our lead-in example. The main difficulty to be dealt with throughout the technical development is that Type 2 structures are inherently two-dimensional, having both a set of domain elements and a set of possible worlds which interact via the probabilistic operators and DL constructors.

1.1 Contributions and Structure of the Paper

The main contribution of this paper is to provide a detailed analysis of the computational complexity of reasoning in various Type 2 probabilistic DLs. In particular, we aim to separate features of our probabilistic DLs that make reasoning hard from those that do not: for probabilistic DLs based on $\mathcal{ALC}$, we are primarily interested in understanding the decidability frontier while, in the case of $\mathcal{EL}$, we are more interested in delineating the border between the tractable cases and the intractable ones. Table 1 shows some of the complexity results obtained, where "$P_{>0}C, P_{=1}C$" means that we restrict the use of the probabilistic concept constructor $P_{\sim n}C$ to the listed combinations of $\sim$ and n, and where each table entry denotes completeness for the displayed complexity class (except for one case in which we know 2ExpTime-hardness, but where decidability is open). In the following, we discuss in detail the structure of the paper and the obtained results.

1. Follicular lymphoma is a differential diagnosis for Mantle cell lymphoma

In Section 2, we start with defining the Type 2 probabilistic DLs that are relevant for this paper, discuss their expressive power and give modelling examples. We also introduce the reasoning tasks concept satisfiability, concept subsumption, knowledge base (KB) consistency, and KB consequence. We then discuss the role of possible worlds with probability zero and argue that, in technical constructions and algorithms, we can disregard them without loss of generality.

Section 3 is concerned with probabilistic extensions of $\mathcal{ALC}$. The most basic such extension is Prob-$\mathcal{ALC}$, extending classical $\mathcal{ALC}$ with the probabilistic concept constructor $P_{\sim n}C$. We use a type elimination procedure to show that all mentioned reasoning tasks are ExpTime-complete in Prob-$\mathcal{ALC}$, thus no more difficult than in non-probabilistic $\mathcal{ALC}$. The proof can easily be adapted to the extension Prob-$\mathcal{ALC}^{\mathsf{lineq}}$ of Prob-$\mathcal{ALC}$ with linear concept inequalities. The procedure also establishes the *uniform model property (UMP)* of Prob-$\mathcal{ALC}^{\mathsf{lineq}}$, meaning that every consistent Prob-$\mathcal{ALC}^{\mathsf{lineq}}$ KB has a uniform model, that is, a model in which all worlds are assigned the same probability. We then give a slightly refined version of the model construction that additionally establishes a *bounded model property (BMP)*, stating that every consistent Prob-$\mathcal{ALC}^{\mathsf{lineq}}$ KB $\mathcal{K}$ has a model in which the number of worlds and the number of domain elements are bounded by $2^{2^{\mathcal{O}(n)}}$, n the size of $\mathcal{K}$. We then observe that the extension Prob-$\mathcal{ALC}^{\mathsf{indep}}$ of Prob-$\mathcal{ALC}$ with independence constraints does not have the UMP. Since Prob-$\mathcal{ALC}^{\mathsf{indep}}$ is a fragment of the extension Prob-$\mathcal{ALC}^{\mathsf{polyeq}}$ of Prob-$\mathcal{ALC}^{\mathsf{lineq}}$ with polynomial concept inequalities, the same applies to Prob-$\mathcal{ALC}^{\mathsf{polyeq}}$. Nevertheless, using a refined algorithm and model construction we are able to show that reasoning in Prob-$\mathcal{ALC}^{\mathsf{polyeq}}$ is still ExpTime-complete. We finally consider versions of Prob-$\mathcal{ALC}$ that admit probabilistic roles. The most basic such extension is Prob-$\mathcal{ALC}_r$, which extends Prob-$\mathcal{ALC}$ with a probabilistic role constructor. We demonstrate that this logic is close to the undecidability frontier by proving that reasoning in both Prob-$\mathcal{ALC}_r^{\mathsf{indep}}$ and Prob-$\mathcal{ALC}_r^{\mathsf{lineq}}$ is undecidable. On the positive side, we show that reasoning in Prob-$\mathcal{ALC}_r$ becomes decidable and 2ExpTime-complete when all probabilistic constructors (both on concepts and roles) are restricted to $P_{>0}$ and $P_{=1}$. In fact, this is an immediate consequence of the observation that the resulting DL Prob-$\mathcal{ALC}_r^{01}$ is a notational variant of the two-dimensional combination of $\mathcal{ALC}$ and the modal logic $\mathsf{S5}$ (Artale, Lutz, & Toman, 2007). The decidability status of Prob-$\mathcal{ALC}_r$ itself remains as an open problem.

In Section 4, we study probabilistic extensions of $\mathcal{EL}$. As in the $\mathcal{ALC}$ case, the basic version Prob-$\mathcal{EL}$ only offers probabilistic concept constructors. Unlike in the $\mathcal{ALC}$ case, though, Prob-$\mathcal{EL}$ is not computationally as well-behaved as one would hope for: reasoning is ExpTime-complete and thus no simpler than in Prob-$\mathcal{ALC}$. The intuitive reason is that extending $\mathcal{EL}$ with probabilities results in *non-convexity*, that is, an implicit form of disjunction becomes available. We start with analyzing this effect in detail. It turns out that we are non-convex even if we only admit a single probabilistic concept constructor $P_{\sim p}$, *for any fixed value* of $\sim \in \{<, \leq, =, >, \geq\}$ and $p \in (0, 1)$. We thus proceed to study the restriction Prob-$\mathcal{EL}^{01}$ of Prob-$\mathcal{EL}$ in which only probabilistic concept constructors of the form $P_{>0}C$ and $P_{=1}C$ are admitted. Note that Prob-$\mathcal{EL}^{01}$ is still a reasonable logic that admits reasoning about possibility (expressed as $P_{>0}C$) and certainty (expressed as $P_{=1}C$). For example, the concept $\mathsf{Patient} \sqcap P_{>0}\exists\mathsf{hadContactWith}(\mathsf{Patient} \sqcap P_{=1}\mathsf{Infected})$ describes the class of patients who possibly were in contact with a patient who certainly was infected. Prob-$\mathcal{EL}^{01}$ can be viewed as a two-dimensional combination of $\mathcal{EL}$ with the conjunctive and positive fragment of the modal logic $\mathsf{S5}$, in the spirit of many-dimensional modal logics (Gabbay, Kurucz, Wolter, & Zakharyaschev, 2003). We use a consequence-based procedure to prove that reasoning in Prob-$\mathcal{EL}^{01}$ is in PTime. As the next step, we consider a different approach to lowering the complexity of Prob-$\mathcal{EL}$, inspired by the observation that many biomedical ontologies such as Snomed CT are *classical TBoxes*, that is, sets of concept definitions $A \equiv D$ with A a concept name. We thus adopt this restricted form of TBox and show that, then, it sometimes becomes possible to use probabilities other than 0 and 1 without losing tractability. In fact, reasoning in Prob-$\mathcal{EL}$ with classical TBoxes remains in PTime when only the constructors $P_{\sim p}$ and $P_{=1}$ are admitted, for any single choice of $\sim \in \{\geq, >, =\}$ and of p; surprisingly, this is not the case for $\sim \in \{<, \leq\}$. The proof again uses a consequence-based procedure. In the final part of Section 4, we extend Prob-$\mathcal{EL}^{01}$ with probabilistic roles. We show that reasoning in the resulting logic Prob-$\mathcal{EL}^{01}_r$ (with unrestricted TBoxes) is PSpace-complete and thus, while not tractable, considerably simpler than reasoning in Prob-$\mathcal{ALC}^{01}_r$. The lower bound is established by a reduction of the word problem of deterministic, polynomially space-bounded Turing machines and the upper bound is proved by a rather subtle version of a consequence-based procedure.

We discuss related work in Section 5. Throughout the paper, we defer some proof details to the appendix. This paper is an extended and refined version of two previous conference papers (Lutz & Schröder, 2010; Gutiérrez-Basulto, Jung, Lutz, & Schröder, 2011). The approach to probabilistic DLs presented in this paper and its conference predecessors has been extended to the guarded fragment of FO (Jung, Lutz, Goncharov, & Schröder, 2014). The non-standard reasoning tasks of computing least common subsumers and most specific concepts have been studied for an extension of Prob-$\mathcal{EL}^{01}$ (Ecke, Peñaloza, & Turhan, 2014). Moreover, the task of reading off a set of concept inclusions from a given probabilistic interpretation has been investigated for a variant of Prob-$\mathcal{ALC}$ (Kriegel, 2015). Finally, Prob-$\mathcal{ALC}$ has been used in decision making to model the background knowledge of the decision maker (Acar, 2014).

2. Type 2 Probabilistic DLs

We introduce the syntax and semantics of the probabilistic DLs studied in this paper, give modelling examples, and make some initial technical observations.

2.1 Concepts, TBoxes, ABoxes

For the rest of the paper, fix countably infinite sets $\mathsf{N_C}$, $\mathsf{N_R}$, and $\mathsf{N_I}$ of *concept names*, *role names*, and *individual names*, respectively. The probabilistic DL Prob-$\mathcal{ALC}$ is obtained from the classical DL $\mathcal{ALC}$ by adding a probabilistic concept constructor. More precisely, *Prob-$\mathcal{ALC}$ concepts* are formed according to the syntax rule[2]

$$C, D ::= A \mid \neg C \mid C \sqcap D \mid \exists r.C \mid P_{\geq n}C$$

where A ranges over $\mathsf{N_C}$, C and D over concepts, r over $\mathsf{N_R}$, and n over rational numbers from the interval $[0, 1]$. Intuitively, the probability restriction $P_{\geq n}C$ describes the class of

2. We use a different and arguably more natural naming scheme than in the predecessor conference paper, where what we call Prob-$\mathcal{ALC}$ here would be called Prob-$\mathcal{ALC}_c$.

objects that are believed to belong to the class C with probability (degree of belief) at least n. We use the usual abbreviations $C \sqcup D$ for $\neg(\neg C \sqcap \neg D)$, $\forall r.C$ for $\neg\exists r.\neg C$, $\top$ for $A \sqcup \neg A$, and $\bot$ for $\neg\top$. We also introduce several additional abbreviations that are related to the probabilistic concept constructor. In particular, $P_{<n}C$ abbreviates $\neg P_{\geq n}C$, $P_{\leq n}C$ abbreviates $P_{\geq 1-n}\neg C$, $P_{>n}C$ abbreviates $P_{<1-n}\neg C$, and $P_{=n}C$ abbreviates $P_{\geq n}C \sqcap P_{\leq n}C$. The only reason for using rational numbers instead of real numbers in the definition of Prob-$\mathcal{ALC}$ is that the former can be finitely represented; an alternative would be to use a finitely representable subset of the real numbers such as the algebraic reals. All constructors and their names are listed in Table 2, including abbreviations.

The probabilistic DL Prob-$\mathcal{EL}$ is obtained from the classical DL $\mathcal{EL}$ by adding the probability restrictions that can be defined in Prob-$\mathcal{ALC}$. Prob-$\mathcal{EL}$ is thus closely related to the fragment of Prob-$\mathcal{ALC}$ in which negation is disallowed, but it is not identical to that fragment because some abbreviations that rely on negation are now included as primitive constructors (namely $\top$ and all probability restrictions). Formally, *Prob-$\mathcal{EL}$ concepts* are thus formed according to the syntax rule

$$C, D ::= \top \mid A \mid C \sqcap D \mid \exists r.C \mid P_{\sim n}C$$

where $\sim$ ranges over $\{\leq, <, =, >, \geq\}$ and all other symbols range are as above.

In DLs, TBoxes are used to formalize an ontology, and ABoxes store instance data. Both in Prob-$\mathcal{ALC}$ and in Prob-$\mathcal{EL}$, a *TBox* is a finite set of *concept inclusions (CIs)* $C \sqsubseteq D$ where C and D are concepts. We sometimes write $C \equiv D$ to abbreviate $C \sqsubseteq D$ and $D \sqsubseteq C$. In some parts of the paper, we will also consider a more restricted form of TBoxes that we call a *classical TBox*. Formally, a classical TBox $\mathcal{T}$ is a finite set of *concept definitions* $A \equiv D$ where A is a concept name and D is a concept such that every concept name appears at most once on the left-hand side of a definition in $\mathcal{T}$. Sometimes, the additional assumption is made that definitions must be acyclic, but we do not require that here. Note that TBoxes do not have any probabilistic features. It would be possible to consider also TBox statements of the form $P_{\sim n}(C \sqsubseteq D)$ expressing the degree of belief that the inclusion $C \sqsubseteq D$ holds with probability $\sim n$. Though this kind of expressive power might be useful for some applications, we do not consider it in the present paper to keep the number of logics under study manageable.

We next introduce ABoxes. In contrast to TBoxes, we deviate from the classical case and add probabilistic features. This is important for applications in which the data comes from noisy or untrusted data sources. A *Prob-$\mathcal{ALC}$ ABox* is an expression formed according to the rule

$$\mathcal{A} ::= C(a) \mid r(a, b) \mid \neg\mathcal{A} \mid \mathcal{A} \wedge \mathcal{A}' \mid P_{\geq n}\mathcal{A}$$

where C, r, and n have the same ranges as above, a, b range over $\mathsf{N_I}$, and $\mathcal{A}, \mathcal{A}'$ range over Prob-$\mathcal{ALC}$ ABoxes. Abbreviations $P_{\sim n}\mathcal{A}$ for $\sim \in \{\leq, <, =, >\}$ are defined in the same way as for concepts. A *knowledge base (KB)* is a pair $\mathcal{K} = (\mathcal{T}, \mathcal{A})$ with $\mathcal{T}$ a TBox and $\mathcal{A}$ an ABox. Note that, in non-probabilistic DLs, ABoxes are normally sets of assertions $C(a)$ and $r(a, b)$. In our definition, we add conjunction and negation as ABox constructors in order to support, for example, speaking about joint probabilities of classical ABox assertions, as in $P_{\geq 0.1}(r(a, b) \wedge B(b))$. It will become clear when the semantics is defined that we do not assume probabilistic independence of ABox assertions, and thus joint probabilities cannot

Name	Syntax	Semantics
concept name	A	$A^{\mathcal{I},w}$
role name	r	$r^{\mathcal{I},w}$
top	$\top$	$\Delta^{\mathcal{I}}$
bottom	$\bot$	$\emptyset$
negation	$\neg C$	$\Delta^{\mathcal{I}} \setminus C^{\mathcal{I},w}$
conjunction	$C \sqcap D$	$C^{\mathcal{I},w} \cap D^{\mathcal{I},w}$
disjunction	$C \sqcup D$	$C^{\mathcal{I},w} \cup D^{\mathcal{I},w}$
existential restriction	$\exists r.C$	$\{d \in \Delta^{\mathcal{I}} \mid \exists e \in C^{\mathcal{I},w} \text{ with } (d,e) \in r^{\mathcal{I},w}\}$
value restriction	$\forall r.C$	$\{d \in \Delta^{\mathcal{I}} \mid \forall e \in \Delta^{\mathcal{I}} : (d,e) \in r^{\mathcal{I},w} \Rightarrow e \in C^{\mathcal{I},w}\}$
probability restriction	$P_{\geq n}C$	$\{d \in \Delta^{\mathcal{I}} \mid p^{\mathcal{I}}_d(C) \geq n\}$

Table 2: Syntax and semantics of Prob-$\mathcal{ALC}$ concept constructors

just be calculated from atomic ones. In Prob-$\mathcal{EL}$, we want to disallow negation also in ABoxes. Thus, a *Prob-$\mathcal{EL}$ ABox* is an expression formed according to the rule

$$\mathcal{A} ::= C(a) \mid r(a,b) \mid \mathcal{A} \wedge \mathcal{A}' \mid P_{\sim n}\mathcal{A}$$

where all symbol ranges are as expected. We will generally drop the 'Prob-$\mathcal{ALC}$' and 'Prob-$\mathcal{EL}$' qualifications and speak only of an ABox as it shall always be clear from the context which version we refer to. We will use $\mathsf{Ind}(\mathcal{A})$ to denote the set of individual names that occur in the ABox $\mathcal{A}$.

For any syntactic object O such as a concept, a TBox, and an ABox, we use $|O|$ to denote the number of symbols needed to write O where concept names, role names, and individual names count as one and probability values are encoded in binary.

2.2 Semantics

The semantics of classical DLs such as $\mathcal{ALC}$ is based on interpretations $\mathcal{I} = (\Delta^{\mathcal{I}}, \cdot^{\mathcal{I}})$, where $\Delta^{\mathcal{I}}$ is a non-empty set called the *domain* and $\cdot^{\mathcal{I}}$ is an *interpretation function* that maps each $A \in \mathsf{N_C}$ to a subset $A^{\mathcal{I}} \subseteq \Delta^{\mathcal{I}}$, each $r \in \mathsf{N_R}$ to a subset $r^{\mathcal{I}} \subseteq \Delta^{\mathcal{I}} \times \Delta^{\mathcal{I}}$, and each $a \in \mathsf{N_I}$ to an element $a^{\mathcal{I}} \in \Delta^{\mathcal{I}}$.[3] We refer the reader to (Baader et al., 2003) for more information about the classical case. To provide a semantics for Type 2 probabilistic DLs, we generalize such classical interpretations to probabilistic interpretations, in analogy to Halpern's generalization of FO structures to Type 2 probabilistic FO structures (Halpern, 1990). As explained in detail in the introduction, this type of semantics is highly appropriate for representing the uncertain aspects of concepts from the application domain.

A *probabilistic interpretation* takes the form

$$\mathcal{I} = (\Delta^{\mathcal{I}}, W, (\mathcal{I}_w)_{w \in W}, \mu)$$

3. The unique name assumption is not made, but would not change any of the results in this paper.

where $\Delta^\mathcal{I}$ is the (non-empty) domain, W a non-empty set of *possible worlds*, μ a discrete probability distribution on W (i.e., one that is defined on all subsets of W, and hence has countable support), and for each $w \in W$, $\mathcal{I}_w$ is a classical DL interpretation with domain $\Delta^\mathcal{I}$ such that $a^{\mathcal{I}_w} = a^{\mathcal{I}_{w'}}$ for all $a \in \mathsf{N_I}$ and $w, w' \in W$. Since $a^{\mathcal{I}_w}$ does not depend on w, we write only $a^\mathcal{I}$. Given a probabilistic interpretation $\mathcal{I}$, we define the following simultaneously by mutual recursion:

1. the *extension* $C^{\mathcal{I},w}$ or $C^{\mathcal{I}_w}$ of a concept C in a world w, and
2. the *probability* $p^\mathcal{I}_d(C)$ of $d \in \Delta^\mathcal{I}$ being a C.

While the recursive cases for Point 1 are given in Table 2, the case for Point 2 is as follows:

$$p^\mathcal{I}_d(C) = \mu(\{w \in W \mid d \in C^{\mathcal{I},w}\}).$$

For the reader's convenience, Table 2 lists the semantics also of the concept constructors that were introduced only as abbreviations.

Via the probability function $p^\mathcal{I}$, a probabilistic interpretation assigns a probabilistic semantics to concepts. The elements of W are the worlds that are considered possible and $\mu(w)$ is the subjective belief in w being identical to the actual world. By definition, $p^\mathcal{I}_d(C)$ thus expresses the subjective belief in d being an instance of C, which replaces the two definitive cases for classical DL interpretations $\mathcal{I}$ where d either is an instance of C (when $d \in C^\mathcal{I}$) or d is not an instance of C (when $d \notin C^\mathcal{I}$).

The semantics of TBoxes is defined as in the classical case, but taking into account the fact that we have several possible worlds (which should all respect the domain knowledge in the TBox). A probabilistic interpretation $\mathcal{I}$ *satisfies* a concept inclusion $C \sqsubseteq D$ (written $\mathcal{I} \models C \sqsubseteq D$) if $C^{\mathcal{I},w} \subseteq D^{\mathcal{I},w}$ for every world w of $\mathcal{I}$. It is a *model* of a TBox $\mathcal{T}$ if it satisfies all concept inclusions in $\mathcal{T}$.

Let $\mathcal{I}$ be a probabilistic interpretation. To give a semantics to ABoxes, we use mutual recursion to define the following:

1. a relation "$\models$" that defines when a world w of $\mathcal{I}$ *satisfies* $\mathcal{A}$ (written $\mathcal{I}, w \models \mathcal{A}$) and
2. the *probability* $p^\mathcal{I}(\mathcal{A})$ that $\mathcal{A}$ is satisfied in a world w.

The recursive cases for Point 1 are given in Table 3 and the case for Point 2 is

$$p^\mathcal{I}(\mathcal{A}) = \mu(\{w \in W \mid \mathcal{I}, w \models \mathcal{A}\}).$$

Note that the syntactically distinct ABox assertions $(P_{\sim n}C)(a)$ and $P_{\sim n}(C(a))$ are semantically equivalent, so that we shall not distinguish them in the sequel. This allows us to simply write $P_{\sim n}C(a)$. We say that $\mathcal{I}$ is a *model* of $\mathcal{A}$ if $\mathcal{I}, w \models \mathcal{A}$ for some world w. It is a *model* of a knowledge base $\mathcal{K} = (\mathcal{T}, \mathcal{A})$ if it is a model of both $\mathcal{T}$ and $\mathcal{A}$.

Note that, while a TBox $\mathcal{T}$ must be satisfied in *all* worlds of a probabilistic interpretation $\mathcal{I}$ for $\mathcal{I}$ to be a model of $\mathcal{T}$, for an ABox $\mathcal{A}$ it suffices to be satisfied in only a *single* world. We believe that this is a natural generalization of the classical semantics. In fact, the most basic decision problem for knowledge bases is consistency, the question whether a given KB has a model. This can be viewed as a satisfiability problem and, from this perspective, our definition of a model of an ABox is in line with Halpern's definition of satisfiability in probabilistic FO. To express that an ABox $\mathcal{A}$ is true with certainty, we can simply write $P_{=1}\mathcal{A}$.

$\mathcal{I}, w \models C(a)$	iff	$a^{\mathcal{I}} \in C^{\mathcal{I},w}$
$\mathcal{I}, w \models r(a,b)$	iff	$(a^{\mathcal{I}}, b^{\mathcal{I}}) \in r^{\mathcal{I},w}$
$\mathcal{I}, w \models \neg\mathcal{A}$	iff	$\mathcal{I}, w \not\models \mathcal{A}$
$\mathcal{I}, w \models \mathcal{A} \wedge \mathcal{A}'$	iff	$\mathcal{I}, w \models \mathcal{A}$ and $\mathcal{I}, w \models \mathcal{A}'$
$\mathcal{I}, w \models P_{\geq n}(\mathcal{A})$	iff	$p^{\mathcal{I}}(\mathcal{A}) \geq n$

Table 3: Semantics of ABoxes

2.3 Examples

We illustrate how Type 2 probabilistic DLs can be used to model medical knowledge, revisiting some of the example classes from SNOMED CT mentioned in the introduction. For example, the class '*probable tubo-ovarian abscess*' (Bodenreider, Smith, & Burgun, 2004) can be modelled in Prob-$\mathcal{ALC}$ as a concept of the form

$$P_{\geq\alpha}\mathsf{TuboOvarianAbscess}$$

which describes findings that are a tubo-ovarian abscess with probability at least α. As a concrete value for α, one would choose a suitable threshold such as 0.75 or 0.9. We note that this concept is not subsumed by $\mathsf{TuboOvarianAbscess}$, reflecting that a probable tubo-ovarian abscess needs not actually be a tubo-ovarian abscess. If we focus on patients instead of on findings, we could use the concept

$$\exists\mathsf{hasAbnormality}.P_{\geq\alpha}\mathsf{TuboOvarianAbscess},$$

to describe patients who have an abnormality (e.g. a sonographic irregularity) that is a tubo-ovarian abscess with probability at least α, and the concept

$$P_{\geq\alpha}\exists\mathsf{hasAbnormality}.\mathsf{TuboOvarianAbscess},$$

to describe patients who have a tubo-ovarian abscess with probability at least α and otherwise do not necessarily have any abnormality. It is interesting to note that uncertain diagnoses may well have definite consequences. For example, Lyme disease is typically treated with antibiotics even when the diagnosis is not entirely certain, due to the combination of the graveness of the disease and the difficulty of diagnosing it with certainty. This is modelled by the following concept inclusion:

$$P_{\geq 0.8}\exists\mathsf{hasDisease}.\mathsf{LymeDisease} \sqsubseteq \exists\mathsf{recommendedTreatment}.\mathsf{Antibiotic}$$

Taking up an example mentioned by Bodenreider et al. (2004), we can formalize the SNOMED CT concept '*basal cell tumour, uncertain whether benign or malignant*', as

$$\mathsf{BasalCellTumour} \sqcap \exists\mathsf{hasStatus}.(P_{\geq\alpha}\mathsf{Benign} \sqcap P_{\alpha}\mathsf{Malignant}).$$

where α is a threshold probability indicating reasonable possibilities, such as $\alpha = 0.1$. As noted in the introduction, uncertainty of instance data is ubiquitous in medicine. Consider for example a scenario where patient John exhibits fatigue symptoms. Since fatigue is an unspecific symptom of Lyme disease, we may model this by an ABox such as

$$P_{=1}\mathsf{Fatigue}(\mathsf{s1}) \wedge P_{=1}\mathsf{hasSymptom}(\mathsf{John}, \mathsf{s1}) \wedge P_{\geq 0.1}\exists \mathsf{hasCause}.\,\mathsf{LymeDisease}(\mathsf{s1}).$$

If, however, John additionally exhibits unclear fever, which is also unspecifically related to Lyme disease, then one will be more inclined to attribute also John's fatigue to the suspected case of Lyme disease, so we could use an ABox such as

$$\begin{gathered} P_{=1}\mathsf{Fatigue}(\mathsf{s1}) \wedge P_{=1}\mathsf{hasSymptom}(\mathsf{John}, \mathsf{s1}) \wedge \\ P_{=1}\mathsf{Fever}(\mathsf{s2}) \wedge P_{=1}\mathsf{hasSymptom}(\mathsf{John}, \mathsf{s2}) \wedge \\ P_{\geq 0.2}(C(\mathsf{s1})) \wedge P_{\geq 0.3}(C(\mathsf{s2})) \wedge P_{\geq 0.15}(C(\mathsf{s1}) \sqcap C(\mathsf{s2})) \end{gathered}$$

where, for readability, we abbreviate $C = \exists \mathsf{hasCause}.\,\mathsf{LymeDisease}$. Note that the ABox additionally indicates that the attributions of the two symptoms to Lyme disease are not independent: the lower bound on their joint probability is higher than the product of the lower bounds on the individual probabilities.

The standard translation from $\mathcal{ALC}$ to FO, as described for example in (Baader et al., 2003), can be extended to a translation from Prob-$\mathcal{ALC}$ to Type 2 probabilistic FO in a straightforward way. Type 2 probabilistic FO is interpreted over structures that strongly resemble our probabilistic interpretations, except that worlds index FO structures instead of DL interpretations. The syntax of Type 2 probabilistic FO admits real-valued terms of the form $w(\phi)$ where ϕ is a formula, denoting the subjective probability that ϕ is satisfied (by the given valuation). Concepts $P_{\geq n}C$ are then translated to the probabilistic FO formula $w(C^{\#}(x)) \geq n$, where $C^{\#}(x)$ is the standard translation of C with free variable x. For example, the concept related to basal cell tumours above translates to

$$\mathsf{BasalCellTumour}(x) \wedge \exists y\,(\mathsf{hasStatus}(x, y) \wedge w(\mathsf{Benign}(y)) \geq \alpha \wedge w(\mathsf{Malignant}(y)) \geq \alpha).$$

The translation of TBoxes and probabilistic ABoxes is just as simple and left to the reader.

2.4 Reasoning Tasks

A concept C is *satisfiable w.r.t. a TBox* $\mathcal{T}$ if there is a model $\mathcal{I}$ of $\mathcal{T}$ and a world w such that $C^{\mathcal{I},w} \neq \emptyset$. Note that this is in line with the definition of a model of an ABox, where we also demand that the ABox is satisfied in at least one world. A concept C is *subsumed by a concept* D *w.r.t. a TBox* $\mathcal{T}$, denoted $\mathcal{T} \models C \sqsubseteq D$, if for all models $\mathcal{I}$ of $\mathcal{T}$ and worlds w, we have $C^{\mathcal{I},w} \subseteq D^{\mathcal{I},w}$. A knowledge base $\mathcal{K}$ is *consistent* if it has a model. An ABox $\mathcal{A}'$ is a *consequence* of a knowledge base $\mathcal{K} = (\mathcal{T}, \mathcal{A})$, written $\mathcal{K} \models \mathcal{A}'$, if for every model $\mathcal{I}$ of $\mathcal{K}$ and world w with $\mathcal{I}, w \models \mathcal{A}$, we have $\mathcal{I}, w \models \mathcal{A}'$.

The four notions just introduced give rise to the decision problems *concept satisfiability*, *concept subsumption*, *KB consistency*, and *KB consequence* in an obvious way. For example, *KB consequence* is the problem to decide, given as input a knowledge base $\mathcal{K}$ and an ABox $\mathcal{A}'$, whether $\mathcal{K} \models \mathcal{A}'$. On first sight, KB consequence might seem to be a slightly unusual reasoning problem. In classical DLs, it is more common to consider instance checking, which is the restriction of KB consequence to ABoxes of the form $C(a)$. Given that we allow probabilistic and Boolean constructors in the ABox, though, KB consequence is a natural generalization of instance checking.

In Prob-$\mathcal{ALC}$ (and its extensions introduced below), concept satisfiability and concept non-subsumption can be reduced to each other in polynomial time and the same is true for KB consistency and KB non-consequence. For example, we have $\mathcal{K} \models \mathcal{A}'$ where $\mathcal{K} = (\mathcal{T}, \mathcal{A})$ if the KB $(\mathcal{T}, \mathcal{A} \wedge \neg\mathcal{A}')$ is inconsistent. Moreover, concept satisfiability and subsumption can be reduced in polynomial time to KB consistency and KB consequence. For Prob-$\mathcal{ALC}$, we will thus typically state lower complexity bounds for concept satisfiability and upper bounds for KB consistency to achieve maximum generality.

In classical $\mathcal{EL}$, concept satisfiability and KB consistency are trivial since every concept is satisfiable w.r.t. every TBox and every KB is consistent. This is not the case in Prob-$\mathcal{EL}$, even if we use only a single probabilistic operator; e.g. the concept $P_{\leq 0.5}A \sqcap P_{\leq 0.5}P_{\leq 0.5}A$ is unsatisfiable. As a consequence, concept subsumption and concept satisfiability are again mutually reducible in polynomial time: $\mathcal{T} \models C \sqsubseteq D$ iff $C \sqcap A$ is unsatisfiable w.r.t. $\mathcal{T} \cup \{A \sqcap D \sqsubseteq \bot\}$ where A is a fresh concept name and $\bot$ abbreviates an unsatisfiable concept, and C is satisfiable w.r.t. $\mathcal{T}$ iff $\mathcal{T} \not\models C \sqsubseteq \bot$. Throughout the paper, though, we will often restrict the available comparisons $\sim \in \{\leq, <, =, >, \geq\}$ and concrete probability values in Prob-$\mathcal{EL}$; for example, Prob-$\mathcal{EL}^{01}$ is the fragment of Prob-$\mathcal{EL}$ that admits only the probability restrictions $P_{>0}A$ and $P_{=1}A$. This fragment behaves like classical $\mathcal{EL}$ in the sense that concept satisfiability and KB consistency are trivial.

2.5 Language Extensions

We introduce and discuss several natural extensions of the basic probabilistic DLs Prob-$\mathcal{ALC}$ and Prob-$\mathcal{EL}$ that are studied in this paper. All extensions can easily be translated to Type 2 probabilistic FO.

2.5.1 Probabilistic Roles

Instead of restricting the application of probabilistic constructors to concepts, one can also allow their application to roles. In the corresponding extension *Prob-*$\mathcal{ALC}_r$ of Prob-$\mathcal{ALC}$, the concept formation rule has the additional case $\exists P_{\sim n} r.C$ where $\sim$ ranges over $\{\leq, <, =, >, \geq\}$. The concept $\exists P_{\sim n} r.C$ combines an existential restriction with the application of the probabilistic constructor $P_{\sim n}$ to the role name r. Similarly as for probabilistic concepts, we use $\forall P_{\sim n} r.C$ as an abbreviation for $\neg\exists P_{\sim n} r.\neg C$. In contrast to the case of probabilistic concepts, it is not possible to define the other incarnations of $\exists P_{\sim n} r.C$ in terms of $\exists P_{\geq n} r.C$ and thus we introduce five primitive constructors, one for each $\sim \in \{\leq, <, =, >, \geq\}$.

To give a semantics for *Prob-*$\mathcal{ALC}_r$ concepts, we define the following simultaneously by mutual recursion:

1. the *extension* $C^{\mathcal{I},w}$ or $C^{\mathcal{I}_w}$ of a concept C in a world w,
2. the *probability* $p_d^{\mathcal{I}}(C)$ of $d \in \Delta^{\mathcal{I}}$ being a C, and
3. the *probability* $p_{d,e}^{\mathcal{I}}(r)$ that $d, e \in \Delta^{\mathcal{I}}$ are related by r.

The definitions for Points 1 and 2 are as for Prob-$\mathcal{ALC}$ except that, in Point 1, the semantics of the additional concept constructors is given by the following clause:

$$(\exists P_{\sim n} r.C)^{\mathcal{I},w} = \{d \in \Delta^{\mathcal{I}} \mid \exists e \in C^{\mathcal{I},w} : p^{\mathcal{I}}_{d,e}(r) \sim n\}$$

For Point 3, we set

$$p^{\mathcal{I}}_{d,e}(r) = \mu(\{w \in W \mid (d,e) \in r^{\mathcal{I},w}\}).$$

To give an example for Prob-$\mathcal{ALC}_r$, consider the SNOMED CT class '*natural death with probable cause suspected*'. We interpret this class as expressing two properties: on the one hand, there is at least one cause that is considered probable, and on the other hand, no cause is certain. Using probabilistic roles, we can model this as the concept

$$\mathsf{NaturalDeath} \sqcap \exists P_{\geq\alpha}\mathsf{hasCause}.\top \sqcap \neg\exists P_{\geq\beta}\mathsf{hasCause}.\top$$

expressing that there is some (unspecified) phenomenon which is believed to be the cause of death with probability at least α, but nothing is believed to be the cause of death with probability more than β, where $\alpha > 0$ and $\beta < 1$ are suitable chosen thresholds that satisfy $\alpha < \beta$.

2.5.2 Linear and Polynomial Inequalities as Concepts and ABoxes

In contrast to Prob-$\mathcal{ALC}$ as introduced above, Type 2 probabilistic FO admits the formation of unrestricted linear and polynomial inequalities over probabilities. This inspires the extensions Prob-$\mathcal{ALC}^{\mathsf{lineq}}$ and Prob-$\mathcal{ALC}^{\mathsf{polyeq}}$ of Prob-$\mathcal{ALC}$, in which (i) concepts $P_{\geq n}C$ are replaced with linear (resp. polynomial) inequalities $\mathcal{E}$ over expressions $P(C)$, C a concept, called *concept inequalities*; and (ii) ABoxes $P_{\geq n}(\mathcal{A})$ are replaced with linear (resp. polynomial) inequalities $\mathcal{E}$ over expressions $P(\mathcal{A})$, $\mathcal{A}$ an ABox, called *ABox inequalities*. The semantics of a concept inequality $\mathcal{E}$ is that $\mathcal{E}^{\mathcal{I},w}$ contains precisely those $d \in \Delta^{\mathcal{I}}$ such that the inequality $\mathcal{E}$ is satisfied when each $P(C)$ in it is replaced by $p^{\mathcal{I}}_d(C)$. The semantics of ABox inequalities is defined analogously. In all inequalities, we allow strict comparisons via "$<$" as well as non-strict comparisons via "$\leq$". In extension of our earlier policy, we regard an ABox assertion of the form $\mathcal{E}(a)$ with $\mathcal{E}$ a concept inequality as an ABox inequality when convenient. For example, $P(A) > P(B)(a)$ is regarded as $P(A(a)) > P(B(a))$.

To illustrate the use of linear inequalities, we make a brief detour to *qualitative* reasoning about probabilities. There are a number of different proposals, of which we consider two. First, Gärdenfors (1975) considers a logic with a binary operator 'more probable than'. This can be captured in Prob-$\mathcal{ALC}^{\mathsf{lineq}}$ since 'more probably being a C than a D' corresponds to the linear concept inequality $P(C) > P(D)$. Second, Herzig (2003) proposes an operator 'more probably than not', i.e., the probability of an event is higher than its complement. In Prob-$\mathcal{ALC}^{\mathsf{lineq}}$, this can be expressed as $P(C) > P(\neg C)$. We note that, in general, there is no consensus as to whether subjective probabilities in medical ontologies should be represented quantitatively as in the previous section or qualitatively, e.g., by defining a 'probable tubo-ovarian abscess' as

$$P(\mathsf{TuboOvarianAbscess}) > c \cdot P(\neg\mathsf{TuboOvarianAbscess})$$

for some constant c (a generalization of Herzig's approach). It may depend on the concrete application which modelling is more appropriate; we note that Prob-$\mathcal{ALC}^{\mathsf{lineq}}$ and Prob-$\mathcal{ALC}^{\mathsf{polyeq}}$ support both views.

Linear inequalities can also be used to express conditional probabilities. For example, a concept constructor for conditional probabilities $P_{\sim n}(C|D)$, with $\sim \in \{\leq, <, =, >, \geq\}$, can be expressed as the linear concept inequality $P(C \sqcap D) \sim n \cdot p(D)$.

Polynomial inequalities are strictly more expressive than linear ones. In particular, they capture several kinds of independence constraints, as discussed in the following.

2.5.3 Independence Constraints

The semantics of Prob-$\mathcal{ALC}$ has no built-in independence of probabilistic events. For example, although one might expect that the ABox

$$P_{\geq 0.5}\mathsf{NearSighted}(\mathsf{John}) \wedge P_{\geq 0.5}\mathsf{Vegetarian}(\mathsf{John})$$

has the consequence $P_{\geq 0.25}(\mathsf{NearSighted} \sqcap \mathsf{Vegetarian})(\mathsf{John})$ since the two involved events are intuitively independent, it is not hard to see that the semantics of Prob-$\mathcal{ALC}$ does not support such a deduction. Polynomial inequalities make it possible to add independence constraints. For example, we can obtain the desired consequence in the example above when we add the TBox statement $\top \sqsubseteq C$, where C is the polynomial concept inequality

$$P(\mathsf{NearSighted} \sqcap \mathsf{Vegetarian}) = P(\mathsf{NearSighted}) \cdot P(\mathsf{Vegetarian}).$$

For use in lower bounds and other negative results, we single out a particular and very simple form of independence constraint as follows.

We use Prob-$\mathcal{ALC}^{\mathsf{indep}}$ to denote the extension of Prob-$\mathcal{ALC}$ obtained by admitting *binary independence constraints* of the form $\mathsf{indep}(C, D)$ in the TBox, with C, D concepts. An interpretation $\mathcal{I}$ *satisfies* $\mathsf{indep}(C, D)$ if for all $d \in \Delta^{\mathcal{I}}$, we have $p_d^{\mathcal{I}}(C) \cdot p_d^{\mathcal{I}}(D) = p_d^{\mathcal{I}}(C \sqcap D)$.

Note that these independence constraints are strictly weaker than polynomial inequalities. For example, the latter also allow to express independence of more than two events.

2.6 Worlds with Probability Zero

Probabilistic interpretations may contain worlds with probability 0 which represent situations that are infinitely improbable, but not impossible per se. The presence or absence of such worlds has an impact on the reasoning problems studied in this paper; for example, the concept $A \sqcap P_{=1}\neg A$ is satisfiable w.r.t. the empty TBox when worlds of probability zero are admitted, but not otherwise. It depends on the intended application whether or not the presence of probability 0 worlds is desired or not. For computational purposes, however, it is more convenient to assume that all worlds have positive probability. We show that this can be done without loss of generality.

A probabilistic interpretation $\mathcal{I}$ is *positive* if $\mu(w) > 0$ for all worlds w of $\mathcal{I}$. This gives rise to positive versions of the introduced reasoning problems concept satisfiability, concept subsumption, KB consistency, and KB consequence in a straightforward way. For example, *positive KB consequence* is the problem to decide, given a KB $\mathcal{K} = (\mathcal{T}, \mathcal{A})$ and an ABox

$\mathcal{A}'$, whether $\mathcal{K} \models^+ \mathcal{A}$, that is, whether for every positive model $\mathcal{I}$ of $\mathcal{K}$ and world w with $\mathcal{I}, w \models \mathcal{A}$, we have $\mathcal{I}, w \models \mathcal{A}'$.

For all of the reasoning problems and all of the probabilistic DLs considered in this paper, the positive version can be reduced in polynomial time to the unrestricted version and vice versa. As an example, we consider KB consequence in Prob-$\mathcal{ALC}$.

Lemma 1. *In Prob-$\mathcal{ALC}$, KB consequence and positive KB consequence are polynomially reducible to each other.*

Proof. The reduction of positive KB consequence to KB consequence is by the observation that $\mathcal{K} = (\mathcal{T}, \mathcal{A}) \models^+ \mathcal{A}'$ iff $(\mathcal{T}, P_{>0}(\mathcal{A} \wedge A(a))) \models P_{>0}(\mathcal{A}' \wedge A(a))$ where A is a fresh concept name and a a fresh invidual name.

"$\Rightarrow$". Let $\mathcal{I}$ be a model of $\mathcal{T}$ with $\mathcal{I}, w \models P_{>0}(\mathcal{A} \wedge A(a))$. Then there is a world v with $\mathcal{I}, v \models \mathcal{A} \wedge A(a)$ and $\mu(v) > 0$. Drop all worlds with probability zero from $\mathcal{I}$ and call the result $\mathcal{I}'$. An easy induction shows that for all Prob-$\mathcal{ALC}$ concepts C and worlds w in $\mathcal{I}'$, we have $C^{\mathcal{I},w} = C^{\mathcal{I}',w}$. In particular, the extensions of probability restrictions do not change since by the semantics they only depend on world of positive probability. Consequently, $\mathcal{I}'$ is a positive model of $\mathcal{T}$ and $\mathcal{I}', v \models \mathcal{A}$. From the assumption that $\mathcal{K} = (\mathcal{T}, \mathcal{A}) \models^+ \mathcal{A}' \wedge A(a)$, we obtain $\mathcal{I}', v \models \mathcal{A}'$, thus $\mathcal{I}, v \models \mathcal{A}' \wedge A(a)$. It follows that $\mathcal{I}, w \models P_{>0}(\mathcal{A}' \wedge A(a))$.

"$\Leftarrow$". Let $\mathcal{I}$ be a positive model of $\mathcal{T}$ with $\mathcal{I}, w \models \mathcal{A}$. Extend $\mathcal{I}$ to interpret also A by taking $A^{\mathcal{I},w} = \{a^{\mathcal{I}}\}$, and $A^{\mathcal{I},v} = \emptyset$ for $v \neq w$. Then $\mathcal{I}, w \models \mathcal{A} \wedge A(a)$, and hence $\mathcal{I}, w \models P_{>0}(\mathcal{A} \wedge A(a))$, so by assumption $\mathcal{I}, w \models P_{>0}(\mathcal{A}' \wedge A(a))$, which by the way we defined $A^{\mathcal{I},w}$ implies $\mathcal{I}, w \models \mathcal{A}'$.

For the reduction of KB consequence to positive KB consequence, let $\mathcal{K} = (\mathcal{T}, \mathcal{A})$ be a knowledge base and $\mathcal{A}'$ an ABox. Pick a fresh concept name $\widehat{A}$ for every concept name A that occurs in $\mathcal{K}$ or $\mathcal{A}'$ and a fresh role name $\widehat{r}$ for every role name r that occurs in $\mathcal{K}$ or $\mathcal{A}'$. For a concept C, let $\widehat{C}$ denote the concept obtained by replacing every occurrence of a concept name A and role name r that is outside the scope of any $P_{\sim n}$ with $\widehat{A}$ and $\widehat{r}$, where the scope of $P_{\sim n}$ is D in $P_{\sim n}D$, and r in $\exists P_{\sim n} r.D$. This replacement operation is lifted to TBoxes and ABoxes in the obvious way. Now put $\widehat{\mathcal{K}} = (\mathcal{T} \cup \widehat{\mathcal{T}}, \widehat{\mathcal{A}})$. We show that $\mathcal{K} \models \mathcal{A}'$ iff $\widehat{\mathcal{K}} \models^+ \widehat{\mathcal{A}}'$.

"$\Leftarrow$". Let $\mathcal{I}$ be a model of $\mathcal{T}$ with $\mathcal{I}, w \models \mathcal{A}$. Construct the interpretation $\widehat{\mathcal{I}}$ by dropping all worlds of probability zero from $\mathcal{I}$ and interpreting the new symbols as $\widehat{A}^{\mathcal{I},v} = A^{\mathcal{I},w}$ and $\widehat{r}^{\mathcal{I},v} = r^{\mathcal{I},w}$ for every world v. One shows by an easy induction that for all concepts C and worlds v of $\widehat{\mathcal{I}}$, we have (i) $C^{\widehat{\mathcal{I}},v} = C^{\mathcal{I},v}$ and (ii) $\widehat{C}^{\widehat{\mathcal{I}},v} = C^{\mathcal{I},w}$. Thus, $\widehat{\mathcal{I}}$ is a positive model of $\mathcal{T} \cup \widehat{\mathcal{T}}$ and by (ii), $\widehat{\mathcal{I}}, v \models \widehat{\mathcal{A}}$ for every world v of $\widehat{\mathcal{I}}$. By assumption, we thus have $\widehat{\mathcal{I}}, v \models \widehat{\mathcal{A}}'$ for every v. From (ii), we get $\mathcal{I}, w \models \mathcal{A}'$ and are done.

"$\Rightarrow$". Let $\widehat{\mathcal{I}}$ be a positive model of $\mathcal{T} \cup \widehat{\mathcal{T}}$ with $\widehat{\mathcal{I}}, w \models \widehat{\mathcal{A}}$. Construct an interpretation $\mathcal{I}$ by adding a new world w_0 with probability zero, and put $A^{\mathcal{I},w_0} = \widehat{A}^{\widehat{\mathcal{I}},w}$, $r^{\mathcal{I},w_0} = \widehat{r}^{\widehat{\mathcal{I}},w}$ for each concept name A and each role name r occurring in $\mathcal{K}$ or $\mathcal{A}'$. By a straightforward induction, one shows that for all concepts C and worlds v of $\widehat{\mathcal{I}}$, we have (i) $C^{\widehat{\mathcal{I}},v} = C^{\mathcal{I},v}$ and (ii) $C^{\mathcal{I},w_0} = \widehat{C}^{\widehat{\mathcal{I}},w}$. By (i) and (ii), $\mathcal{I}$ is a model of $\mathcal{T}$ and by (ii), $\mathcal{I}, w_0 \models \mathcal{A}$. By assumption, $\mathcal{I}, w_0 \models \mathcal{A}'$, which by (ii) implies $\widehat{\mathcal{I}}, w \models \widehat{\mathcal{A}}'$. $\square$

For the remainder of the paper, we will always assume positive interpretations unless explicitly noted otherwise. Instead of "$\models^+$", from now on we simply write "$\models$"; similarly, by 'consequence' we mean positive consequence, etc.

3. Prob-$\mathcal{ALC}$ and Extensions

We analyse the computational complexity of reasoning with Prob-$\mathcal{ALC}$ and its extensions. We start with showing that, in Prob-$\mathcal{ALC}$, all considered reasoning tasks are ExpTime-complete and briefly observe that the algorithm used to establish the upper bound (for KB consistency) can easily be adapted to its extension Prob-$\mathcal{ALC}^{\mathsf{lineq}}$. We then show that the algorithm can also be adapted to Prob-$\mathcal{ALC}^{\mathsf{polyeq}}$, which requires some less straightforward modifications. We finally study Prob-$\mathcal{ALC}_r$ and show that all reasoning tasks are 2ExpTime-complete when only the probability values zero and one are admitted and undecidable in Prob-$\mathcal{ALC}_r^{\mathsf{indep}}$ and Prob-$\mathcal{ALC}_r^{\mathsf{lineq}}$ (when probability values are unrestricted). Decidability of non-extended Prob-$\mathcal{ALC}_r$ with unrestricted probability values remains open. We also consider some interesting model properties such as the uniform model property and the bounded model property.

3.1 Prob-$\mathcal{ALC}$

ExpTime-hardness of concept satisfiability in Prob-$\mathcal{ALC}$ is inherited from classical $\mathcal{ALC}$. To show that all considered reasoning tasks are ExpTime-complete, it thus suffices to establish an ExpTime upper bound for KB consistency in Prob-$\mathcal{ALC}$. We do this by developing a procedure in the style of Pratt's *type elimination* (Pratt, 1979) which checks for the existence of a decomposed representation of a model that we call a quasimodel. To deal with probabilities, the procedure involves solving systems of linear inequalities.

Let $\mathcal{K} = (\mathcal{T}, \mathcal{A})$ be the knowledge base whose consistency is to be decided. We assume w.l.o.g. that the TBox has the form $\mathcal{T} = \{\top \sqsubseteq C_\mathcal{T}\}$ and use

- $\mathsf{ccl}(\mathcal{K})$ to denote the *concept closure* of $\mathcal{K}$, i.e., the set of all concepts that are used in $\mathcal{K}$ (possibly as a subconcept) and their negations, and
- $\mathsf{acl}(\mathcal{K})$ to denote the *ABox closure* of $\mathcal{K}$, i.e., the set of all sub-ABoxes of $\mathcal{A}$ and all ABoxes of the form $C(a)$ with $C \in \mathsf{ccl}(\mathcal{K})$ and $a \in \mathsf{Ind}(\mathcal{A})$, closed under single negation.

Note that the cardinality of $\mathsf{ccl}(\mathcal{K})$ is bounded by $2|\mathcal{K}|$ and the size of $\mathsf{acl}(\mathcal{K})$ by $4|\mathcal{K}|$. An *ABox type* for $\mathcal{K}$ is a subset $t \subseteq \mathsf{acl}(\mathcal{K})$, which describes a model of an ABox at a single possible world. An ABox type t is required to satisfy the following conditions:

1. $C_\mathcal{T}(a) \in t$ for all $a \in \mathsf{Ind}(\mathcal{A})$;
2. $(\neg C)(a) \in t$ iff $C(a) \notin t$, for all $(\neg C)(a) \in \mathsf{acl}(\mathcal{K})$;
3. $(C \sqcap D)(a) \in t$ iff $C(a), D(a) \in t$, for all $(C \sqcap D)(a) \in \mathsf{acl}(\mathcal{K})$;
4. $\neg\mathcal{A}' \in t$ iff $\mathcal{A}' \notin t$, for all $\neg\mathcal{A}' \in \mathsf{acl}(\mathcal{K})$;
5. $\mathcal{A}' \wedge \mathcal{A}'' \in t$ iff $\mathcal{A}', \mathcal{A}'' \in t$, for all $\mathcal{A}' \wedge \mathcal{A}'' \in \mathsf{acl}(\mathcal{K})$;

6. if $C(b) \in t$ and $r(a,b) \in t$, then $\exists r.C(a) \in t$, for all $\exists r.C(a) \in \mathsf{acl}(\mathcal{T})$.

Let $\mathfrak{A}_\mathcal{K}$ denote the set of all ABox types for $\mathcal{K}$. Note that each interpretation $\mathcal{I}$ and world w of $\mathcal{I}$ realize an ABox type

$$t_w^\mathcal{I} = \{\mathcal{A}' \in \mathsf{acl}(\mathcal{K}) \mid \mathcal{I}, w \models \mathcal{A}'\}.$$

We now define types for elements of probabilistic interpretations that are not from the ABox. For uniform notation, we fix an individual name $a_\varepsilon \in \mathsf{Ind}(\mathcal{A})$ that we use to represent such 'anonymous' elements. An *element type* for $\mathcal{K}$ is a subset $t \subseteq \{C(a_\varepsilon) \mid C \in \mathsf{ccl}(\mathcal{K})\}$ that respects Conditions 1 to 3 in the definition of ABox types. Since element types refer only to the fixed individual name a_ε, we will often drop that name altogether and simply view an element type as a subset of $\mathsf{ccl}(\mathcal{K})$. Let $\mathfrak{T}_\mathcal{K}$ denote the set of all element types for $\mathcal{K}$. Note that each interpretation $\mathcal{I}$, world w of $\mathcal{I}$, and element $d \in \Delta^\mathcal{I}$ realize an element type

$$t_w^\mathcal{I}(d) = \{C \in \mathsf{ccl}(\mathcal{K}) \mid d \in C^{\mathcal{I},w}\}.$$

A *quasimodel* for $\mathcal{K}$ is a pair $Q = (T, T')$ with $T \subseteq \mathfrak{A}_\mathcal{K}$ and $T' \subseteq \mathfrak{T}_\mathcal{K}$. The general idea is that T is the set of ABox types realized in the probabilistic interpretation described by the quasimodel, and T' is the set of element types realized in it. Of course, not every possible pair (T, T') describes the set of types realized in an interpretation. For this reason, we are interested in quasimodels that satisfy two additional conditions, one per dimension. The first condition concerns the DL dimension. A type $t \in T \cup T'$ is *saturated* in Q if for each $\exists r.C(a) \in t$, there exists a $t' \in T'$ with $t' \supseteq \{C\} \cup \{D \mid \forall r.D(a) \in t\}$.

It is more complicated to deal with the probabilistic dimension. To deal with it, it is helpful to think of a probabilistic interpretation $\mathcal{I}$ as consisting of a set of slices. In particular, we refer to the set of ABox types $t_w^\mathcal{I}$, w a world of $\mathcal{I}$, as the *ABox slice* of $\mathcal{I}$. Similarly, for each $d \in \Delta^\mathcal{I}$ we refer to the set of element types $t_w^\mathcal{I}(d)$, w a world of $\mathcal{I}$, as the *d-slice* of $\mathcal{I}$ and as an *element slice* if the precise identitiy of d is not important.

For each $t_0 \in T$, we define a system of linear inequalities $\mathcal{E}(t_0, T)$ over the variables $(x_t)_{t \in T}$ that expresses constraints on the probability x_t of each ABox type $t \in T$ in the ABox slice of a probabilitic interpretation:

1. probabilistic ABoxes in t_0 are satisfied:
   - for each $P_{\geq n}\mathcal{A}' \in t_0$: $\sum_{t \in T \mid \mathcal{A}' \in t} x_t \geq n$
   - for each $\neg P_{\geq n}\mathcal{A}' \in t_0$: $\sum_{t \in T \mid \mathcal{A}' \in t} x_t < n$
2. probabilities sum up to one: $\sum_{t \in T} x_t = 1$
3. all realized types agree on probablistic ABoxes with t_0:
   $x_t = 0$ for all t such that $(t \setminus t_0) \cup (t_0 \setminus t)$ contains an ABox of the form $P_{\geq n}\mathcal{A}'$;
4. t_0 has positive probability: $x_{t_0} > 0$.

For each $t_0 \in T'$, we define a corresponding system of inequalities $\mathcal{E}(t_0, T')$ literally in the same way as for ABox types, only that T is replaced with T'.

One might be tempted to think that, in Point 4 above, we should demand that *all* x_t are positive because we are working with positive interpretations. However, variable x_t being assigned value zero just means that the type t does not actually occur in the described slice. Thus, solving the system of inequalities actually selects for us a subset of T (resp. T') that is realized in the slice; note that this is particularly natural in the context of element slices since it would clearly not be reasonable to demand that every element type from T' occurs in every element slice. Also note that Points 1 and 3 implicitly also deal with the probabilistic concept constructor since we do not distinguish between $(P_{\geq n}C)(a)$ (where such a constructor is used) and the probabilistic ABox $P_{\geq n}(C(a))$.

Let us get back to our quasimodel $Q = (T, T')$ for $\mathcal{K}$. We say that a type $t \in T$ (resp. $t' \in T'$) is *coherent* in Q if the system of inequalities $\mathcal{E}(t, T)$ (resp. $\mathcal{E}(t', T')$) has a non-negative solution over the reals; note that there is a non-negative solution over the reals iff there is such a solution over the rationals (Schrijver, 1986). We call Q *proper* for $\mathcal{K}$ if every $t \in T \cup T'$ is saturated and coherent in Q, and there is a $t \in T$ with $\mathcal{A} \in t$.

Lemma 2. *$\mathcal{K}$ is consistent iff there is a proper quasimodel for $\mathcal{K}$.*

Proof. "⇒". Let $\mathcal{I}, w_\mathcal{A} \models \mathcal{K}$. Define a quasimodel $Q = (T, T')$ by setting

$$T = \{t_w^\mathcal{I} \mid w \in W\} \qquad \text{and} \qquad T' = \{t_w^\mathcal{I}(d) \mid w \in W, d \in \Delta^\mathcal{I}\}.$$

By the definition of Q and the semantics, each $t \in T \cup T'$ is saturated in Q. To establish coherence, we start with ABox types. Thus, let $t \in T$. We want to show that $\mathcal{E}(t, T)$ has a solution, which we write as a map δ assigning values to variables. Choose a $w \in W$ such that $t_w^\mathcal{I} = t$. For each $t' \in T$, define

$$\delta(x_{t'}) = \sum_{w' \in W \mid t_{w'}^\mathcal{I} = t'} \mu(w').$$

Using the definition of δ and the semantics, it can be verified that δ is a solution for $\mathcal{E}(t, T)$. In particular, $\delta(x_t) > 0$ since the corresponding sum contains the positive summand $\mu(w)$. The argument for element types $t \in T'$ is similar: choose $w \in W$ and $d \in \Delta^\mathcal{I}$ such that $t_w^\mathcal{I}(d) = t$, and for each $t' \in T'$, put

$$\delta(x_{t'}) = \sum_{w' \in W \mid t_{w'}(d)^\mathcal{I} = t'} \mu(w').$$

We have $\mathcal{A} \in t_{w_\mathcal{A}}^\mathcal{I}$ and $t_{w_\mathcal{A}}^\mathcal{I} \in T$, thus Q is proper.

"⇐": Let $Q = (T, T')$ be a proper quasimodel for $\mathcal{K}$. For each $t \in T$ (resp. $t \in T'$), the system of linear inequalities $\mathcal{E}(t, T)$ (resp. $\mathcal{E}(t, T')$) has a non-negative rational solution. For each $t \in T \cup T'$, fix such a solution δ_t. For a positive rational number r, we use $\mathsf{den}(r)$ to denote the denominator of the irreducible form of r (in particular, $\mathsf{den}(0) = 1$). Put

$$c := 1 / \prod_{(t,t') \in (T \times T) \cup (T' \times T')} \mathsf{den}(\delta_t(x_{t'}))$$

We now construct an interpretation as follows. Choose a set W such that $|W| = 1/c$ and put $\mu(w) = c$ for all $w \in W$. Note that every world has positive probability, as required by the positive interpretations that we work with.

In the remainder of the construction, we start with defining the ABox slice of the desired interpretation and then repeatedly add element slices to satisfy existential restrictions. Before we start, we highlight two properties that are central for this construction:

(a) for each $t_0 \in T$ and each $w_0 \in W$, there is a mapping $\tau : W \to T$ such that $\tau(w_0) = t_0$ and for each $t \in T$, we have $\sum_{w \in W | \tau(w)=t} \mu(w) = \delta_{t_0}(x_t)$.

(b) for each $t_0 \in T'$ and each $w_0 \in W$, there is a mapping $\tau : W \to T'$ such that $\tau(w_0) = t_0$ and for each $t \in T'$, we have $\sum_{w \in W | \tau(w)=t} \mu(w) = \delta_{t_0}(x_t)$.

A proof is straightforward once one observes that for all $(t_0, t) \in (T \times T) \cup (T' \times T')$, there is an integer $n \geq 0$ with $c \cdot n = \delta_{t_0}(x_t)$ and that for each $t_0 \in T$, we have $\sum_{t \in T} \delta_{t_0}(x_t) = 1$ and $\delta_{t_0}(x_{t_0}) > 0$, and correspondingly for each $t_0 \in T'$.

The set of domain elements and the interpretation of concept and role names is constructed inductively along with mappings $\pi_d : W \to \mathfrak{T}_\mathcal{K}$, indexed over domain elements d, such that the following properties are satisfied for all $d \in \Delta^\mathcal{I}$ and all $w, w' \in W$:

(P1) $\pi_d(w) \in T'$ or there are $t \in T$ and $a \in \mathsf{Ind}(\mathcal{A})$ such that $\pi_d(w) = \{C(a_\varepsilon) \mid C(a) \in t\}$;

(P2) $\pi_d(w)$ and $\pi_d(w')$ agree on concepts of the form $P_{\geq n}C$.

The intuitive meaning of $\pi_d(w) = t$ is that we have chosen that d should realize the type t in world w. The steps of the construction are as follows:

- There is a $t_\mathcal{A} \in T$ such that $\mathcal{A} \in t$. To start the construction, we put $\Delta^\mathcal{I} = \mathsf{Ind}(\mathcal{A})$ and pick some $w_0 \in W$. By (a), we find a mapping $\tau : W \to T$ such that $\tau(w_0) := t_\mathcal{A}$ and for each $t \in T$, we have $\sum_{w \in W | \tau(w)=t} \mu(w) = \delta_{t_\mathcal{A}}(x_t)$. Further put

  - $a^\mathcal{I} = a$ for each $a \in \mathsf{N_I}$;
  - $A^{\mathcal{I},w} = \{a \mid A(a) \in \tau(w)\}$ for each concept name A and $w \in W$;
  - $r^{\mathcal{I},w} = \{(a, b) \mid r(a, b) \in \tau(w)\}$ for each role name r and $w \in W$.

  For all $a \in \mathsf{Ind}(\mathcal{A})$ and $w \in W$, put $\pi_a(w) = \{C(a_\varepsilon) \mid C(a) \in \tau(w)\}$. (P1) is satisfied by definition of π_a. (P2) is satisfied since for all $w', w'' \in W$, $\delta_{t_\mathcal{A}}(x_{\tau(w')})$ and $\delta_{t_\mathcal{A}}(x_{\tau(w'')})$ must be positive by the Properties of τ, and thus the inequalities in Point 3 of the definition of $\mathcal{E}_{t_\mathcal{A},T}$ yield agreement of $\pi_d(w')$ and $\pi_d(w'')$ on concepts of the form $P_{\geq n}C$.

- Repeat the following step indefinitely and fairly. Choose $w \in W$, $d \in \Delta^\mathcal{I}$, and $\exists r.C \in \pi_d(w)$ such that there is no $e \in \Delta^\mathcal{I}$ with $(d, e) \in r^{\mathcal{I},w}$ and $C \in \pi_e(w)$. By saturatedness and (P1), there is a $t \in T'$ such that $t \supseteq \{C\} \cup \{D \mid \forall r.D \in \pi_d(w)\}$. Add a fresh element e to $\Delta^\mathcal{I}$. By (b) there is a mapping $\pi_e : W \to T'$ such that $\pi_e(w) = t$ and for each $t' \in T'$, we have $\sum_{w' \in W | \pi_e(w')=t'} \mu(w') = \delta_t(x_{t'})$. Put $r^{\mathcal{I},w} = r^{\mathcal{I},w} \cup \{(d, e)\}$. For all concept names A and $w' \in W$, put $A^{\mathcal{I},w'} = A^{\mathcal{I},w'} \cup \{e\}$ if $A \in \pi_e(w')$. Using similar arguments as in the base case, it can be verified that (P1) and (P2) are satisfied.

One then has a *truth lemma* stating that

$$C^{\mathcal{I},w} = \{d \in \Delta^{\mathcal{I}} \mid C \in \pi_d(w)\}$$

for all $w \in W$ and all $C \in \mathsf{ccl}(\mathcal{K})$, proved by induction on the structure of C. The base case (concept names) holds by construction, and the Boolean cases are trivial. The case for existential restrictions is straightforward by the iterative construction above, noting that it is applied fairly. Finally, consider the case $C = P_{\geq n}D$. Let $w \in W$, and let $d \in \Delta^{\mathcal{I}}$. Then

$$d \in (P_{\geq n}D)^{\mathcal{I},w} \text{ iff } \mu\{w' \in W \mid d \in D^{\mathcal{I},w'}\} \geq n.$$

Assume first that $d \notin \mathsf{Ind}(\mathcal{A})$, i.e. d has been added in the iterative construction above, in a step for a world $\widehat{w}$ selecting an element type $\widehat{t}$ for d in $\widehat{w}$. Then we have

$$\begin{aligned}
\mu\{w' \in W \mid d \in D^{\mathcal{I},w'}\} &= \mu\{w' \in W \mid D \in \pi_d(w')\} && \text{(by induction)}\\
&= \sum_{t \in T' \mid D \in t} \mu\{w' \mid \pi_d(w') = t\}\\
&= \sum_{t \in T' \mid D \in t} \delta_{\widehat{t}}(x_t) && \text{(by construction).}
\end{aligned}$$

Thus

$$d \in (P_{\geq n}D)^{\mathcal{I},w} \text{ iff } \sum_{t \in T' \mid D \in t} \delta_{\widehat{t}}(x_t) \geq n.$$

Since $\delta_{\widehat{t}}(x_t)$ satisfies the inequalities in Point 1 of the definition of $\mathcal{E}(\widehat{t}, T')$, it follows that $d \in (P_{\geq n}D)^{\mathcal{I},w}$ iff $P_{\geq n}D \in \widehat{t}$. Note that $\pi_d(\widehat{w}) = \widehat{t}$ and thus it remains to show that $P_{\geq n}D \in \pi_d(\widehat{w})$ iff $P_{\geq n}D \in \pi_d(w)$, which is a consequence of (P2).

The remaining case is that $d \in \mathsf{Ind}(\mathcal{A})$, i.e., d comes from the base case of the construction of $\Delta^{\mathcal{I}}$, for some map $\tau : W \to T$. We then have

$$\begin{aligned}
\mu\{w' \in W \mid d \in D^{\mathcal{I},w'}\} &= \mu\{w' \in W \mid D \in \pi_d(w')\} && \text{(by induction)}\\
&= \mu\{w' \in W \mid D(d) \in \tau(w')\} && \text{(by construction)}\\
&= \sum_{t \in T \mid D(d) \in t} \mu\{w' \mid \tau(w') = t\}\\
&= \sum_{t \in T \mid D(d) \in t} \delta_{t_{\mathcal{A}}}(x_t) && \text{(by construction)}
\end{aligned}$$

and can proceed to argue in the same way as in the previous case. This finishes the proof of the truth lemma.

By the truth lemma, $\mathcal{I}, w \models \mathcal{T}$ for all w. Similarly, we prove that $\mathcal{I}, w \models \mathcal{A}'$ iff $\mathcal{A}' \in \tau(w)$ for every world w, with τ as in the base of the construction of $\Delta^{\mathcal{I}}$. The proof is by induction on the structure of $\mathcal{A}' \in \mathsf{acl}(\mathcal{K})$: base cases are by the truth lemma (for $D(a)$) and by construction (for $r(a, b)$), respectively, and Boolean cases are trivial. The case for $P_{\geq n}\mathcal{A}'$ is very similar to the case $d \in \mathsf{Ind}(\mathcal{A})$ in the calculation for $P_{\geq n}D$ above. It now follows that $\mathcal{I}, w_0 \models \mathcal{A}$, as required. $\square$

Thus, we can decide consistency of a given KB $\mathcal{K}$ by checking the existence of a proper quasimodel for $\mathcal{K}$. This can be done as follows. Start with the quasimodel $Q = (\mathfrak{A}_\mathcal{K}, \mathfrak{T}_\mathcal{K})$ and then repeatedly delete types from both components that are not saturated or not coherent. If a type t with $\mathcal{A} \in t$ survives in the first component after the sets have stabilized, answer 'satisfiable'. Otherwise, answer 'unsatisfiable'. Note that this algorithm attempts to compute a maximal quasimodel in which all types are saturated and coherent as a greatest fixpoint by approximation from above. For a given pair (T, T') and an ABox or element type t, it can be decided in polynomial time whether t is saturated and coherent; for the latter, we use the fact that solvability of mixed systems of linear inequalities[4] can be decided in polynomial time (Schröder, 2007). Consequently and since there are only single exponentially many ABox types and element types, the algorithm runs in ExpTime. We have thus shown the following.

Theorem 3. *In Prob-$\mathcal{ALC}$, concept satisfiability, concept subsumption, KB consistency, and KB consequence are* ExpTime*-complete.*

We note that the proof of Theorem 3 easily extends to Prob-$\mathcal{ALC}^{\mathsf{lineq}}$. The only necessary modifications concern the systems of inequalities $\mathcal{E}(t_0, T)$ and consist in replacing ABoxes of the form $P_{\geq n}\mathcal{A}'$ with linear ABox inequalities. Thus, Points 1 and 3 from the definition of $\mathcal{E}(t_0, T)$ now read as follows:

1′. linear ABox inequalities in t_0 are satisfied: $\mathcal{E}(t_0, T)$ contains every linear ABox inequality from t_0 with every expression $P(\mathcal{A}')$ replaced by $\sum_{t \in T | \mathcal{A}' \in t} x_t$;

3′. all realized types agree on linear ABox inequalities with t_0:
$x_t = 0$ for all t such that $(t \setminus t_0) \cup (t_0 \setminus t)$ contains a linear ABox inequality.

We shall see shortly that Theorem 3 can even be extended to Prob-$\mathcal{ALC}^{\mathsf{polyeq}}$. Before we consider this case, though, we make explicit some interesting consequences of the proof of Theorem 3. In fact, the proof of Lemma 2 establishes a *uniform model property (UMP)*: every consistent Prob-$\mathcal{ALC}$ KB has a *uniform* model $\mathcal{I}$, i.e., the probability distribution μ of $\mathcal{I}$ satisfies $\mu(w) = \mu(w')$ for all $w, w' \in W$. Our constructions do not directly yield a *finite model property (FMP)* as the domains of the constructed models may be infinite (whereas the set of worlds is finite). As discussed in the following, this can be fixed by a slight modification which then even yields a *bounded model property (BMP)*. We use the following result about linear programming.

Proposition 4 (Chvátal 1983, Fagin, Halpern, & Megiddo 1990)**.** *If a system of r linear inequalities has a non-negative solution, then it has a non-negative solution with at most r entries positive. If moreover all coefficients are intergers of length ℓ, then we can additionally assume that the size of each member of the solution is $\mathcal{O}(r\ell + r\log(r))$.*

Here, the *length* of an integer denotes the number of bits of its binary representation and the *size* of a rational number is the sum of the lengths of the binary representations of the numerator and denominator in its irreducible form. By this lemma, the number $c = |W|$ of worlds in the models we construct in the above proof is bounded by $2^{2^{\mathcal{O}(|\mathcal{K}|)}}$: c is the

4. I.e., systems that comprise strict inequalities based on "$<$" and non-strict inequalities based on "$\leq$".

product of exponentially many denominators $\mathsf{den}(\delta_t(x_{t'}))$, whose length is, by the previous lemma, polynomial in $r\ell$ where r is the number of inequalities and ℓ is the maximal length of coefficients in $\mathcal{E}(t,T)$ (resp. $\mathcal{E}(t,T')$). Clearly, ℓ is polynomial in $|\mathcal{K}|$. Moreover, r can be made polynomial in $|\mathcal{K}|$ by replacing the exponentially many equations from Point 3 of the definition of $\mathcal{E}(t,T)$ with a single equation $x_{t_0} + \cdots + x_{t_k} = 0$ where $t_0,\ldots,t_k$ are all types that qualify as t in the original Point 3. In summary, c is of length $2^{\mathcal{O}(|\mathcal{K}|)}$, so that $c \in 2^{2^{\mathcal{O}(|\mathcal{K}|)}}$. It is then simple to obtain the same bound on the number of domain elements: in the model construction, we can stop adding new domain elements as soon as in every world, there is an element of every type; since every new domain element adds at least one new type to at least one world, we need at most $2^{\mathcal{O}(|\mathcal{K}|)} \cdot 2^{2^{\mathcal{O}(|\mathcal{K}|)}} = 2^{2^{\mathcal{O}(|\mathcal{K}|)}}$ domain elements. We thus obtain the following combination of the UMP and a BMP.

Corollary 5. *Every consistent Prob-$\mathcal{ALC}^{\mathsf{lineq}}$ KB has a uniform model in which the number of worlds and the number of domain elements are both bounded by $2^{2^{\mathcal{O}(|\mathcal{K}|)}}$.*

We have no proof that a super-*polynomial* number of worlds (nor a super-exponential number of domain elements) can actually be enforced.

3.2 Polynomial Inequalities

The ExpTime upper bound just established can be generalized to Prob-$\mathcal{ALC}^{\mathsf{polyeq}}$. It has already been shown in the previous section how the systems of inequalities $\mathcal{E}(t_0,T)$ can be modified to account for linear inequalities, see Points $1'$ and $3'$. Exactly the same modification works also for Prob-$\mathcal{ALC}^{\mathsf{polyeq}}$, with the effect that these systems now consist of polynomial inequalities instead of of linear ones. To emphasize this difference, we denote the modified systems with $\mathcal{E}^p(t_0,T)$. We are interested in finding non-negative solutions over the reals, which in the case of polynomial inequalities does not coincide with the existence of a non-negative solution over the rationals.

There are now two difficulties to be addressed. The first one is that polynomial systems of inequalities cannot be solved in polynomial time (unless P=NP). In fact, the best known upper bound is PSpace (Canny, 1988) and the best known lower bound is NP. The exact complexity is a notorious open problem which has even resulted in the introduction of the complexity class $\exists\mathbb{R}$ that consists of all problems mutually reducible in polynomial time with solving polynomial systems of inequalities (Schaefer, 2010). It would thus seem that the methods employed in the previous section yield only an ExpSpace upper bound for Prob-$\mathcal{ALC}^{\mathsf{polyeq}}$. To push the bound back down to ExpTime, we first make the following observation (Fagin, Halpern, & Megiddo, 1990).

Lemma 6. *If a system $\mathcal{E}^p(t_0,T)$ has a solution, T a set of ABox types or of element types, then it has a solution with at most $|\mathcal{K}|^2$ variables positive.*

Proof. Let δ be a solution for $\mathcal{E}^p(t_0,T)$. For each $\mathcal{A} \in \mathsf{acl}(\mathcal{K})$, set $\delta(\mathcal{A}) = \sum_{t\in T \mid \mathcal{A}\in t} \delta(x_t)$. The system $\widehat{\mathcal{E}}(t_0,T)$ of linear inequalities is obtained from $\mathcal{E}^p(t_0,T)$ by replacing the inequalities in Point $1'$ with

$$\sum_{t\in T \mid \mathcal{A}\in t} x_t = \delta(\mathcal{A}) \qquad \text{for all } \mathcal{A} \in \mathsf{acl}(\mathcal{K})$$

and the inequalities in Point 3 with a single one as in the proof of Corollary 5. By construction, the system $\widehat{\mathcal{E}}(t_0, T)$ has a solution. Since $\widehat{\mathcal{E}}(t_0, T)$ consists of at most $4|\mathcal{K}|^2 + 3$ linear inequalities, the first part of Proposition 4 (which also applies when coefficients are real numbers as in $\widehat{\mathcal{E}}(t_0, T)$) yields the existence of a solution with at most $\mathcal{O}(|\mathcal{K}|^2)$ positive values. By construction of $\widehat{\mathcal{E}}(t_0, T)$, this solution is also a solution of $\mathcal{E}^p(t_0, T)$. $\square$

Instead of checking a system $\mathcal{E}^p(t_0, T)$ directly for solvability, Lemma 6 allows us to consider all systems $\mathcal{E}^p(t_0, T^*)$ with $T^* \subseteq T$ such that $|T^*| \leq 4|\mathcal{K}|^2 + 3$ and $t_0 \in T^*$: If any of these systems has a solution δ, then we obtain a solution for $\mathcal{E}^p(t_0, T)$ by setting $\delta(x_t) = 0$ for all $t \in T \setminus T^*$. Conversely, solvability of $\mathcal{E}^p(t_0, T)$ and Lemma 6 implies the existence of a solution δ for $\mathcal{E}^p(t_0, T)$ and a subset $T^* \subseteq T$ such that $|T^*| \leq 4|\mathcal{K}|^2 + 3$ and $\delta(x_t) = 0$ for all $t \in T \setminus T^*$. Clearly, δ is also a solution for $\mathcal{E}^p(t_0, T^*)$. Each system $\mathcal{E}^p(t_0, T^*)$ is of size polynomial in $|\mathcal{K}|$, so there are exponentially many such systems. Thus, the solvability of $\mathcal{E}^p(t_0, T)$ can be checked in ExpTime.

The second problem that arises is the loss of the UMP, which means that the model construction underlying Lemma 2, which assumes a rational solution and produces uniform models, no longer works.

Lemma 7. *Prob-$\mathcal{ALC}^{\mathsf{indep}}$ does not have the UMP.*

Proof. Consider the knowledge base $\mathcal{K} = (\mathcal{T}, \{D(a)\})$ with

$$\mathcal{T} = \{\mathsf{indep}(A_1, A_2), \mathsf{indep}(A_2, A_3), \mathsf{indep}(A_1, A_3)\}$$
$$D = P_{=\frac{1}{2}}(A_1 \sqcap A_2) \sqcap P_{=\frac{1}{2}}(A_2 \sqcap A_3) \sqcap P_{=\frac{1}{2}}(A_1 \sqcap A_3).$$

Let $\mathcal{I}$ be a model of $\mathcal{K}$. We show that $\mathcal{I}$ cannot be uniform. For brevity, put $p_i = p^{\mathcal{I}}_{a^{\mathcal{I}}}(A_i)$ for $i = 1, 2, 3$. Since $\mathcal{I}$ is a model of $\mathcal{T}$ and of D, we must clearly have $p_i p_j = 1/2$ for $i \neq j$. It follows that all p_i are non-zero, and hence that $p_1 = p_2 = p_3$ (e.g. $p_2 = p_3$ follows from $p_1 p_2 = p_1 p_3$ and $p_1 \neq 0$), so $p_1 = p_2 = p_3 = 1/\sqrt{2}$. But in uniform models, all probabilities are rational, so $\mathcal{I}$ is not uniform.

It remains to show that $\mathcal{K}$ has a model at all. To construct such a model $\mathcal{I}$, take $\Delta^{\mathcal{I}} = \{d\}$ and the set of worlds $W = W_0^3$ where W_0 has two worlds $w_\bot$, $w_\top$. Equip W_0 with the probability measure μ_0 given by $\mu_0(w_\top) = 1/\sqrt{2}$, and take μ to be the arising product measure on W, that is, $\mu(v_1, v_2, v_3) = \prod_{i=1,2,3} \mu_0(v_i)$ for all $v_1, v_2, v_3 \in W$. Then let $a^{\mathcal{I}} = d$ satisfy A_i in a world (v_1, v_2, v_3) iff $v_i = w_\top$. It can be verified that $\mathcal{I}$ is a model of $\mathcal{K}$. $\square$

We now describe a more subtle model construction that replaces the one used in the proof of Lemma 2 and works also for real-valued solutions. Of course, it constructs non-uniform models. Proper quasimodels for Prob-$\mathcal{ALC}^{\mathsf{polyeq}}$ are defined exactly as for Prob-$\mathcal{ALC}$, except that the generalized systems $\mathcal{E}^p(t, T)$ and $\mathcal{E}^p(t, T')$ are used.

Lemma 8. *$\mathcal{K}$ is consistent iff there is a proper quasimodel for $\mathcal{K}$.*

Proof. "⇒" is straightforward, as in the proof of Lemma 2. We prove the "⇐" direction. In particular, we show how to define W and μ such that the Properties (a) and (b) used in the proof of Lemma 2 still hold.

Choose a non-negative solution δ_t for each system $\mathcal{E}(t,T)$, $t \in T$, and each system $\mathcal{E}(t,T')$, $t \in T'$. Let S be the set of all values that appear in the chosen solutions. We use $\mathsf{f}(x)$ to denote the fractional part of a real number x. Now let M be the subset of $[0,1]$ that contains 0, 1, and all numbers of the form

$$\mathsf{f}(i \cdot s_0 + (s_1 + \cdots + s_\ell))$$

where $s_0, \ldots, s_\ell \in S$ and $i \in \mathbb{N}$ are such that $i \cdot s_0 < 1$ and $s_1 + \cdots + s_\ell \leq 1$. Clearly, M is finite. Let $m_1, \ldots, m_k$ be the enumeration of M in ascending order (thus $m_1 = 0$ and $m_k = 1$). Define

$$W = \{(m_i, m_{i+1}) \mid 1 \leq i < k\}.$$

These worlds partition the interval $[0,1]$ into subintervals, which give rise to

$$\mu(m_i, m_{i+1}) = m_{i+1} - m_i \text{ for all } (m_i, m_{i+1}) \in W.$$

We have to show that (a) and (b) hold for W and μ as just defined. Because the proofs are almost identical, we concentrate on (a). Let $t_0 \in T$ and $w_0 = (m_{i_0}, m_{i_0+1}) \in W$. We have to define a mapping $\tau : W \to T$ such that $\tau(w_0) = t_0$ and

$$\sum_{w \in W \mid \tau(w) = t} \mu(w) = \delta_{t_0}(x_t) \text{ for all } t \in T. \tag{$*$}$$

Let $t_0, \ldots, t_p$ be an enumeration of all types that are assigned positive probability by δ_{t_0} (where t_0 is the type to be assigned to w_0). Let x_0 be the largest multiple of $\delta_{t_0}(x_{t_0})$ with $x_0 \leq m_{i_0}$. For $1 \leq i \leq p+1$, put

$$x_i = x_0 + \delta_{t_0}(x_{t_0}) + \cdots + \delta_{t_0}(x_{t_{i-1}}).$$

Note that $\mathsf{f}(x_i) \in M$ for all $0 \leq i \leq p+1$. Intuitively, $x_0, \ldots, x_{p+1}$ provides an assignment of a subinterval of $[x_0, x_0+1]$ to each type from the list $t_0, \ldots, t_p$ such that the subinterval assigned to t_i is of length $\delta_{t_0}(t)$ and $[x_0, x_1]$ is the subinterval assigned to t_0. We use the $[x_0, 1]$-part of this assignment to define the restriction of τ to worlds from the interval $[x_0, 1]$ and the $[1, x_0+1]$-part to define the restriction of τ to worlds from the interval $[0, x_0]$. Thus put $\tau(m_i, m_{i+1}) = t_\ell$ whenever one of the following holds:

1. $x_0 \leq m_i$, $x_\ell \leq m_i$, and $m_{i+1} \leq x_{\ell+1}$;
2. $m_i < x_0$, $x_\ell \leq m_i + 1$, and $m_{i+1} + 1 \leq x_{\ell+1}$.

It can be verified that $(*)$ is satisfied using that $x_{i+1} - x_i = \delta_{t_0}(t_i)$ and that $\mathsf{f}(x_i) \in M$, that is, the subintervals defined by $x_0, \ldots, x_{p+1}$ are unions of consecutive subintervals defined by W and μ, up to the shift by -1 in Point 2. It thus remains to show that $\tau(w_0) = t_0$. By choice, $x_0 \leq m_{i_0}$. Moreover, m_{i_0+1} is the least element of M strictly above m_{i_0}, thus $m_{i_0+1} - m_{i_0} \leq x_1 - x_0$ and Point 1 yields $\tau(m_{i_0}, m_{i_0+1}) = t_0$. □

We have thus established the following result.

Theorem 9. *In Prob-$\mathcal{ALC}^{\mathsf{polyeq}}$, concept satisfiability, concept subsumption, KB consistency, and ABox consequence are* ExpTime*-complete.*

3.3 Probabilistic Roles

Reasoning in Prob-$\mathcal{ALC}_r$ is much more challenging than in Prob-$\mathcal{ALC}$. In fact, we do not know whether it is decidable, but we can prove undecidability if, in addition to probabilistic roles, we allow binary independence constraints or linear concept inequalities. We also observe that the fragment Prob-$\mathcal{ALC}_r^{01}$ of Prob-$\mathcal{ALC}_r$ that allows only the probability restrictions $P_{>0}C$ and $P_{=1}C$ is decidable, and that reasoning with it is 2ExpTime-complete. We start with the latter, exploiting a very closely relationship between Prob-$\mathcal{ALC}_r^{01}$ and the modal description logic $\mathsf{S5}_{\mathcal{ALC}}$ in the version with modal roles.

We recall that $\mathsf{S5}_{\mathcal{ALC}}$ extends the syntax of $\mathcal{ALC}$ with concepts of the form $\Diamond C$ and with *modal roles* $\Box r$ and $\Diamond r$ which can be used inside existential and universal restrictions and where r is a role name (Gabbay et al., 2003; Artale, Lutz, & Toman, 2007). The semantics of $\mathsf{S5}_{\mathcal{ALC}}$ is quite similar to that of Prob-$\mathcal{ALC}_r$. An $\mathsf{S5}_{\mathcal{ALC}}$-interpretation is defined like a probabilistic interpretation except that it does not comprise a probability distribution over worlds. The interpretation of $\Diamond C$ in a world w in such an interpretation $\mathcal{I}$, with set W of worlds, is then given by

$$(\Diamond C)^{\mathcal{I},w} = \{d \in \Delta^{\mathcal{I}} \mid \exists w' \in W : d \in C^{\mathcal{I},w}\},$$

and that of modal roles $\Diamond r$, $\Box r$ by

$$(\exists \Diamond r.C)^{\mathcal{I},w} = \{d \in \Delta^{\mathcal{I}} \mid \exists e \in C^{\mathcal{I}}\, \exists w' \in W : (d,e) \in r^{\mathcal{I},w'}\}$$
$$(\exists \Box r.C)^{\mathcal{I},w} = \{d \in \Delta^{\mathcal{I}} \mid \exists e \in C^{\mathcal{I}}\, \forall w' \in W : (d,e) \in r^{\mathcal{I},w'}\}.$$

TBoxes are then defined as for $\mathcal{ALC}$ and Prob-$\mathcal{ALC}$. We are not going to consider ABoxes in the context of $\mathsf{S5}_{\mathcal{ALC}}$.

As has already been stated above, Prob-$\mathcal{ALC}_r^{01}$ is the fragment of Prob-$\mathcal{ALC}_r$ that only admits probabilities zero and one. More specifically, we only admit probability restrictions $P_{>0}C$ and existential restrictions $\exists P_{>0}r.C$, and $\exists P_{=1}r.C$; note that $P_{=1}C$ is then definable as $\neg P_{>0}\neg C$, $P_{=0}C$ as $P_{=1}\neg C$, and $P_{<1}C$ as $\neg P_{=1}C$. In the literature, there are known cases where this restriction of probability values regains decidability of undecidable probabilistic logics, for example, in probabilistic CTL (Brázdil, Forejt, Kretínský, & Kucera, 2008).

There is a straightforward equivalence-preserving translation of $\mathsf{S5}_{\mathcal{ALC}}$-TBoxes into first-order logic (Gabbay et al., 2003); therefore, all reasoning problems in $\mathsf{S5}_{\mathcal{ALC}}$ remain unaffected if we restrict our attention to at most countable models. Now it is easy to see that over such models, $\mathsf{S5}_{\mathcal{ALC}}$ and Prob-$\mathcal{ALC}_r^{01}$ are essentially identical: we can equip every at most countable set W of worlds with a discrete probability distribution μ such that $\mu(w) > 0$ for all $w \in W$; the exact numerical values $\mu(w)$ are irrelevant for the semantics of Prob-$\mathcal{ALC}_r^{01}$. Under this conversion of models, the concept $P_{>0}C$ of Prob-$\mathcal{ALC}_r^{01}$ is equivalent to the concept $\Diamond C$ of $\mathsf{S5}_{\mathcal{ALC}}$, $\exists P_{>0}r.C$ is equivalent to $\exists\Diamond r.C$, and $\exists P_{=1}r.C$ is equivalent to $\exists\Box r.C$. It is already known that, in $\mathsf{S5}_{\mathcal{ALC}}$, concept satisfiability and subsumption are 2ExpTime-complete (Artale, Lutz, & Toman, 2007). By what was said above, these results transfer to Prob-$\mathcal{ALC}_r^{01}$.

Theorem 10. *(Artale, Lutz, & Toman, 2007) In Prob-$\mathcal{ALC}_r^{01}$, concept satisfiability and concept subsumption are 2*ExpTime*-complete.*

It seems clear that this result extends to KB consistency and consequence, but to the best of our knowledge these reasoning problems have not been studied for $\mathsf{S5}_{\mathcal{ALC}}$. We skip them here, too.

We now turn our attention to the extensions of Prob-$\mathcal{ALC}_r$ with binary independence constraints and with linear concept inequalities, proving undecidability by reduction of the halting problem of two-register machines. We briefly recall the relevant definitions. A two-register machine M is similar to a Turing machine. It also has an internal state taken from a finite set of possible states, but instead of a tape, it has two registers that contain non-negative integers. In one step, the machine can increment the content of one of the registers or test whether the content of the register is zero and if not then decrement it. In the second case, the successor state depends on whether the tested register was zero or not. There is a designated halting state, and M halts if it encounters this state. Formally:

Definition 11. A (deterministic) *two-register machine (2RM)* is a pair $M = (Q, P)$ with $Q = \{q_0, \dots, q_\ell\}$ a set of *states* and $P = I_0, \dots, I_{\ell-1}$ a sequence of *instructions*. By definition, q_0 is the *initial state*, and q_ℓ the *halting state*. For all $i < \ell$,

- either $I_i = +(p, q_j)$ is an *incrementation instruction* with $p \in \{1, 2\}$ a register and q_j the subsequent state;
- or $I_i = -(p, q_j, q_k)$ is a *decrementation instruction* with $p \in \{1, 2\}$ a register, q_j the subsequent state if register p contains 0, and q_k the subsequent state otherwise.

A *configuration* of M is a triple (q, m, n), with q the current state and $m, n \in \mathbb{N}$ the register contents. We write $(q_i, n_1, n_2) \Rightarrow_M (q_j, m_1, m_2)$ if one of the following holds:

- $\mathcal{I}_i = +(p, q_j)$, $m_p = n_p + 1$, and $m_{\overline{p}} = n_{\overline{p}}$, where $\overline{1} = 2$ and $\overline{2} = 1$;
- $\mathcal{I}_i = -(p, q_j, q_k)$, $n_p = m_p = 0$, and $m_{\overline{p}} = n_{\overline{p}}$;
- $\mathcal{I}_i = -(p, q_k, q_j)$, $n_p > 0$, $m_p = n_p - 1$, and $m_{\overline{p}} = n_{\overline{p}}$.

The *computation* of M on input $(n, m) \in \mathbb{N}^2$ is the unique longest (possibly infinite) configuration sequence $(p_0, n_0, m_0) \Rightarrow_M (p_1, n_1, m_1) \Rightarrow_M \cdots$ such that $p_0 = q_0$, $n_0 = n$, and $m_0 = m$.

The *halting problem* for 2RMs is to decide, given a 2RM M, whether its computation on input $(0, 0)$ is finite (which implies that its last state is q_ℓ).

Since 2RMs are Turing complete, their halting problem is undecidable.

Theorem 12. *In Prob-$\mathcal{ALC}_r^{\mathsf{indep}}$ and Prob-$\mathcal{ALC}_r^{\mathsf{lineq}}$, concept satisfiability, concept subsumption, KB consistency, and ABox consequence are undecidable.*

Proof. We first consider Prob-$\mathcal{ALC}_r^{\mathsf{lineq}}$. We reduce the halting problem for 2RMs to concept satisfiability by transforming a 2RM $M = (Q, P)$ into a TBox $\mathcal{T}_M$ and a concept I such that I is not satisfiable w.r.t. $\mathcal{T}_M$ iff M halts. More precisely, every model of $\mathcal{T}_M$ and I describes an infinite computation of M on $(0, 0)$ via an infinite r-chain of probability 1, i.e., there are $d_0, d_1, \ldots \in \Delta^{\mathcal{I}}$ such that $p^{\mathcal{I}}_{d_i, d_{i+1}}(r) = 1$ for all $i \geq 0$; register contents are described using two concept names R_1 and R_2: at each element d_j, register i has value n if $p^{\mathcal{I}}_{d_j}(R_i) = \frac{1}{2^n}$. Conversely, every infinite computation gives rise to a model of $\mathcal{T}_M$ and I. We use the following signature to encode computations on the chain $d_0, d_1, \ldots$:

- concept names $Q_0, \ldots, Q_\ell$ encode the current state;
- the concept names R_1 and R_2 for representing register contents;
- concept names S_1 and S_2 are used at d_i to describe the register contents at d_{i+1}, encoded in the same way as in the case of R_1 and R_2.

We define the TBox $\mathcal{T}_M$ step by step, along with explanations:

- Start in state q_0 and with the registers containing zero:

$$I \sqsubseteq Q_0 \sqcap P_{=1}(R_1) \sqcap P_{=1}(R_2)$$

- Enforce the infinite chain:

$$\top \sqsubseteq \exists P_{=1} r.\top$$

- Incrementation instructions are executed correctly: for all $I_i = +(p, q_j)$,

$$\begin{aligned} Q_i &\sqsubseteq \forall r.Q_j \\ Q_i &\sqsubseteq P(R_p) = 2 \cdot P(S_p) \\ Q_i &\sqsubseteq P(R_{\overline{p}}) = P(S_{\overline{p}}) \\ S_\ell &\sqsubseteq \forall r.R_\ell \\ \neg S_\ell &\sqsubseteq \forall r.\neg R_\ell \end{aligned}$$

  where ℓ ranges over $\{1, 2\}$;

- Decrementation instructions are executed correctly: for all $I_i = -(p, q_j, q_k)$,

$$\begin{aligned} Q_i \sqcap P_{=1} R_p &\sqsubseteq \forall r.Q_j \\ Q_i \sqcap P_{=1} R_p &\sqsubseteq P(R_1) = P(S_1) \\ Q_i \sqcap P_{=1} R_p &\sqsubseteq P(R_2) = P(S_2) \\ Q_i \sqcap \neg P_{=1} R_p &\sqsubseteq \forall r.Q_k \\ Q_i \sqcap \neg P_{=1} R_p &\sqsubseteq P(S_p) = 2 \cdot P(R_p) \\ Q_i \sqcap \neg P_{=1} R_p &\sqsubseteq P(R_{\overline{p}}) = P(S_{\overline{p}}) \end{aligned}$$

- The halting state q_ℓ is never reached, and thus the computation is infinite:

$$\top \sqsubseteq \neg Q_\ell$$

It is not difficult to prove that the computation of M on $(0, 0)$ is finite iff I is satisfiable w.r.t. $\mathcal{T}_M$.

We can adapt the reduction to Prob-$\mathcal{ALC}_r^{\mathsf{indep}}$ as follows. To describe incrementation, we introduce an auxiliary concept name H_p for each $p \in \{1, 2\}$ and then replace the second concept inclusion in the incrementation block of the Prob-$\mathcal{ALC}_r^{\mathsf{lineq}}$ reduction with

$$\begin{aligned} &\mathsf{indep}(H_p, R_p) \\ &Q_i \sqsubseteq P_{=\frac{1}{2}} H_p \sqcap P_{=1}((R_p \sqcap H_p) \leftrightarrow S_p). \end{aligned}$$

Note that, together with $P_{=\frac{1}{2}} H_p$, the independence constraint guarantees that the probability of $H_p \sqcap R_p$ is exactly half the probability of R_p. To describe decrementation, we introduce an auxiliary concept name H'_p for each $p \in \{1, 2\}$ and then replace the second concept inclusion in the decrementation block with

$$\begin{array}{l} \mathsf{indep}(H'_p, S_P) \\ Q_i \sqcap \neg P_{=1} R_p \sqsubseteq P_{=\frac{1}{2}} H'_p \sqcap P_{=1}((S_p \sqcap H'_p) \leftrightarrow R_p). \end{array}$$

□

4. Prob-$\mathcal{EL}$ and Variants

We study probabilistic DLs that are based on $\mathcal{EL}$ instead of $\mathcal{ALC}$, aiming in particular to identify cases of lower complexity. Our first result is a strong negative one: for almost all probability restrictions such as $P_{>0.4}C$ and $P_{<0,9}C$, Prob-$\mathcal{EL}$ is ExpTime hard even if only this one probability restriction, with its fixed comparison type and probability value, is available. The intuitive reason is that these probabilistic DLs are not convex, that is, they allow to define a form of disjunction. We thus proceed to study Prob-$\mathcal{EL}^{01}$, the version of Prob-$\mathcal{EL}$ in which only probability restrictions of the form $P_{>0}C$ and $P_{=1}C$ are admitted. This logic is convex and we show that it admits reasoning in polynomial time. We also consider restricting the TBox formalism to so-called classical TBoxes and show that, in this case, also probability values except zero and one can be admitted without losing tractability. Finally, we extend Prob-$\mathcal{EL}^{01}$ with probabilistic roles and establish PSpace-completeness of reasoning.

4.1 Lower Bounds

The aim of this section is to establish the following strong lower bound, showing that already many simple fragments of Prob-$\mathcal{EL}$ are computationally as hard as Prob-$\mathcal{ALC}$. For all $\sim \in \{\leq, <, =, >, \geq\}$ and $p \in [0, 1]$, we use Prob-$\mathcal{EL}^{\sim p}$ to denote the fragment of Prob-$\mathcal{EL}$ that admits only probability restrictions of the form $P_{\sim p}C$.

Theorem 13. *For all $p \in (0, 1)$ and $\sim \in \{\leq, <, =, >, \geq\}$, as well as for $\sim p \in \{<1, =0\}$, concept subsumption in Prob-$\mathcal{EL}^{\sim p}$ is ExpTime-hard.*

To prove Theorem 13, we show that each relevant logic Prob-$\mathcal{EL}^{\sim p}$ is *non-convex* in the sense that there are a general TBox $\mathcal{T}$ and concepts $C, D_1, \ldots, D_n$, $n \geq 2$ such that $\mathcal{T} \models C \sqsubseteq D_1 \sqcup \ldots \sqcup D_n$, but $\mathcal{T} \not\models C \sqsubseteq D_i$ for all i. Once non-convexity is established, standard proof techniques (Baader et al., 2005) can be used to reduce concept satisfiability in $\mathcal{ALC}$ relative to general TBoxes to subsumption in Prob-$\mathcal{EL}^{\sim p}$, see Appendix A.3 for details of how the reduction works in this case.

We start with the case $p \in (0, 1)$ and $\sim \in \{<, \leq\}$, or $\sim p \in \{<1, =0\}$, where non-convexity can be attained even if the TBox $\mathcal{T}$ is empty. In fact, we have

$$\emptyset \models \top \sqsubseteq P_{\sim p} A \sqcup P_{\sim p} P_{\sim p} A$$

since $d \notin (P_{\sim p} A)^{\mathcal{I}, w}$ implies $d \in (P_{=0} P_{\sim p} A)^{\mathcal{I}, w}$ and thus by choice of $\sim$ and p also $(P_{\sim p} P_{\sim p} A)^{\mathcal{I}, w}$, but we clearly also have

$$\emptyset \not\models \top \sqsubseteq P_{<p} A \qquad \text{and} \qquad \emptyset \not\models \top \sqsubseteq P_{<p} P_{<p} A.$$

For the cases where $\sim \in \{=,>,\geq\}$ and $p \in (0,1)$, non-convexity is slightly more difficult to establish. We concentrate here on the case $\sim = \geq$ and defer the cases $\sim \in \{=,>\}$, which are minor variations, to the appendix. So consider a logic Prob-$\mathcal{EL}^{\geq p}$ with $p \in (0,1)$. We begin with the subcase $p \leq 0.5$. Pick $k > 0$ such that $k \cdot p > 1$ and put

$$\begin{aligned}
\mathcal{T} &= \{A_i \sqcap A_j \sqsubseteq P_{\geq p} B_{ij} \mid 1 \leq i < j \ \leq k\} \\
C &= P_{\geq p} A_1 \sqcap \ldots \sqcap P_{\geq p} A_k \\
D_{ij} &= P_{\geq p} B_{ij}.
\end{aligned}$$

We show that the above witnesses non-convexity. Intuitively, the probabilities stipulated by C sum up to > 1, thus some of the A_i have to overlap in some world, but it is not clear which A_i these are. The TBox makes an overlap between A_i and A_j in some world globally visible by enforcing satisfaction of the marker concept $P_{\geq p} B_{ij}$, which then holds in *all* worlds.

Lemma 14. *$\mathcal{T} \models C \sqsubseteq \bigsqcup_{1 \leq i<j \leq k} D_{ij}$, but $\mathcal{T} \not\models C \sqsubseteq D_{ij}$ for $1 \leq i < j \leq k$.*

Proof. For the former claim, let $\mathcal{I}$ be a model of $\mathcal{T}$ and $d \in C^{\mathcal{I},w}$. Since $d \in (P_{\geq p} A_i)^{\mathcal{I},w}$ for $1 \leq i \leq k$ and $k \cdot p > 1$, there is a world v with $d \in (A_i \sqcap A_j)^{\mathcal{I},w}$ for some i, j with $1 \leq i < j \leq k$. It follows that $d \in D_{ij}^{\mathcal{I},v}$, thus $d \in D_{ij}^{\mathcal{I},w}$.

For the latter claim, fix i_0, j_0 with $1 \leq i_0 < j_0 \leq k$. We construct a model $\mathcal{I} = (\Delta^{\mathcal{I}}, W, (\mathcal{I}_w)_{w \in W}, \mu)$ by setting $\Delta^{\mathcal{I}} = \{d\}$, $W = \{w_1, w_2\}$, $\mu(w_1) = \mu(w_2) = 0.5$, taking d to satisfy A_{i_0} exactly in world w_2, and A_i exactly in world w_1 for $i \neq i_0$. For $1 \leq i < j < k$, we then ensure that d satisfies $P_{\geq p} B_{ij}$ only when this is required by the TBox; to this end, we take $B_{ij}^{\mathcal{I},w} = \emptyset$ if $i_0 \in \{i,j\}$, and $B_{ij}^{\mathcal{I},w} = \{d\}$ otherwise, for all $w \in W$. Then $\mathcal{I}$ is, by construction, a model of $\mathcal{T}$. To see that $\mathcal{T} \not\models C \sqsubseteq D_{i_0 j_0}$, let $w \in W$; then clearly $p_d^{\mathcal{I}}(A_i) = 0.5 \geq p$, i.e., $d \in C^{\mathcal{I},w}$, and there is no world v with $d \in B_{i_0 j_0}^{\mathcal{I},v}$, so $d \notin D_{i_0 j_0}^{\mathcal{I},w}$. □

For the remaining subcase $1 > p > 0.5$, the main idea is to use the constructor $P_{\geq p} C$ to simulate $P_{\geq q} C$, for some $q < 0.5$. Let n be the smallest integer such that $n > \frac{1}{2(1-p)}$ and put $q = pn - n + 1$. Note that $n \geq 2$ and $0 < q < 0.5$ (to see that $q > 0$ note that since $\frac{1}{2(1-p)} > 1$, $n < \frac{1}{1-p}$). Intuitively, we use the fact that

$$P_{\geq p} X_1 \sqcap \ldots \sqcap P_{\geq p} X_n \sqsubseteq P_{\geq q}(X_1 \sqcap \ldots \sqcap X_n)$$

which allows us to simply redo the above reduction for the case $p < 0.5$ with q in place of p. Thus, let $k > 0$ be such that $k \cdot q > 1$ and define

$$\begin{aligned}
\mathcal{T} &= \{A_{i1} \sqcap \ldots \sqcap A_{in} \sqsubseteq A_i \mid 1 \leq i \leq k\} \cup \\
&\quad \ \{A_i \sqcap A_j \sqsubseteq P_{\geq p} B_{ij} \mid 1 \leq i < j \ \leq k\} \\
C &= \bigsqcap_{1 \leq i \leq k} \bigsqcap_{1 \leq j \leq n} P_{\geq p} A_{ij} \\
D_{ij} &= P_{\geq p} B_{ij}.
\end{aligned}$$

Indeed, the above witnesses non-convexity:

Lemma 15. *$\mathcal{T} \models C \sqsubseteq \bigsqcup_{1 \le i < j \le k} D_{ij}$, but $\mathcal{T} \not\models C \sqsubseteq D_{ij}$ for $1 \le i < j \le k$.*

Proof. For the former claim, let $\mathcal{I}$ be a model of $\mathcal{T}$ and $d \in C^{\mathcal{I},w}$. We first verify the following:

Claim. $d \in (P_{\ge q} A_i)^{\mathcal{I},w}$ for $1 \le i \le k$.

Proof of the Claim. Assume this is not the case for some i, i.e., $d \in (P_{<q} A_i)^{\mathcal{I},w}$. Define

$$S := \sum_{1 \le j \le n} p_d^{\mathcal{I}}(A_{ij}).$$

On the one hand, we must have $S \ge p \cdot n$ since $d \in C^{\mathcal{I},w}$. On the other hand, we can compute S from $\mathcal{I}$ as follows:

$$S = \sum_{w \in W} \mu(w) \cdot n_w \tag{1}$$

where $n_w = |\{j \mid 1 \le j \le n, d \in A_{ij}^{\mathcal{I},w}\}|$. Let $W_1 = \{w \in W \mid n_w = n\}$ and put $W_2 = W \setminus W_1$; in particular, $n_w < n$ for all $w \in W_2$. As, by assumption, $d \in (P_{<q} A_i)^{\mathcal{I},w}$ and $A_{i1} \sqcap \ldots \sqcap A_{in} \sqsubseteq A_i \in \mathcal{T}$, we obtain $p_d^{\mathcal{I}}(A_{i1} \sqcap \ldots \sqcap A_{in}) < q$, and thus $\mu(W_1) < q$. Equation (1) yields

$$\begin{aligned} S = \sum_{w \in W} \mu(w) n_w &= \sum_{w \in W_1} \mu(w) n_w + \sum_{w \in W_2} \mu(w) n_w \\ &\le n\mu(W_1) + (n-1)\mu(W_2) < nq + (n-1)(1-q) = n + q - 1 \end{aligned}$$

It remains to apply $q = pn - n + 1$ to obtain $S < pn$. Thus, we have $pn \le S < pn$, contradiction. This finishes the proof of the claim.

We can now continue as in the proof of Lemma 14: since $k \cdot q > 1$, there is a world $v \in W$ with $d \in (A_i \sqcap A_j)^{\mathcal{I},v}$ for some i, j with $1 \le i < j \le k$. It follows that $d \in D_{ij}^{\mathcal{I},v}$, thus $d \in D_{ij}^{\mathcal{I},w}$ for all $w \in W$.

For the latter claim, fix i_0, j_0 with $1 \le i_0 < j_0 \le k$. To show that $\mathcal{T} \not\models C \sqsubseteq D_{i_0 j_0}$, we construct a model $\mathcal{I} = (\Delta^{\mathcal{I}}, W, (\mathcal{I}_w)_{w \in W}, \mu)$ of $\mathcal{T}$ with $\Delta^{\mathcal{I}} = \{d\}$ and $W = \{w_{1\ell}, w_{2\ell} \mid 1 \le \ell \le n\}$ such that $d \in C^{\mathcal{I},v}$ and $d \notin D_{i_0 j_0}^{\mathcal{I},v}$ for any $v \in W$ by putting for all $1 \le i \le k$ and $1 \le \ell, \ell' \le n$:

$$A_i^{\mathcal{I}_{w_{1\ell}}} := \begin{cases} \{d\} & \text{if } i \ne i_0 \\ \emptyset & \text{otherwise} \end{cases} \qquad A_i^{\mathcal{I}_{w_{2\ell}}} := \begin{cases} \{d\} & \text{if } i = i_0 \\ \emptyset & \text{otherwise} \end{cases}$$

$$A_{i\ell'}^{\mathcal{I}_{w_{1\ell}}} := \begin{cases} \{d\} & \text{if } i \ne i_0 \text{ or } \ell \ne \ell' \\ \emptyset & \text{otherwise} \end{cases} \qquad A_{i\ell'}^{\mathcal{I}_{w_{2\ell}}} := \begin{cases} \{d\} & \text{if } i = i_0 \text{ or } \ell \ne \ell' \\ \emptyset & \text{otherwise} \end{cases}$$

and for all $1 \le i < j \le k$, $1 \le \ell \le n$:

$$\begin{aligned} B_{ij}^{\mathcal{I}_{w_{1\ell}}} &:= B_{ij}^{\mathcal{I}_{w_{2\ell}}} := \begin{cases} \emptyset & \text{if } i_0 \in \{i, j\} \\ \{d\} & \text{otherwise} \end{cases} \\ \mu(w_{1\ell}) &:= \mu(w_{2\ell}) := \tfrac{1}{2n}. \end{aligned}$$

It is not hard to verify that $\mathcal{I}$ is a model of $\mathcal{T}$. Note that for $1 \leq i \leq k$ and $1 \leq \ell \leq n$, we have

$$p_d^{\mathcal{I}}(A_{i\ell}) = \frac{1}{2} + (n-1) \cdot \frac{1}{2n} = 1 - \frac{1}{2n}$$

since for every $A_{i\ell}$, there is a $b \in \{1,2\}$ such that d satisfies $A_{i\ell}$ in $w_{b\ell'}$ for all ℓ' and in all $w_{(3-b)\ell'}$ whenever $\ell \neq \ell'$. Using the fact that $n > \frac{1}{2(1-p)}$ we get

$$p_d^{\mathcal{I}}(A_{i\ell}) = 1 - \frac{1}{2n} > 1 - \frac{1}{2 \cdot \frac{1}{2(1-p)}} = 1 - \frac{2(1-p)}{2} = p,$$

and thus, $d \in C^{\mathcal{I},v}$ for any $v \in W$. Finally, it is easy to check that $d \notin D_{i_0 j_0}^{\mathcal{I},v}$ for any $v \in W$. □

4.2 Prob-$\mathcal{EL}^{01}$ is Tractable

In analogy with the logic Prob-$\mathcal{ALC}^{01}$ considered in Section 3.3, we define Prob-$\mathcal{EL}^{01}$ to be the fragment of Prob-$\mathcal{EL}$ that only admits probability restrictions of the form $P_{>0}C$ and $P_{=1}C$. Note that none of the logics listed in Theorem 13 is a fragment of Prob-$\mathcal{EL}^{01}$. Our main result in this section is that, in Prob-$\mathcal{EL}^{01}$, KB consequence can be decided in PTime and is thus no more difficult than in the underlying description logic $\mathcal{EL}$.

Theorem 16. *In Prob-$\mathcal{EL}^{01}$, concept subsumption and KB consequence are in* PTime.

In the remainder of this section, we prove Theorem 16 by giving a polynomial time algorithm for KB consequence in Prob-$\mathcal{EL}^{01}$. It is convenient to assume that all involved ABoxes and TBoxes are in a certain normal form, namely that ABoxes take the form

$$\mathcal{A}_0 \wedge \bigwedge_{i=1}^{k} P_{>0}\, \mathcal{A}_i \wedge P_{=1}(\mathcal{A}_{=1})$$

where $\mathcal{A}_0, \dots, \mathcal{A}_k$ and $\mathcal{A}_{=1}$ are conjunctions of concept assertions of the form $A(a)$ (with A a concept name) and of role assertions; and that TBoxes are sets of concept inclusions of the form

$$X_1 \sqcap \cdots \sqcap X_n \sqsubseteq A \quad \text{or} \quad A \sqsubseteq X,$$

where A denotes a concept name and the X_i and X denote *basic concepts*, that is, either $\top$, a concept name, or concepts of the form $P_{\sim n}A$ or $\exists r.A$ with A a concept name.

We briefly argue that any ABox and TBox can be converted into the required form in polynomial time, preserving KB consequence. Recall that, in Prob-$\mathcal{EL}^{01}$, an ABox is built up from concept and role assertions, conjunction, and a probability constructor. Given such an ABox $\mathcal{A}$, we can first remove nested applications of the probability constructor and pull this constructor out of conjunctions by applying the transformations

$$P_{\sim p}P_{\approx q}\mathcal{A} \rightsquigarrow P_{\approx q}\mathcal{A} \qquad P_{\sim p}(\mathcal{A}_1 \wedge \cdots \wedge \mathcal{A}_k \wedge P_{\approx q}\mathcal{A}) \rightsquigarrow P_{\sim p}(\mathcal{A}_1 \wedge \cdots \wedge \mathcal{A}_k) \wedge P_{\approx q}\mathcal{A}$$

and then push conjunctions inside the $P_{=1}$ constructor by applying the transformation

$$P_{=1}\mathcal{A}_1 \wedge P_{=1}\mathcal{A}_2 \rightsquigarrow P_{=1}(\mathcal{A}_1 \wedge \mathcal{A}_2).$$

These transformations are clearly equivalence-preserving. It remains to replace concept assertions $C(a)$ where C is a compound concept with the assertion $A_C(a)$ where A_C is a fresh concept name, and to add $A_C \equiv C$ to the TBox. Let $(\mathcal{T}, \mathcal{A})$ be a KB and $\mathcal{B}$ an ABox, and let $(\mathcal{T}', \mathcal{A}')$ and $\mathcal{B}'$ be the result of the above transformations. It is easy to show that $(\mathcal{T}, \mathcal{A}) \models \mathcal{B}$ iff $(\mathcal{T}', \mathcal{A}') \models \mathcal{B}'$.

We do not give details for establishing the TBox normal form since this is standard for $\mathcal{EL}$-like logics. Indeed, probability restrictions do not pose any additional challenges as they can be treated in the same way as existential restrictions (Baader et al., 2005). Note that the transformation of a TBox $\mathcal{T}$ into a normalized TBox $\mathcal{T}'$ introduces fresh concept names, but as long as the ABoxes $\mathcal{A}$ and $\mathcal{B}$ do not use these concept names, we have $(\mathcal{T}, \mathcal{A}) \models \mathcal{B}$ iff $(\mathcal{T}', \mathcal{A}) \models \mathcal{B}$.

We now develop the announced polynomial time algorithm for KB consequence in Prob-$\mathcal{EL}^{01}$. It be viewed as "consequence-based" in the style of recently proposed efficient procedures (Baader et al., 2005; Kazakov, 2009). Let $\mathcal{K} = (\mathcal{T}, \mathcal{A})$ be a KB such that for some ABox $\mathcal{B}$ (all in normal form), we want to decide whether $\mathcal{K} \models \mathcal{B}$. We show how to build, in polynomial time, a model $\mathcal{I}$ of $\mathcal{K}$ that is *universal* in the sense that for all ABoxes $\mathcal{B}$, we have $\mathcal{K} \models \mathcal{B}$ iff $\mathcal{I}, v_{\mathcal{A}} \models \mathcal{B}$, for a certain fixed world $v_{\mathcal{A}}$. Since the latter can be verified in polynomial time, this yields Theorem 16. Interestingly, we can fix up-front the set of worlds of $\mathcal{I}$, its domain, and its probability distribution μ:

$$
\begin{aligned}
W &= \{v_{\mathcal{A}}, v_{=1}\} \cup \mathsf{Wit}; \\
\Delta^{\mathcal{I}} &= \mathsf{Ind}(\mathcal{A}) \cup (\mathsf{CN} \times W); \\
\mu(w) &= 1/|W| \text{ for all } w \in W.
\end{aligned}
$$

where CN denotes the set of concept names in $\mathcal{K}$ and Wit consists of

- worlds $v_B, v'_B, v_{a,B}, v'_{a,B}$ for all concepts of the form $P_{>0}B$ that occur in $\mathcal{T}$ and all ABox individuals $a \in \mathsf{Ind}(\mathcal{A})$, and
- worlds $v_{\mathcal{A}'}$ for all ABoxes of the form $P_{>0}\mathcal{A}'$ that occur in $\mathcal{A}$.

Each domain element and world has a particular purpose. We call domain elements of the form (A, w) *anonymous* since they are not identified by an individual name in $\mathcal{A}$. An anonymous element (A, w) is used to witness existential restrictions of the form $\exists r.A$ in world w. We are going to make the ABox $\mathcal{A}$ true in world $v_{\mathcal{A}}$. World $v_{=1}$ will make true (only) facts that hold with probability one. And, finally, the worlds in Wit are used as witnesses for the eponymous concepts $P_{>0}B$ and ABoxes $P_{>0}\mathcal{A}_i$.

We start with constructing an interpretation $\mathcal{I}_0 = (\Delta^{\mathcal{I}}, W, (\mathcal{I}_{0w})_{w \in W}, \mu)$ that is a model of $\mathcal{A}$, but not necessarily of $\mathcal{T}$. This is done by setting, for all concept names A, role names r and worlds $w \in W$:

$$
\begin{aligned}
A^{\mathcal{I}_0,w} = {} & \{a \mid A(a) \in \mathcal{A}_0, w = v_{\mathcal{A}}\} \cup \\
& \{a \mid A(a) \in \mathcal{A}_{=1}\} \cup \\
& \{a \mid A(a) \in \mathcal{A}_i \wedge w = v_{\mathcal{A}_i}\};
\end{aligned}
\qquad
\begin{aligned}
r^{\mathcal{I}_0,w} = {} & \{(a,b) \mid r(a,b) \in \mathcal{A}_0, w = v_{\mathcal{A}}\} \cup \\
& \{(a,b) \mid r(a,b) \in \mathcal{A}_{=1}\} \cup \\
& \{(a,b) \mid r(a,b) \in \mathcal{A}_i \wedge w = v_{\mathcal{A}_i}\}.
\end{aligned}
$$

Individual names are interpreted as themselves, that is, $a^{\mathcal{I}_0} = a$. We then *complete* $\mathcal{I}_0$ by exhaustively applying the rules in Figure 1, which ensures that we also satisfy $\mathcal{T}$. We need

R1 If $X_1 \sqcap \cdots \sqcap X_n \sqsubseteq A \in \mathcal{T}$, and $d \in X_i^{\mathcal{I},w}$ for $1 \leq i \leq n$, then $A^{\mathcal{I},w} := A^{\mathcal{I},w} \cup \{d\}$.

R2 If $A \sqsubseteq P_{>0}B \in \mathcal{T}$, $d \in A^{\mathcal{I},w}$, and $d \notin (P_{>0}B)^{\mathcal{I},w}$, then $B^{\mathcal{I},\hat{w}} := B^{\mathcal{I},\hat{w}} \cup \{d\}$ where $\hat{w} = \mathsf{wit}(d,w,B)$.

R3 If $A \sqsubseteq P_{=1}B \in \mathcal{T}$ and $d \in A^{\mathcal{I},w}$ then $B^{\mathcal{I},v} := B^{\mathcal{I},v} \cup \{d\}$ for all $v \in W$.

R4 If $A \sqsubseteq \exists r.B \in \mathcal{T}$ and $d \in A^{\mathcal{I},w}$, then $r^{\mathcal{I},w} := r^{\mathcal{I},w} \cup \{(d,(B,w))\}$ and $B^{\mathcal{I},w} := B^{\mathcal{I},w} \cup \{(B,w)\}$.

Figure 1: Completion rules for Prob-$\mathcal{EL}^{01}$.

to define the witness function $\mathsf{wit}(d,w,B)$ used in Rule **R2**, which returns the world where the concept $P_{>0}B$ is witnessed when it is true at element d in world w:

$$\mathsf{wit}(a,w,B) = \begin{cases} v_{a,B} & \text{if } w \neq v_{a,B}; \\ v'_{a,B} & \text{otherwise;} \end{cases} \qquad \text{and} \qquad \mathsf{wit}((A,v),w,B) = \begin{cases} v_B & \text{if } v \neq v_B; \\ v'_B & \text{otherwise.} \end{cases}$$

The use of primed and non-primed worlds in wit avoids having to realize B in the same world where $P_{>0}B$ appears (which would result in unsound consequences).

The interplay of the rules in Figure 1 is rather subtle and can only be fully appreciated by going through the proof details of Lemma 17 below. Here, we point out some interesting aspects. Note that **R4** chooses *different* elements (A,w_i) to satisfy the same existential restriction $\exists r.A$ in different worlds w_1,w_2. The reason is that, if we would always use the same element d as a witness for $\exists r.A$, then we might 'accidentally' make B true at d with probability one, which would result in unsound consequences. Also note that the witness function makes sure that there is no combination of an element and a world such that this combination is used as a witness both by **R2** and by **R4** (which would also result in unsound consequences). Let us illustrate this for the case of some anonymous element (A,w). Assume that **R2** makes B true at (A,w) in world v to witness a probability restriction $P_{>0}B$. By definition of **R2** and of wit, we have (i) $w \neq v_B$ and $v = v_B$ or (ii) $w = v_B$ and $v = v'_B$. In Case (i), (A,w) cannot be used as a witness by **R4** in world $v = v_B$ since it can only be used as such in world $w \neq v_B$. In Case (ii), (A,w) cannot be used as a witness by **R4** in world $v = v'_B$ since it can only be used as such in world $w = v_B$.

Let $\mathcal{I}$ be the resulting interpretation. Since the completion rules do not introduce additional domain elements or worlds and the number of worlds and domain elements in $\mathcal{I}_0$ is linear in the size of $\mathcal{K}$, it is easy to verify that rule application terminates after polynomially many steps. We show that $\mathcal{I}$ is a universal model of $(\mathcal{T},\mathcal{A})$ in the sense of the following lemma.

Lemma 17. *For all ABoxes $\mathcal{B}$, $(\mathcal{T},\mathcal{A}) \models \mathcal{B}$ iff $\mathcal{I}, v_{\mathcal{A}} \models \mathcal{B}$.*

Proof. For the "$\Rightarrow$" direction, assume that $\mathcal{I}, v_{\mathcal{A}} \not\models \mathcal{B}$. By definition of $\mathcal{I}_0$, we have $\mathcal{I}_0, v_{\mathcal{A}} \models \mathcal{A}$ and, since $\mathcal{I}$ extends $\mathcal{I}_0$, also $\mathcal{I}, v_{\mathcal{A}} \models \mathcal{A}$. Moreover, it is straightforward to verify that $\mathcal{I}$ is a model of $\mathcal{T}$ (if any CI in $\mathcal{T}$ would not be satisfied, there would still be a rule that we could apply). This yields $(\mathcal{T},\mathcal{A}) \not\models \mathcal{B}$.

For the "$\Leftarrow$" direction, the main step is to show by induction on the number of rule applications that a certain set of invariants is preserved. For formulation the invariants, we introduce some notation. For every world $w \in W$, we define an ABox $\mathcal{A}_w$ that consists of the following assertions:

- $X(a) \in \mathcal{A}_w$ if $a \in X^{\mathcal{I},w}$, for all $a \in \mathsf{Ind}(\mathcal{A})$ and basic concepts X;
- $r(a,b) \in \mathcal{A}_w$ if $(a,b) \in r^{\mathcal{I},w}$, for all $a, b \in \mathsf{Ind}(\mathcal{A})$.

We say that an anonymous domain element (A, v) is **R2**-*active in* w if **R2** has added (A, v) to the extension of some B in world w (this can only be the case when $v = w$); similarly, (A, v) is **R4**-*active in* w if **R4** has added (A, v) to the extension of some B in w (this can only be the case when $v \neq w$ and w is a world of the form v_B or v'_B). If none of the above applies, (A, v) is called *inactive in* w. Intuitively, a domain element (A, v) becomes active in w whenever (A, v) witnesses a concept of the form $P_{>0}B$ or $\exists r.B$.

In the appendix, we show that rule application preserves the following invariants, for all Prob-$\mathcal{EL}^{01}$-concepts $C = X_1 \sqcap \ldots \sqcap X_k$, each X_i a basic concept, and all anonymous domain elements $(A, v) \in \Delta^{\mathcal{I}}$:

1. $\mathcal{K} \models \mathcal{A}_{v_\mathcal{A}}$;
2. $\mathcal{K} \models P_{=1}\mathcal{A}_{v_{=1}}$;
3. $\mathcal{K} \models P_{>0}\mathcal{A}_w$ for all $w \in W$;
4. if (A, v) is inactive in w and $(A, v) \in C^{\mathcal{I},w}$, then $\mathcal{T} \models A \sqsubseteq P_{=1}C$;
5. if (A, v) is **R4**-active in v and $(A, v) \in C^{\mathcal{I},v}$, then $\mathcal{T} \models A \sqsubseteq C$;
6. if (A, v) is **R2**-active in $w \in \{v_B, v'_B\}$ and $(A, v) \in C^{\mathcal{I},w}$, then $\mathcal{T} \models B \sqcap P_{>0}A \sqsubseteq C$.

Given these invariants, it is straightforward to show "$\Leftarrow$": Assume $\mathcal{I}, v_\mathcal{A} \models \mathcal{B}$ and $\mathcal{B} = \mathcal{B}_0 \wedge \bigwedge_{i=1}^{\ell} P_{>0}\mathcal{B}_i \wedge P_{=1}\mathcal{B}_{=1}$. Thus, we have $\mathcal{B}_0 \subseteq \mathcal{A}_{v_\mathcal{A}}$, $\mathcal{B}_{=1} \subseteq \mathcal{A}_w$ for all $w \in W$, and for each i, there is some w with $\mathcal{B}_i \subseteq \mathcal{A}_w$. Invariants 1-3 then imply $\mathcal{K} \models \mathcal{B}_0$, $\mathcal{K} \models P_{>0}\mathcal{B}_i$ for all i, and $\mathcal{K} \models P_{=1}\mathcal{B}_{=1}$, and hence $\mathcal{K} \models \mathcal{B}$. $\square$

4.3 Classical TBoxes and Tractability

We have seen in the previous sections that reasoning in Prob-$\mathcal{EL}$ is ExpTime-hard as soon as we admit any single probability restriction of the form $P_{\sim p}$ with $\sim \in \{<, \leq, =, >, \geq\}$ and $p \in (0, 1)$, and that it becomes tractable if we only allow probability restrictions of the form $P_{>0}C$ and $P_{=1}C$. The goal of this section is to show that we can sometimes regain tractability for probability values other than zero and one if we allow only *classical* TBoxes.

We use Prob-$\mathcal{EL}^{\sim p;=1}_{\mathsf{class}}$ to denote the restriction of Prob-$\mathcal{EL}$ obtained by admitting only the probability restrictions $P_{=1}C$ and $P_{\sim p}$, and only classical TBoxes. Our main result is that reasoning is tractable when $\sim \in \{>, \geq\}$ and $p \in (0, 1)$. In this case, concept satisfiability and KB consistency are trivial: in the interpretation $\mathcal{I}$ that has only a single element d and a single world w with $\mu(w) = 1$ and such that $A^{\mathcal{I},w} = \{d\}$ and $r^{\mathcal{I},w} = \{(d, d)\}$ for all concept names A and role names r, all concepts, TBoxes, and ABoxes are satisfied. We thus concentrate on concept subsumption in the following upper bound.

Theorem 18. *For all $\sim \in \{>, \geq\}$ and $p \in (0,1)$, concept subsumption in Prob-$\mathcal{EL}^{\sim p;=1}_{\mathsf{class}}$ is in* PTime.

Recall that a classical TBox $\mathcal{T}$ (subsequently just a TBox) is a finite set of *concept definitions* $A \equiv D$, where A is a concept name and D is a concept description, and additionally every concept name appears at most once on the left-hand side of a definition. A concept name A is *defined* in $\mathcal{T}$ if there is a concept definition $A \equiv C \in \mathcal{T}$, and *primitive* otherwise. We can restrict our attention to the subsumption of defined concept names since $\mathcal{T} \models C \sqsubseteq D$ iff $\mathcal{T} \cup \{A_C \equiv C, A_D \equiv D\} \models A_C \sqsubseteq A_D$, with A_C, A_D fresh concept names. We can also assume that the input TBox is normalized to a set of concept definitions of the form

$$A \equiv P_1 \sqcap \ldots \sqcap P_n \sqcap C_1 \sqcap \ldots \sqcap C_m$$

with $n, m \geq 0$, and where $P_1, \ldots, P_n$ are primitive concept names and $C_1, \ldots, C_m$ are of the form $P_{\sim p}A$, $P_{=1}A$, or $\exists r.A$ with A a defined concept name. Note that the top concept is completely normalized away. It is well-known that such a normalization can be achieved in polynomial time for $\mathcal{EL}$ (Baader, 2003), and probability restrictions do not pose additional challenges as they can be treated like existential restrictions.

Observe first that for $p \in \{0, 1\}$ the theorem is a corollary of the PTime result in the previous section. In the following, we thus assume $0 < p < 1$. We prove the desired PTime upper bound using a consequence-based procedure. In contrast to the algorithm given in the previous section, though, we do not build up an interpretation, but rather complete the TBox with all of its relevant consequences. We start with some notation. For a given TBox $\mathcal{T}$ and a defined concept name A in $\mathcal{T}$, we refer by C_A to the *defining concept* for A in $\mathcal{T}$, i.e., $A \equiv C_A \in \mathcal{T}$. Moreover, we deliberately confuse the concept $C_A = D_1 \sqcap \ldots \sqcap D_k$ with the set $\{D_1, \ldots, D_k\}$. We define the set of concepts that are *certain for* C_A as

$$\mathsf{cert}(C_A) = \{P_*B \mid P_*B \in C_A\} \cup \bigcup_{P_{=1}B \in C_A} \{C_B\}$$

where, here and in what follows, P_* ranges over $P_{=1}$ and $P_{\sim p}$. Intuitively $\mathsf{cert}(C_A)$ contains concepts that hold with probability 1 whenever A is satisfied in some world.

The algorithm starts with the normalized input TBox and then exhaustively applies the completion rules displayed in Figure 2. As a general proviso, each rule can be applied only if it adds a concept that occurs in $\mathcal{T}$ and actually changes the TBox, e.g., **R1** can only be applied when $\exists r.B'$ occurs in $\mathcal{T}$ and $\exists r.B' \notin C_A$. The following lemma establishes the main property of the algorithm.

Lemma 19. *For all defined concept names A, B, we have $\mathcal{T} \models A \sqsubseteq B$ iff, after exhaustive rule application, $C_B \subseteq C_A$.*

Proof. We start with the "$\Leftarrow$"-direction. Using the semantics, it is straightforward to show that the rules are sound, i.e., that if a TBox $\mathcal{T}_2$ is obtained from a TBox $\mathcal{T}_1$ by a single rule application, then $\mathcal{T}_1$ and $\mathcal{T}_2$ are equivalent in the sense that they have the same models. It suffices to show that every model of $\mathcal{T}_1$ is also a model of $\mathcal{T}_2$. To this end, let $\mathcal{I} = (\Delta^\mathcal{I}, W, (\mathcal{I}_w)_{w \in W}, \mu)$ be a model of $\mathcal{T}_1$ and assume $d \in \Delta^\mathcal{I}$ and $w \in W$ such that $d \in A^{\mathcal{I},w}$ and hence $d \in (C_A)^{\mathcal{I},w}$. We show that if a concept D is added to C_A, then we also have $d \in D^{\mathcal{I},w}$. We are then done since this implies $(C_A)^{\mathcal{I},w} = (C_A \sqcap D)^{\mathcal{I},w}$ and because $\mathcal{I}$ is a model of $\mathcal{T}_1$.

R1	If $\exists r.B \in C_A$ and $C_{B'} \subseteq C_B$, then replace $A \equiv C_A$ with $A \equiv C_A \cup \{\exists r.B'\}$.
R2	If $P_{=1}B \in C_A$, then replace $A \equiv C_A$ with $A \equiv C_A \cup C_B$.
R3	If $P_{=1}B \in C_A$, then replace $A \equiv C_A$ with $A \equiv C_A \cup \{P_{\sim p}B\}$.
R4	If $P_{\sim p}B \in C_A$ and $D \in \mathsf{cert}(C_B)$, then replace $A \equiv C_A$ with $A \equiv C_A \cup \{D\}$.
R5	If $C_B \subseteq \mathsf{cert}(C_A)$, then replace $A \equiv C_A$ with $A \equiv C_A \cup \{P_{=1}B\}$.
R6	If $P_{\sim p}B \in C_A$ and $C_{B'} \subseteq \mathsf{cert}(C_A) \cup C_B$, then replace $A \equiv C_A$ with $A \equiv C_A \cup \{P_{\sim p}B'\}$.

Figure 2: TBox completion rules for subsumption in Prob-$\mathcal{EL}^{\sim p;=1}$.

R1 $D = \exists r.B'$. Since $\exists r.B \in C_A$, we have $d \in (\exists r.B)^{\mathcal{I},w}$, i.e., there is some element $e \in \Delta^{\mathcal{I}}$ with $e \in B^{\mathcal{I},w}$ and $(d,e) \in r^{\mathcal{I},w}$. As $\mathcal{I}$ is a model of $\mathcal{T}_1$, we also have $e \in (C_B)^{\mathcal{I},w}$ and, by the assumption $C_{B'} \subseteq C_B$, we obtain $e \in (C_{B'})^{\mathcal{I},w}$. Again, as $\mathcal{T}_1 \models \mathcal{I}$, we have $e \in B'^{\mathcal{I},w}$. The semantics yields $d \in (\exists r.B')^{\mathcal{I},w}$.

R2 $D \in C_B$. Since $P_{=1}B \in C_A$, we have $d \in (P_{=1}B)^{\mathcal{I},w}$. By the semantics, $d \in B^{\mathcal{I},w}$ and thus $d \in (C_B)^{\mathcal{I},w}$.

R3 $D = P_{\sim p}B$. Since $P_{=1}B \in C_A$, we have $d \in (P_{=1}B)^{\mathcal{I},w}$. By the semantics, $d \in (P_{\sim p}B)^{\mathcal{I},w}$.

R4 $D \in C_B$. Since $P_{\sim p}B \in C_A$, we have $d \in (P_{\sim p}B)^{\mathcal{I},w}$. Hence, there is some world $v \in W$ with $d \in B^{\mathcal{I},v}$. By definition of $\mathsf{cert}(C_B)$ and the semantics, we have that $d \in D^{\mathcal{I},w'}$ for all $D \in C_B$ and $w' \in W$. Thus, in particular, $d \in D^{\mathcal{I},w}$ for all $D \in C_B$.

R5 $D = P_{=1}B$. By definition of $\mathsf{cert}(C_A)$ and the semantics, we have that $d \in D^{\mathcal{I},v}$ for all $D \in C_A$ and $v \in W$. As $C_B \subseteq \mathsf{cert}(C_a)$, we have that $d \in (C_B)^{\mathcal{I},v}$ for all $v \in W$. As $\mathcal{I}$ is a model of $\mathcal{T}_1$, we get $d \in B^{\mathcal{I},v}$ for all $v \in W$. The semantics yields $d \in (P_{=1}B)^{\mathcal{I},w}$.

R6 $D = P_{\sim p}B'$. Since $P_{\sim p}B \in C_A$, we have $d \in (P_{\sim p}B)^{\mathcal{I},w}$. By the semantics, $p_d^{\mathcal{I}}(C_B) \sim p$. By definition of $\mathsf{cert}(C_A)$ and the semantics, we have $p_d^{\mathcal{I}}(C_A) = 1$. Together, this implies $p_d^{\mathcal{I}}(C_A \cup C_B) \sim p$. As by assumption $C_{B'} \subseteq C_A \cup C_B$, we also have $p_d^{\mathcal{I}}(C_{B'}) \sim p$ and hence $d \in (P_{\sim p}B')^{\mathcal{I},w}$.

For "$\Rightarrow$", let $\mathcal{T}$ be the TBox after exhaustive application of the rules of Figure 2. We construct a model $\mathcal{I}$ of $\mathcal{T}$ that satisfies a concept inclusion $A \sqsubseteq B$ only if $C_B \subseteq C_A$. Let Def denote the set of defined concept names in $\mathcal{T}$. Moreover, let Sub denote all concepts of the form P (primitive concept name), $\exists r.A$, $P_{=1}A$, and $P_{\sim p}A$ that occur in $\mathcal{T}$. We first fix some constants that will be used to define the probabilities of worlds in the model $\mathcal{I}$:

- first fix $\alpha, \alpha' \in (0,1)$ such that $\frac{\alpha}{2} < \alpha' < \alpha < p$ (this is possible because $p > 0$);
- next fix an integer $m \geq 2$ such that

$$(p - \alpha') + \frac{1 - (p - \alpha' + 3|\mathsf{Def}| \cdot \frac{\alpha}{2})}{m} < p$$

(this can be done simply by choosing m sufficiently large since $p - \alpha' < p$);

- finally, choose an integer $k \geq 2$ such that

$$(p - \alpha') \cdot \frac{k-1}{k} + \alpha > p$$

(this can be done by choosing k sufficiently large since $(p - \alpha') + \alpha > p$).

Start defining an interpretation $\mathcal{I} = (\Delta^{\mathcal{I}}, W, (\mathcal{I}_w)_{w \in W}, \mu)$ by putting:

$$\begin{aligned}
W &= \{\delta_{Ai}, 1_j, p_\ell \mid A \in \mathsf{Def}, i \in \{1,2,3\}, 1 \leq j \leq m, 1 \leq \ell \leq k\} \\
\Delta^{\mathcal{I}} &= \{(A, v) \mid A \in \mathsf{Def}, v \in W\} \\
\mu(p_\ell) &= \frac{p - \alpha'}{k} \\
\mu(\delta_{Ai}) &= \frac{\alpha}{2} \\
\mu(1_j) &= \frac{1 - (p - \alpha' + 3|\mathsf{Def}| \cdot \frac{\alpha}{2})}{m}
\end{aligned}$$

Note first that μ is indeed a probabilitiy distribution: $\mu(\{p_\ell \mid 1 \leq \ell \leq k\}) = p - \alpha'$, $\mu(\{\delta_{Ai} \mid A \in \mathsf{Def}, i \in \{1,2,3\}\}) = 3|\mathsf{Def}|\frac{\alpha}{2}$, and $\mu(\{1_j \mid 1 \leq j \leq m\}) = 1-(p-\alpha'+3|\mathsf{Def}|\frac{\alpha}{2})$; and thus, $\mu(W) = 1$. Moreover, we observe the following two important properties of μ:

(P1) for any set V of worlds that contains at least $k-1$ of the worlds p_ℓ and at least two of the worlds δ_{Ai}, the probabilities sum up to more than p;

(P2) any set of worlds whose probabilities sum up to a value $> p$ includes at least two worlds from $W \setminus \{p_\ell \mid 1 \leq \ell \leq k\}$.

Using the condition in the choice of k, it is easy to see that Property (P1) is satisfied. For Property (P2), define $V = \{p_\ell \mid 1 \leq \ell \leq k\}$ and observe that $\mu(V) = p - \alpha'$. It suffices to show that, when adding a single world to the set of worlds V, the probability will stay below p. By choice of α and α', it is clear that $\mu(V \cup \{\delta_{Ai}\}) = p - \alpha' + \frac{\alpha}{2} < p$. By choice of m, we further have

$$\mu(V \cup \{1_j\}) = p - \alpha' + \frac{(1 - (p - \alpha' + 3|\mathsf{Def}| \cdot \frac{\alpha}{2}))}{m} < p$$

and are done.

To define concept and role memberships, first define a map $\pi : (\Delta^{\mathcal{I}} \times W) \to 2^{\mathsf{Sub}}$ such that each set $\pi(\cdot, \cdot)$ is minimal with the following conditions satisfied for all $A \in \mathsf{Def}$ and $v, w \in W$:

1. $C_A \subseteq \pi((A,w),w)$
2. if $P_* B \in C_A$, then $P_* B \in \pi((A,w),v)$;
3. if $P_{=1} B \in C_A$, then $C_B \subseteq \pi((A,w),v)$;
4. if $P_{>p} B \in C_A$, then $C_B \subseteq \pi((A,w),p_\ell)$ for all $\ell \leq k$ when $w \notin \{p_\ell \mid \ell \leq k\}$;
5. if $P_{>p} B \in C_A$, then $C_B \subseteq \pi((A,p_i),p_\ell)$; for all $\ell \leq k$ with $i \neq \ell$;
6. if $P_{>p} B \in C_A$, then $C_B \subseteq \pi((A,w),\delta_{B1})$ and $C_B \subseteq \pi((A,w),\delta_{B2})$ when $w \notin \{\delta_{B1},\delta_{B2},\delta_{B3}\}$;
7. if $P_{>p} B \in C_A$, then $C_B \subseteq \pi((A,\delta_{Bi}),\delta_{Bj})$ for all distinct $i,j \in \{1,2,3\}$.

Now we simply read off from π the interpretation of the defined concept names A, primitive concept names P, and role names r as

$$\begin{aligned} A^{\mathcal{I},w} &= \{d \in \Delta^{\mathcal{I}} \mid C_A \subseteq \pi(d,w)\} \\ P^{\mathcal{I},w} &= \{d \in \Delta^{\mathcal{I}} \mid P \in \pi(d,w)\} \\ r^{\mathcal{I},w} &= \{(d,(B,w)) \in \Delta^{\mathcal{I}} \times \Delta^{\mathcal{I}} \mid \exists r.B \in \pi(d,w)\}. \end{aligned}$$

Intuitively, by 4 and 5, concepts of the form $P_{>p}B$ are made true in (at least) $k-1$ worlds p_ℓ, and, by 6 and 7, two additional worlds δ_{Ai}, which suffices by (P1). Conversely, the construction ensures that an element (A,w) satisfies B in at most one world from $W \setminus \{p_\ell \mid 1 \leq l \leq k\}$ if $P_{>p}B, P_{=1}B \notin C_A$. Consequently, $(A,w) \notin (P_{>p}B)^{\mathcal{I},v}$, by (P2). Worlds of the form 1_j just fill up the remaining space and, by the choice of m, have sufficiently small probability $\mu(1_j)$ to ensure (P1) and (P2). Based on this intuition, we establish the following claim in Appendix C.

Claim. $(A,w) \in D^{\mathcal{I},v}$ iff $D \in \pi((A,w),v)$ for all $D \in \mathsf{Sub}$, $(A,w) \in \Delta^{\mathcal{I}}$, and $v \in W$.

Using the claim, we now show that $\mathcal{I}$ is a model of $\mathcal{T}$. Let $B \in \mathsf{Def}$; we have to show that $\mathcal{I}$ satisfies $B \equiv C_B$. We have $(A,w) \in B^{\mathcal{I},v}$ iff (by definition of $\mathcal{I}$) $C_B \subseteq \pi((A,w),v)$ iff (by the claim) $(A,w) \in D^{\mathcal{I},v}$ for all $D \in C_B$ iff (by the semantics) $(A,w) \in (C_B)^{\mathcal{I},w}$.

To show that $\mathcal{I}$ satisfies a concept inclusion $A \sqsubseteq B$ only if $C_B \subseteq C_A$, assume that $\mathcal{I} \models A \sqsubseteq B$. By Condition 1 in the definition of π, the definition of $A^{\mathcal{I}}$, and the claim, we have $(A,1_1) \in A^{\mathcal{I},1_1}$. Thus, $(A,1_1) \in B^{\mathcal{I},1_1}$. By definition of $B^{\mathcal{I}}$, we obtain $C_B \subseteq \pi((A,1_1),1_1)$. By choice of π (since Conditions 1–3 only these speak about the world 1_1), it follows that for every $D \in C_B$, we have $D \in C_A$ or, for some B', both $P_{=1}B' \in C_A$ and $D \in C_{B'}$. In the latter case, non-applicability of rule **R2** yields $D \in C_A$. In summary, we thus have $C_B \subseteq C_A$ as required.

This finishes the proof in case $\sim\, = \,>$. For the case $\sim\, = \,\geq$, one can show soundness of the rules **R1**–**R6** along the lines of "$\Leftarrow$" above. For "$\Rightarrow$", we can, in fact, use the *same* interpretation $\mathcal{I}$. For seeing this, observe that the distribution μ underlying $\mathcal{I}$ satisfies the following variations of (P1) and (P2), which suffice to prove the claim above.

(P1′) for any set V of worlds that contains at least $k-1$ of the worlds in $\{p_\ell \mid \ell \leq k\}$ and at least two distinct δ_{Ai}, δ_{Bj} the probabilities sum up to at least p;

(P2′) any set of worlds whose probabilities sum up to a value $\geq p$ includes at least two worlds from $W \setminus \{p_\ell \mid \ell \leq k\}$.

The rest of the proof is then identical. □

The completion algorithm requires only polynomial time since each rule application extends the TBox, and both the number of concept definitions and the number of conjuncts in each concept definition are bounded by the size of the original TBox. This finishes the proof of Theorem 18.

It is interesting to note that the proof of Theorem 18 is based on exactly the same algorithm, for all $\sim \in \{\geq, >\}$ and $p \in (0,1)$. It follows that there is in fact only *one single logic* Prob-$\mathcal{EL}^{\sim p;=1}$, for all such $\sim$ and p. Formally, given a Prob-$\mathcal{EL}^{\sim p;=1}$-concept C, $\approx \in \{\geq, >\}$ and $q \in (0,1)$, let $C_{\approx q}$ denote the result of replacing each subconcept $P_{\sim p}D$ in C with $P_{\approx q}D$ in C and similarly for Prob-$\mathcal{EL}^{\sim p;=1}$-TBoxes $\mathcal{T}$.

Theorem 20. *Let $p, q > 0$, $\sim, \approx \in \{>, \geq\}$. Given a Prob-$\mathcal{EL}^{\sim p;=1}$-TBox $\mathcal{T}$ and concept names A, B, we have $\mathcal{T} \models A \sqsubseteq B$ iff $\mathcal{T}_{\approx q} \models A \sqsubseteq B$.*

Consequently, the (potentially difficult!) choice of a concrete $\sim \in \{\geq, >\}$ and $p \in (0,1)$ is moot. In fact, it might be more intuitive to replace the constructor $P_{\sim p}C$ with a constructor $\mathcal{L}\,C$ that describes elements which 'are likely to be a C', and to replace $P_{=1}C$ with the constructor $\mathcal{C}\,C$ to describe elements that 'are certain to be a C'. Other approaches to logics of likelihood have been proposed, for example, by Halpern and Rabin (1987) and Herzig (2003). Note that the case $p = 0$ is different from the cases considered above: for example, we have

$$\mathcal{T}_\emptyset \models P_{>p}\exists r.A \sqsubseteq P_{>p}\exists r.P_{>p}A$$

if $p = 0$, but not otherwise. In the spirit of the constructors $\mathcal{C}$ and $\mathcal{L}$, $P_{>0}C$ can be replaced with a constructor $\mathcal{P}C$ that describes elements for which 'it is possible that they are a C'. For example, the Snomed CT concepts 'definite thrombus' and 'possible thrombus' can then be written as $\mathcal{C}\,\mathsf{Thrombus}$ and $\mathcal{P}\,\mathsf{Thrombus}$ (although we speculate that the Snomed CT designers mean 'likely' rather than 'possible').

The reader might ask whether it is possible to extend Theorem 18 to $\sim \in \{<, \leq\}$. The following theorem answers this question to the negative.

Theorem 21. *For all $\sim \in \{<, \leq\}$ and $p \in (0,1)$, concept satisfiability in Prob-$\mathcal{EL}^{\sim p}$ is* NP*-hard and concept subsumption is* coNP*-hard, even when the TBox is empty.*

Proof. It suffices to consider concept satisfiability since there is a straightforward reduction to the complement of concept subsumption: C is satisfiable w.r.t. $\mathcal{T}$ iff $\mathcal{T} \not\models C \sqsubseteq A$, A a fresh concept name. We prove NP-hardness of concept satisfiability by a reduction of the satisfiability problem of propositional formulae in conjunctive normal form (CNF).

Let $\varphi = \varphi_1 \wedge \cdots \wedge \varphi_n$ be a propositional formula in CNF with $\varphi_i = \ell_{i1} \vee \cdots \vee \ell_{ik_i}$ where each literal ℓ_{ij} is either a variable x or a negated variable $\neg x$. Assume moreover that the variables in φ are $x_1, \ldots, x_m$ and introduce corresponding concept names $X_1, \ldots, X_m$.

Define the concept C_φ as:

$$C_\varphi = P_{\sim p}C_1 \sqcap \cdots \sqcap P_{\sim p}C_n \qquad \text{with}$$

$$C_i = f(\neg\ell_{i1}) \sqcap \cdots \sqcap f(\neg\ell_{ik_i}) \qquad \text{where } f(\ell) = \begin{cases} P_{\sim p}X_j & \text{if } \ell = x_j; \\ P_{\sim p}(P_{\sim p}X_j) & \text{if } \ell = \neg x_j \end{cases}$$

and double negations cancel out. It can be verified that C_φ is satisfiable iff φ is satisfiable. In particular, if C_φ is satisfied in a model $\mathcal{I}$ at an element d in a world w, then

$$d \in (P_{\sim p}(f(\neg\ell_{i1}) \sqcap \cdots \sqcap f(\neg\ell_{ik_i})))^{\mathcal{I},w}$$

for $1 \leq i \leq n$; since each $f(\neg\ell_{ij})$ is of the form $P_{\sim p}D_{ij}$ and concepts of this form can only have probability zero or one, there must be at least one $f(\neg\ell_{ij})$ that has probability zero and thus $d \notin f(\neg\ell_{ij})^{\mathcal{I},w}$. By choice of f, this implies $d \in f(\ell_{ij})^{\mathcal{I},w}$. We can thus construct a satisfying assignment for φ by assigning true to all variables x such that $d \in f(x)^{\mathcal{I},w}$. The converse direction is left to the reader. □

This result can be explained again by (non-)convexity. The proof of Theorem 21 uses the non-convexity witness for $<, \leq$ from the proof of Theorem 13 (in the definition of $f(\ell)$). In contrast, the non-convexity witnesses for $>, \geq$ from the proof of that theorem cannot be expressed using a classical TBox; in fact we have the following.

Corollary 22. *For all $\sim \in \{>, \geq\}$ and $p \in (0,1)$, Prob-$\mathcal{EL}^{\sim p;=1}_{\mathsf{class}}$ is convex for subsumption, that is, $\mathcal{T} \models C \sqsubseteq D_1 \sqcup D_2$ implies $\mathcal{T} \models C \sqsubseteq D_1$ or $\mathcal{T} \models C \sqsubseteq D_2$ for all classical Prob-$\mathcal{EL}^{\sim p;=1}$-TBoxes $\mathcal{T}$ and Prob-$\mathcal{EL}^{\sim p;=1}$-concepts C, D_1, D_2.*

Proof. As in the case of deciding concept subsumption, we can assume that C, D_1, D_2 are concept names, see the remark after Theorem 18. Assume that $\mathcal{T} \models C \sqsubseteq D_1 \sqcup D_2$. An inspection of the proof of the completeness part of Lemma 19 yields, for every defined concept name A, the existence of a probabilistic interpretation $\mathcal{I} = (\Delta^{\mathcal{I}}, W, (\mathcal{I}_w)_{w \in W}, \mu)$, a domain element $d_A \in \Delta^{\mathcal{I}}$ and a world $w_A \in W$ with $\mu(w_A) > 0$ such that $d_A \in A^{\mathcal{I},w_A}$ and for all defined concept names B, we have $d_A \in B^{\mathcal{I},w_A}$ if and only if $\mathcal{T} \models A \sqsubseteq B$. Choosing $A = C$, we obtain $d_A \in C^{\mathcal{I},w_A}$, but $d_A \notin D_1^{\mathcal{I},w_A}$ and $d_A \notin D_2^{\mathcal{I},w_A}$, which witnesses $\mathcal{T} \not\models C \sqsubseteq D_1 \sqcup D_2$. □

4.4 Prob-$\mathcal{EL}$ with Probabilistic Roles

We study the extension Prob-$\mathcal{EL}_r$ of Prob-$\mathcal{EL}$ that admits probabilistic roles. Using the non-convexity results for Prob-$\mathcal{EL}$ from Section 4.1 and the undecidability results for Prob-$\mathcal{ALC}_r$ from Section 3.3, it is not hard to exhibit a polynomial time reduction from concept subsumption in Prob-$\mathcal{ALC}_r$ to concept subsumption in Prob-$\mathcal{EL}_r$, following the strategy from Appendix A.3. Since the decidability status of the former remains open, we cannot hope to clarify decidability of the latter. Inspired by the positive results from Section 3.3, we thus concentrate on the fragment Prob-$\mathcal{EL}^{01}_r$ of Prob-$\mathcal{EL}_r$, where only the probability restrictions $P_{>0}C$ and $P_{=1}C$ are admitted, and only the probabilistic roles $P_{>0}r$ and $P_{=1}r$ are allowed.

We start with establishing a lower bound which shows that, unlike in the case without probabilistic roles, we cannot hope for tractability. As a preparation for the lower bound proof, we observe that, just like Prob-$\mathcal{ALC}_r^{01}$, also Prob-$\mathcal{EL}_r^{01}$ does not enjoy the finite model property.

Lemma 23. *Prob-$\mathcal{EL}_r^{01}$ lacks the FMP.*

Proof. Let

$$\mathcal{T} = \{ \begin{array}{rcl} \top & \sqsubseteq & P_{>0}A \\ \top & \sqsubseteq & \exists P_{=1}r.\top \\ & & \end{array} \quad \begin{array}{rcl} \exists r.A & \sqsubseteq & A' \\ \exists r.A' & \sqsubseteq & A' \\ & & \end{array} \quad \begin{array}{rcl} A \sqcap A' & \sqsubseteq & B \\ \exists r.B & \sqsubseteq & B \\ P_{>0}B & \sqsubseteq & B \; \} \end{array}$$

Then we have $\mathcal{T} \not\models \top \sqsubseteq B$. In fact, there is a model $\mathcal{I}$ of $\mathcal{T}$ with $B^{\mathcal{I}} = \emptyset$: $\mathcal{I}$ contains infinitely many elements arranged in an r-chain of probability one, that is, $\Delta^{\mathcal{I}} = \{d_0, d_1, \dots\}$ and $p^{\mathcal{I}}_{d_i,d_{i+1}}(r) = 1$ for all $i \geq 0$; moreover, $\mathcal{I}$ has infinitely many worlds $w_0, w_1, \dots$ such that $A^{\mathcal{I},w_i} = \{d_i\}$. If A' is interpreted in a minimal way, then $(A \sqcap A') = \emptyset$, and thus we can choose $B^{\mathcal{I}} = \emptyset$. However, if we take a model $\mathcal{J}$ of $\mathcal{T}$ that has only a finite number of domain elements, then $\mathcal{J}$ must contain an r-cycle which implies $(A \sqcap A') \neq \emptyset$, thus $B^{\mathcal{J}} \neq \emptyset$; similarly, if $\mathcal{J}$ has only a finite number of worlds, then there must be an infinite r-chain that satisfies A in more than one world (in fact, in infinitely many), which again implies $(A \sqcap A') \neq \emptyset$. □

The proof of the following hardness result can be considered an extension of the proof of Lemma 23. We confine ourselves with a sketch here and give details in Appendix D.

Theorem 24. *In Prob-$\mathcal{EL}_r^{01}$, concept subsumption is* PSPACE*-hard even if probability restrictions $P_{=1}C$ are disallowed and only one of $\exists P_{=1}r.C$ and $\exists P_{>0}r.C$ is admitted.*

Proof (sketch). We concentrate on the case where $\exists P_{=1}r.C$ is admitted. The proof is by reduction of the word problem of deterministic, polynomially space-bounded Turing machines. Let M be such a machine, x an input to M of length n, and $m = p(n)$ the space bound of M on x. We construct a TBox $\mathcal{T}$ and concept C_0 such that $\mathcal{T} \models \top \sqsubseteq C_0$ iff M accepts x. The basic idea is that each model $\mathcal{I}$ of $\mathcal{T}$ will contain an infinite r-chain of probability one as in the proof of Lemma 23. For every element on the path, there will be a world w such that the concept memberships of d at w represent the initial configuration of M on x. When going backwards in the chain but staying in the world w, as in the propagation of A' in the proof of Lemma 23, the concept memberships evolve according to the computation of M on x. Since computations start at arbitrarily far distance from the root of the chain, it follows that for each configuration c that is encountered during the computation, there is a world w where the concept memberships of the root represent c. It is then easy to use the probability restiction $P_{>0}C$ in C_0 to check whether any of these configurations is accepting. □

The main result of this section is that the lower bound given in Theorem 24 is tight, that is, concept subsumption in Prob-$\mathcal{EL}_r^{01}$ is PSPACE-complete. To prove the upper bound, we once more devise a consequence-based procedure. It is once more presented in a different way: while the procedure in Section 4.2 constructs a model and the one in Section 4.3

completes the TBox, for the current goal it is more convenient to build a dedicated data structure. Another difference to the previous sections is that the algorithm here is non-deterministic, relying on Savitch's theorem for determinization. We assume that the TBox is in the same normal form as in Section 4.2. Thus, a *basic concept* is a concept of the form $\top$, A, $P_{>0}A$, $P_{=1}A$, or $\exists\alpha.A$, where A is a concept name and, here and in what follows, α is a *role*, i.e., of the form r, $P_{>0}r$, or $P_{=1}r$ with r a role name. In the latter two cases, we call the role α a *probabilistic* role. Every concept inclusion in the input TBox is required to be of the form $X_1 \sqcap \ldots \sqcap X_n \sqsubseteq X$ with $X_1, \ldots, X_n, X$ basic concepts.

Let $\mathcal{T}$ be the input TBox in normal form, CN the set of concept names that occur in $\mathcal{T}$, BC the set of basic concepts in $\mathcal{T}$, and ROL the set of roles in $\mathcal{T}$. Our algorithm maintains the following data structures:

- a mapping Q that associates with each $A \in \mathsf{CN}$ a subset $Q(A) \subseteq \mathsf{BC}$ such that $\mathcal{T} \models A \sqsubseteq X$ for all $X \in Q(A)$;
- a mapping Q_{cert} that associates with each $A \in \mathsf{CN}$ a subset $Q_{\mathsf{cert}}(A) \subseteq \mathsf{BC}$ such that $\mathcal{T} \models A \sqsubseteq P_{=1}X$ for all $X \in Q_{\mathsf{cert}}(A)$;
- a mapping R that associates with each probabilistic role $\alpha \in \mathsf{ROL}$ a binary relation $R(\alpha)$ on CN such that $\mathcal{T} \models A \sqsubseteq P_{>0}(\exists\alpha.B)$ for all $(A, B) \in R(\alpha)$.

The above mappings can also be understood as an abstract description of a model of $\mathcal{T}$: each set $Q(A)$ lists mandatory concept memberships of a domain element d in a world w that satisfies $d \in A^{\mathcal{I},w}$, each set $Q_{\mathsf{cert}}(A)$ list mandatory concept memberships that hold with probability one for every domain element that satisfies A in some world, and R describes the structure of the probabilistic roles in the following sense: when $(A, B) \in R(\alpha)$, then $d \in A^{\mathcal{I},w}$ implies that there is an element e such that e is an α-successor of d (with probability one) and, in some world v, satisfies all concepts in $Q(B)$. Note that non-probabilistic roles are not represented in $R(\cdot)$; we will treat them in a more implicit way later on.

The data structures are initialized as follows, for all $A \in \mathsf{CN}$ and probabilistic roles α:

$$Q(A) = \{\top, A\} \qquad Q_{\mathsf{cert}}(A) = \{\top\} \qquad R(\alpha) = \emptyset.$$

The sets $Q(\cdot)$, $Q_{\mathsf{cert}}(\cdot)$, and $R(\cdot)$ are then repeatedly extended by the application of several rules. Before we can introduce these rules, we need some preliminaries. As the first step, Figure 3 presents a set of rules that serves the purpose of saturating a set of concepts Γ; these rules are not directly applied to our data structure, but they will later on serve an important purpose in its completion. We use $\mathsf{cl}(\Gamma)$ to denote the set of concepts that is the result of exhaustively applying the displayed rules to Γ, where any rule can only be applied if the added concept is in BC, but not yet in Γ. It is not hard to see that rule application terminates after polynomially many steps.

The rules that are used for completing the data structures $Q(\cdot)$, $Q_{\mathsf{cert}}(\cdot)$, and $R(\cdot)$ are more complex and refer to 'backtraces' through these data structures.

Definition 25. A *backtrace to A_n* is a finite sequence $S, A_1, \alpha_2, A_2, \ldots, \alpha_n, A_n$ where

1. each $A_i \in \mathsf{CN}$ and each $\alpha_i \in \mathsf{ROL}$ is a probabilistic role;

R1	If $X_1 \sqcap \ldots \sqcap X_n \sqsubseteq X \in \mathcal{T}$ and $X_1, \ldots, X_n \in \Gamma$, then add X to Γ.
R2	If $P_{=1}A \in \Gamma$, then add A to Γ.
R3	If $A \in \Gamma$, then add $P_{>0}A$ to Γ.
R4	If $\exists P_{=1}r.A \in \Gamma$, then add $\exists r.A$ and $\exists P_{>0}r.A$ to Γ.
R5	If $\exists r.A \in \Gamma$, then add $\exists P_{>0}r.A$ to Γ.
R6	If $\exists \alpha.A \in \Gamma$ and $B \in Q(A)$, then add $\exists \alpha.B$ to Γ.

Figure 3: Saturation rules for $\mathsf{cl}(\Gamma)$.

2. $S = B$ for some $P_{>0}B \in Q(A_1)$ or $S = (r, B)$ for some $(A_1, B) \in R(P_{>0}r)$;
3. $(A_i, A_{i-1}) \in R(\alpha_i)$ for $1 < i \leq n$.

If t is a backtrace of length n and $k \leq n$, we use t_k to denote the backtrace $S, A_1, \alpha_2, \ldots, \alpha_k, A_k$.

The purpose of a backtrace is illustrated in Figure 4. Assume we are interested in basic concepts that must be true at the element d_n in world w_1, and that we know from the $R(\cdot)$ data structure that d_i must be connected to d_{i-1} via the probabilistic role α_i (in *every* world), for $1 < i \leq n$, and that $Q(A_1)$ tells us that d_1 must be in $P_{>0}B$. Then there is a world w_2 where d_1 is in B and via CIs such as $\exists \alpha_i.X \sqsubseteq Y$, this might enforce additional concepts to be true at the elements $d_2, \ldots, d_n$ in w_2 (c.f. the Turing machine computations in the proof of Theorem 24). Now, new concepts at d_n in w_2 might imply additional concepts to be true at d_n in w_1 via CIs of the form $P_{>0}X \sqsubseteq Y$, and the purpose of backtraces is to take care of exactly such memberships. In this example, the world w_2 was generated by a concept of the form $P_{>0}B$, which is the first case of Point 2 of Definition 25. The second case is that w_2 is generated by a concept of the form $\exists P_{>0}r.A$. The backwards propagation along the $d_n, \ldots, d_1$ chain in w_2, from which backtraces get their name, is captured by the following notion. Note that the rules **R1** to **R6** that we had already introduced are used in every step of this inductive definition.

Definition 26. Let $t = S, A_1, \alpha_2, \ldots, \alpha_n, A_n$ be a backtrace of length n. Then the *type* $\Gamma(t) \subseteq \mathsf{BC}$ *of* t is

$$\Gamma(t) = \begin{cases} \mathsf{cl}(Q_{\mathsf{cert}}(A_1) \cup \{B\}) & \text{if } t = B, A_1; \\ \mathsf{cl}(Q_{\mathsf{cert}}(A_1) \cup \{\exists r.B' \in \mathsf{BC} \mid B' \in Q_{\mathsf{cert}}(B)\}) & \text{if } t = (r, B), A_1; \\ \mathsf{cl}(Q_{\mathsf{cert}}(A_n) \cup \{\exists \alpha_n.B' \in \mathsf{BC} \mid B' \in \Gamma(t_{n-1})\}) & \text{if } n > 1. \end{cases}$$

In terms of Figure 4, $\Gamma(t)$ is the set of concepts that we know must be true at element d_n in world w_2.

We now return to the description of the algorithm. Figure 5 shows the rules used for completing the data structures $Q(\cdot)$, $Q_{\mathsf{cert}}(\cdot)$, and $R(\cdot)$. Rules **S1** to **S5** are rather straightforward to understand. Rules **S6** and **S7** implement the propagation of information

$$
\begin{tikzcd}
 & A_n & A_{n-1} & & A_2 & A_1 & \\
w_1 & d_n \arrow[r, "\alpha_n"'] & d_{n-1} \arrow[r, "\alpha_{n-1}"'] & \cdots \arrow[r, "\alpha_3"'] & d_2 \arrow[r, "\alpha_2"'] & d_1 & P_{>0}B \arrow[d, bend left] \\
w_2 & d_n \arrow[r, "\alpha_n"'] & d_{n-1} \arrow[r, "\alpha_{n-1}"'] & \cdots \arrow[r, "\alpha_3"'] & d_2 \arrow[r, "\alpha_2"'] & d_1 & B
\end{tikzcd}
$$

Figure 4: Backtrace illustration.

S1 Apply **R1-R6** to $Q(A)$ and $Q_{\mathsf{cert}}(A)$.
S2 If $P_{\sim p}B \in Q(A)$,
then add $P_{\sim p}B$ to $Q_{\mathsf{cert}}(A)$.
S3 If $C \in Q_{\mathsf{cert}}(A)$,
then add $P_{=1}C$ and C to $Q(A)$.
S4 If $\exists\alpha.B \in Q(A)$ with α a probabilistic role,
then add (A, B) to $R(\alpha)$.
S5 If $(A_1, A_2) \in R(\alpha)$ and $B \in Q_{\mathsf{cert}}(A_2)$,
then add $\exists\alpha.B$ to $Q_{\mathsf{cert}}(A_1)$.
S6 If t is a backtrace to B and $P_{\sim p}A \in \Gamma(t)$,
then add $P_{\sim p}A$ to $Q_{\mathsf{cert}}(B)$.
S7 If t is a backtrace to B and $\exists\alpha.A \in \Gamma(t)$ with α a probabilistic role,
then add (B, A) to $R(\alpha)$.

Figure 5: The rules for completing the data structure.

along backtraces, as described above. The algorithm for deciding subsumption starts with the initial data structures defined above and then exhaustively applies the rules shown in Figure 5. To decide whether $\mathcal{T} \models A \sqsubseteq B$, it then simply checks whether $B \in Q(A)$ after the algorithm terminated.

Lemma 27. *For all concept names A_0, B_0, $\mathcal{T} \models A_0 \sqsubseteq B_0$ iff, after exhaustive rule application, $B_0 \in Q(A_0)$.*

Proof. For the "$\Leftarrow$" direction, we show that the algorithm preserves the following invariants of the data structures:

$$C \in Q(A) \text{ implies } \mathcal{T} \models A \sqsubseteq C \quad \textbf{(inv1)}$$
$$C \in Q_{\mathsf{cert}}(A) \text{ implies } \mathcal{T} \models A \sqsubseteq P_{=1}C \quad \textbf{(inv2)}$$
$$(A, B) \in R(\alpha) \text{ implies } \mathcal{T} \models A \sqsubseteq P_{>0}(\exists\alpha.B) \quad \textbf{(inv3)}$$

The proof is by a straightforward induction on the number of applications of the rules in Figure 5, details are given in Appendix E. A crucial step is to realize that for every backtrace

t to B, we have $\mathcal{T} \models B \sqsubseteq P_{>0}(\bigsqcap \Gamma(t))$.[5] Assume now that $B_0 \in Q(A_0)$. Invariant (**inv1**) implies $\mathcal{T} \models A_0 \sqsubseteq B_0$ which finishes the proof of "$\Leftarrow$".

For "$\Rightarrow$", we construct a probabilistic model $\mathcal{I} = (\Delta^{\mathcal{I}}, W, (\mathcal{I}_w)_{w \in W})$ of $\mathcal{T}$ such that there is a world $w \in W$ and a domain element $d \in \Delta^{\mathcal{I}}$ with $d \in A_0^{\mathcal{I},w}$ but $d \notin B_0^{\mathcal{I},w}$. The construction of $\mathcal{I}$ proceeds in a step-wise fashion, in each step extending the domain and/or the set of worlds. We thus define sequences $\Delta^{\mathcal{I}_0}, \Delta^{\mathcal{I}_1}, \dots,$ $W_0, W_1, \dots$ along with partial maps $\pi_0, \pi_1, \dots,$ $\pi_i : \Delta^{\mathcal{I}_i} \times W_i \to 2^{\mathsf{BC}}$, that record desired memberships in basic concepts; role memberships will later be reconstructed from the names of domain elements. The final sets $\Delta^{\mathcal{I}}$ and W are then obtained in the limit. The elements of each set $\Delta^{\mathcal{I}_i}$ are sequences of triples (α, w, A) where $\alpha \in \mathsf{ROL}$ is a role, $w \in W_i$ is a world, and $A \in \mathsf{CN}$ is a concept name. For such a sequence σ, we use σ_j to denote the prefix of σ that consists of the first j triples. A domain element of the form $\sigma \cdot (\alpha, w, A)$ serves the purpose of witnessing the existential restriction $\exists \alpha'.A'$ at σ in world w.

Also each world of $\mathcal{I}$ serves a clear purpose and, intuitively, corresponds to a backtrace. In particular, (most) worlds will be of the form (σ, S) for some $\sigma \in \Delta^{\mathcal{I}}$ and S either B or (r, B) for some concept name B and role name r. For establishing the close correspondence between worlds and backtraces, we define a function δ that maps a world (σ, S) to the sequence $S, A_n, \hat{\alpha}_n, \dots, \hat{\alpha}_2, A_1$ where $\sigma = (\alpha_1, w_1, A_1) \cdots (\alpha_n, w_n, A_n)$ and $\hat{\alpha} = \alpha$ if α is a probabilistic role and $\hat{\alpha} = P_{>0}r$ if α is the role name r. We will show that the world w witnesses the existence of the backtrace $\delta(w)$. To start the construction of $\mathcal{I}$, put

- $\Delta^{\mathcal{I}_0} = \{(\alpha, \varepsilon, A_0)\}$ where α is any role (not important) and A_0 is the concept name from the left-hand side of the subsumption which is to be checked;
- $W_0 = \{\varepsilon, 0\}$,
- $\pi((\alpha, \varepsilon, A_0), \varepsilon) = Q(A_0)$ and $\pi((\alpha, \varepsilon, A_0), 0) = Q_{\mathsf{cert}}(A_0)$.

For the induction step, we start by putting $\Delta^{\mathcal{I}_i} = \Delta^{\mathcal{I}_{i-1}}$, $W_i = W_{i-1}$, and $\pi_i = \pi_{i-1}$, and then apply the following rules:

1. If $\exists \alpha.A \in \pi_i(\sigma, w)$ for some $\sigma \in \Delta^{\mathcal{I}_i}$ and $w \in W_i$, then put $\sigma' := \sigma \cdot (\alpha, w, A)$ and
   (a) add σ' to $\Delta^{\mathcal{I}_i}$ if it does not exist yet, and in this case:
   (b) put $\pi_i(\sigma', w) = Q(A)$ and $\pi_i(\sigma', v) = Q_{\mathsf{cert}}(A)$ for all $v \in W_i \setminus \{w\}$.
2. If $\sigma = (\alpha_1, w_1, A_1) \cdots (\alpha_n, w_n, A_n) \in \Delta^{\mathcal{I}_i}$ and $P_{>0}B \in Q(A_n)$, then
   (a) add (σ, B) to W_i if it does not exist yet, and in this case:
   (b) put $\pi_i(\sigma_j, (\sigma, B)) = \Gamma(\delta(\sigma, B)_{n-j+1})$ for all $1 \leq j \leq n$; and
   (c) put $\pi_i(\sigma' \cdot (\alpha, w, A), (\sigma, B)) = Q_{\mathsf{cert}}(A)$ for all $\sigma' \cdot (\alpha, w, A) \in \Delta^{\mathcal{I}_i}$ that are *not* a prefix of σ.
3. If $\sigma = (\alpha_1, w_1, A_1) \cdots (\alpha_n, w_n, A_n) \in \Delta^{\mathcal{I}_i}$ and $(A_n, B) \in R(P_{>0}r)$, then

5. We denote with $\bigsqcap X$ the conjunction of all concepts in X.

(a) add (σ, r, B) to W_i if it does not exist yet, and in this case:

(b) put $\pi_i(\sigma_j, (\sigma, r, B)) = \Gamma(\delta(\sigma, (r, B))_{n-j+1})$ for all $1 \leq j \leq n$; and

(c) put $\pi_i(\sigma' \cdot (\alpha, w, A), (\sigma, r, B)) = Q_{\mathsf{cert}}(A)$ for all $\sigma' \cdot (\alpha, w, A) \in \Delta^{\mathcal{I}_i}$ that are *not* a prefix of σ.

Note that once $\pi(\sigma, w)$ was set via the application of some rule, it will not be reset by a later rule application. It is also easy to verify that each pair σ, w is assigned a set $\pi(\sigma, w)$.

Finally, put $\Delta^{\mathcal{I}} = \bigcup_{i \geq 0} \Delta^{\mathcal{I}_i}$, $W = \bigcup_{i \geq 0} W_i$, and $\pi = \bigcup_{i \geq 0} \pi_i$. Define μ such that $\mu(w) > 0$ for all $w \in W$ and $\sum_{w \in W} \mu(w) = 1$. If W is finite this is clearly possible; otherwise assign the probabilities 1/2, 1/4, 1/8, ... to (an enumeration of) the worlds. It remains to define the interpretation of concept and role names:

$$
\begin{aligned}
A^{\mathcal{I},w} &= \{\sigma \in \Delta^{\mathcal{I}} \mid A \in \pi(\sigma, w)\} \\
r^{\mathcal{I},w} &= \{(\sigma, \sigma \cdot (P_{>0} r, v, A)) \mid \sigma \cdot (P_{>0} r, v, A) \in \Delta^{\mathcal{I}}, w = (\sigma, r, A)\} \cup \\
&\qquad \{(\sigma, \sigma \cdot (r, w, A)) \mid \sigma \cdot (r, w, A) \in \Delta^{\mathcal{I}}\} \cup \\
&\qquad \{(\sigma, \sigma \cdot (P_{=1} r, v, A)) \mid \sigma \cdot (P_{=1} r, v, A) \in \Delta^{\mathcal{I}}\}
\end{aligned}
$$

This finishes the model construction. For finishing the "$\Rightarrow$"-direction, we need the following Claim, proved in Appendix E.

Claim. For all $\sigma \in \Delta^{\mathcal{I}}$, $w \in W$, and $C \in \mathsf{BC}$, we have $\sigma \in C^{\mathcal{I},w}$ iff $C \in \pi(\sigma, w)$.

To see that $\mathcal{I}$ is a model of $\mathcal{T}$, assume $X_1 \sqcap \ldots \sqcap X_n \sqsubseteq X \in \mathcal{T}$, and let $\sigma \in \Delta^{\mathcal{I}}$ and $w \in W$ such that $\sigma \in X_i^{\mathcal{I},w}$ for all $1 \leq i \leq n$. Due to the Claim, we have $X_i \in \pi(\Sigma, w)$ for all i. By construction, $\pi(\sigma, w)$ is closed under cl for all σ, w. Thus, rule **R1** implies $X \in \pi(\sigma, w)$, and by the Claim, we have $\sigma \in X^{\mathcal{I},w}$.

Consider now $\sigma_0 = (\alpha, \varepsilon, A_0)$ and $w_0 = \varepsilon$. By construction, we have that $\pi(\sigma_0, w_0) = Q(A_0)$ and thus, by assumption, $A_0 \in \pi(\sigma, w_0)$ but $B_0 \notin \pi(\sigma_0, w_0)$. The Claim now yields $\sigma \in A_0^{\mathcal{I},w_0}$ but $\sigma \notin B_0^{\mathcal{I},w_0}$. □

We now argue that the algorithm can be implemented using only polynomial space. First, it is easy to see that there can be only polynomially many rule applications: every rule application extends the data structures $Q(\cdot)$, $Q_{\mathsf{cert}}(\cdot)$, and $R(\cdot)$, but these structures consist of polynomially many sets, each with at most polynomially many elements. It thus remains to verify that each rule application can be executed using only polynomial space. This is obvious for rules **R1**-**R6** and **S1**-**S5**, but not completely so for the rules **S6** and **S7**, which involve backtraces. Since the two rules are very similar, we concentrate on **S6**. All we need to know for applying the rule is whether there is a backtrace t to B such that $P_{\sim p} A \in \Gamma(t)$. We will show below that we can concentrate on backtraces t whose length is bounded by $|\mathcal{T}| \cdot 2^{|\mathcal{T}|}$. We can thus verify the existence of the desired backtrace $t = S, A_1, \alpha_2, A_2, \ldots, \alpha_n, A_n$ by guessing it step-by-step, from S towards A_n, simultaneously computing the types $\Gamma(t_i)$. At each time, we keep only a single pair A_i, α_i and the single type $\Gamma(t_i)$ in memory. When we reach a situation where A_i is the concept name B that we are interested in for applying **S6** and $\Gamma(t_i)$ contains the concept $P_{\sim p} A$ that **S6** looks for, then we can stop guessing and apply the rule. We also maintain a binary counter and stop in a rejecting configuration when the length of the guessed trace has exceeded $|\mathcal{T}| \cdot 2^{|\mathcal{T}|}$. By Savitch's theorem, this non-deterministic algorithm can be determinized.

Proposition 28. *If there is a backtrace t to B with type $\Gamma(t)$, then there is a backtrace $\hat{t}$ to B with $\Gamma(\hat{t}) = \Gamma(t)$ and length at most $M := |\mathcal{T}| \cdot 2^{|\mathcal{T}|}$.*

Proof. Let $t = S, A_1, \alpha_2, \ldots, \alpha_n, A_n$ be a backtrace to A_n with $n > M$. Consider the sequence $(A_1, \Gamma(t_1)), \ldots, (A_n, \Gamma(t_n))$ of concept names with their corresponding types. Note that there are at most $2^{|\mathcal{T}|}$ possible types and at most $|\mathcal{T}|$ concept names. Since $n > M$, the pigeon hole principle implies that there are $1 \leq i < j \leq n$ with $A_i = A_j$ and $\Gamma(t_i) = \Gamma(t_j)$. It should be clear that the sequence

$$t' = S, A_1, \alpha_2, \ldots, \alpha_i, A_i, \alpha_{j+1}, A_{j+1}, \ldots, A_n$$

is, in fact, a backtrace to B and $\Gamma(t') = \Gamma(t)$. Obviously, t' is shorter than t. If the length of t' is at most M, we are done; otherwise, put $t := t'$ and repeat the above steps. $\square$

Summing up, we have obtained the following result.

Theorem 29. *In Prob-$\mathcal{EL}_r^{01}$, concept subsumption is* PSpace*-complete.*

As a byproduct, the proof of Lemma 27 yields convexity of Prob-$\mathcal{EL}_r^{01}$.

Corollary 30. *Prob-$\mathcal{EL}_r^{01}$ is convex for subsumption, that is, $\mathcal{T} \models C \sqsubseteq D_1 \sqcup D_2$ implies $\mathcal{T} \models C \sqsubseteq D_1$ or $\mathcal{T} \models C \sqsubseteq D_2$.*

Proof. We prove the contrapositive. Take a TBox $\mathcal{T}$ and concept names C, D_1, and D_2 such that $\mathcal{T} \not\models C \sqsubseteq D_1$ and $\mathcal{T} \not\models C \sqsubseteq D_2$. By Lemma 27, we get that $D_1 \notin Q(C)$ and $D_2 \notin Q(C)$. The model $\mathcal{I} = (\Delta^{\mathcal{I}}, W, (\mathcal{I}_w)_{w \in W}, \mu)$ of $\mathcal{T}$ constructed in the proof of Lemma 27 features an individual $d \in \Delta^{\mathcal{I}}$ and a world w with $\mu(w) > 0$ and $d \in D^{\mathcal{I},w}$ iff $D \in Q(C)$ for all basic concepts D. Thus, we have $d \in C^{\mathcal{I},w}$ but $d \notin D_1^{\mathcal{I},w}$ and $d \notin D_2^{\mathcal{I},w}$ and hence $d \notin (D_1 \sqcup D_2)^{\mathcal{I},w}$. This proves that $\mathcal{T} \not\models C \sqsubseteq D_1 \sqcup D_2$. $\square$

5. Related Work

There is a large number of proposals for probabilistic DLs that differ widely in several fundamental aspects, among them the following: Is the aim to represent statistical or subjective probabilities (or both)? Is a probabilistic KB associated with one single probability distribution or does it only specify constraints on the set of admissible probability distributions? Are probabilities applied to concepts, roles, TBox statements, or ABox assertions? What type of probabilistic independence is supported? The decisions that we have taken in the design of Prob-$\mathcal{ALC}$ were guided by two aims: to satisfy the needs of modelling uncertainty of domain concepts in ontologies such as Snomed CT and to avoid the intricacies of non-monotonicity. Because of the many proposals for probabilistic DLs and their huge divergence regarding the mentioned aspects, it is impossible to give a fully detailed account of all related work. Instead, we mention the most relevant lines of research and refer the reader also to the survey papers by Lukasiewicz and Straccia (2008) and de Salvo Braz, Amir, and Roth (2008).

An important feature of Prob-$\mathcal{ALC}$ KBs is that they are not associated with one single probability distribution, but only impose constraints on the set of admissible distributions. This approach was pioneered by Fagin, Halpern, and Megiddo (1990) and has first been used in a DL context by Heinsohn (1994) and Jaeger (1994), who, however, concentrate mainly on statistical rather than subjective probabilities. In contrast to these proposals, many other lines of research on probabilistic DLs associate a probabilistic KB with a single probability distribution. In particular, this is the case for probabilistic DLs that combine a classical description logic with a Bayesian network or some other graphical model (Koller, Levy, & Pfeffer, 1997; Yelland, 2000; d'Amato, Fanizzi, & Lukasiewicz, 2008; Ceylan & Peñaloza, 2014; Mauá & Cozman, 2015; Cozman, Polastro, Takiyama, & Revoredo, 2014). Another widely used way to impose a single probability distribution is to include strong independence assumptions in the semantics, instead of specifying them in a graphical model (Jung & Lutz, 2012; Riguzzi, Bellodi, Lamma, & Zese, 2015; Peñaloza & Potyka, 2016). The latter proposals, and several others among the ones already mentioned, additionally differ from Prob-$\mathcal{ALC}$ in that the probability distribution is over possible *knowledge bases* instead of over possible *worlds*. Moreover, all the approaches to probabilistic DLs mentioned above allow the application of probabilities only to TBox statements and to ABox assertions, but not to concept and role names. We emphasize that this results in rather different expressive power than what is offered by Prob-$\mathcal{ALC}$: as illustrated in Section 2, applying probabilities to concepts and roles is essential for modelling uncertainty of domain concepts in ontologies.

A prominent approach to probabilistic DLs is due to Lukasiewicz (2008), who proposes a probabilistic extension of expressive DLs such as $\mathcal{SHIQ}$. These logics have received more attention in the sense of an implementation and the discussion of expressive power (Klinov, Parsia, & Sattler, 2009; Klinov & Parsia, 2013). As in the related work discussed above, P-$\mathcal{SHIQ}$ allows the application of conditional probabilities only to TBox statements and to ABox assertions, but not to concept and role names. The semantics of P-$\mathcal{SHIQ}$ does not fix a single distribution on worlds, somewhat similarly to our proposal. However, the P-$\mathcal{SHIQ}$ approach is technically very different from ours. It has its root in default logic in the sense that it is based on Lehmann's lexicographic entailment and thus non-monotonic in nature.

Our Type 2 probabilistic DLs are similar in semantics and in spirit to temporal description logics (Artale, Kontchakov, Lutz, Wolter, & Zakharyaschev, 2007; Lutz, Wolter, & Zakharyaschev, 2008; Gabbay et al., 2003; Gutiérrez-Basulto, Jung, & Schneider, 2014). These temporal DLs have a two-dimensional semantics with one dimension being the "DL dimension" and the other one being time; likewise, our probabilistic DLs have two dimensions which are the DL dimension and the probabilistic (that is: possible worlds) dimension.

Work less directly related to ours includes probabilistic versions of OWL (da Costa & Laskey, 2006; Ding, Peng, & Pan, 2006), probabilistic versions of Datalog$^{\pm}$ (Gottlob, Lukasiewicz, Martinez, & Simari, 2013), probabilistic description logic programs (Lukasiewicz, 2007), and weighted DLs based on Markov Logic (Niepert, Noessner, & Stuckenschmidt, 2011).

Another line of research that one may consider related to probabilistic DLs are *fuzzy description logics*. However, it is important to understand that while probabilistic DLs are about *uncertainty*, fuzzy DLs focus on the orthogonal aspect of *vagueness*: the central notion is *how much* some individual, say john, belongs to some concept, say Rich. In contrast, in subjective probabilities john is either Rich or he is not and one is concerned with the degree of belief in the former being the case. Due to this inherent semantical difference, we do not discuss fuzzy DLs in depth and instead refer to the survey by Lukasiewicz and Straccia (2008); additionally, an overview of more recent work is given by Borgwardt, Distel, and Peñaloza (2015).

6. Conclusion

We have proposed a novel family of probabilistic DLs that are derived in a straightforward way from Halpern's Type 2 probabilistic FO. We have also provided a substantial analysis of the complexity of reasoning in these logics, considering standard reasoning tasks such as concept subsumption and KB consistency. For simplicity, we have concentrated on probabilistic DLs that extend the basic classical DLs $\mathcal{ALC}$ and $\mathcal{EL}$. We conjecture, though, that our probabilistic DLs, and in particular the most well-behaved variants Prob-$\mathcal{ALC}$ and Prob-$\mathcal{EL}^{01}$, can be extended with many standard DL constructors without a significant impact on their computational properties. The most challenging technical problem left open is whether reasoning in Prob-$\mathcal{ALC}_r$ is decidable when neither linear inequalities nor independence constraints are admitted.

An interesting direction for future research is to consider query answering under ontologies in data-centric applications with uncertain data, in the spirit of probabilistic databases. While our possible worlds semantic is in line with standard approaches to probabilistic databases (Suciu, Olteanu, Ré, & Koch, 2011; Dalvi & Suciu, 2012; Fuhr & Rölleke, 1997), in such a setup one typically wants to make strong independence assumptions on the data (that is, between ABox assertions) which we have not considered in this paper. First steps into this direction have been made (Jung & Lutz, 2012; Ceylan & Peñaloza, 2015; Gottlob et al., 2013). Moreover, a very attractive setup for probabilistic reasoning about data is to combine *statistical* probabilities in the TBox with *subjective* probabilities in the data, thus transforming knowledge about statistics into subjective beliefs about concrete individuals. This form of reasoning is semantically challenging and inherently nonmonotonic and it has a long history (Grove, Halpern, & Koller, 1994). In the context of probabilistic FO, there are at least two interesting proposals that address this problem (Bacchus, Grove, Koller, & Halpern, 1992; Koller & Halpern, 1996). Transferring these setups to probabilistic DLs in a way that is both semantically adequate and computationally well-behaved is an interesting and challenging problem.

Acknowledgements

This work forms part of the DFG project *Probabilistic description logic as a fragment of probabilistic first-order logic* (LU1417/1-1 and SCHR1118/6-1). The first author was supported by the M8 PostDoc project TS-OBDA of the University of Bremen's Excellence Initiative.

Appendix A. Prob-$\mathcal{EL}$ Lower Bounds

We have to deal with the cases where $\sim\, \in \{=, >\}$ and $p \in (0, 1)$. As in the main part of the paper, we distinguish the subcases where $p \leq 0.5$ and where $p > 0.5$.

A.1 Non-convexity for $\sim\, \in \{=, >\}$ and $p \leq 0.5$

Lemma 14 remains true when we replace $\geq$ with $=$ in $\mathcal{T}$, C, and the D_{ij}. The first part of the proof still goes through, without any modifications. For the model construction, we need a slight modification when $p < 0.5$: we add a new world w_3 with $A_i^{\mathcal{I},w_3} := B_{ij}^{\mathcal{I},w_3} := \emptyset$ and update the probability distribution by putting

$$\begin{aligned} \mu(w_1) &= \mu(w_2) = p \\ \mu(w_3) &= 1 - 2p. \end{aligned}$$

When we replace $\geq$ with $>$ and assume $p < 0.5$, the proof of Lemma 14 goes through without any modifications. However, the case $\sim\, = \,>$ and $p = 0.5$ requires a slightly different construction, since $P_{>0.5}A_1 \sqcap P_{>0.5}A_2$ does already imply that $A_1 \sqcap A_2$ holds in some world. We therefore use intersections of three concepts instead of just two. Put

$$\begin{aligned} \mathcal{T} &= \{A_i \sqcap A_j \sqcap A_k \sqsubseteq P_{>p}B_{ijk} \mid 1 \leq i < j < k \leq 4\} \\ C &= P_{>p}A_1 \sqcap \ldots \sqcap P_{>p}A_4 \\ D_{ijk} &= P_{>p}B_{ijk}. \end{aligned}$$

It is only slightly more complicated to show that the above witnesses non-convexity. We begin with showing

$$\mathcal{T} \models C \sqsubseteq D_{123} \sqcup D_{124} \sqcup D_{134} \sqcup D_{234}.$$

Let $\mathcal{I}$ be a model of $\mathcal{T}$, and let $d \in C^{\mathcal{I},w}$. Since $d \in (P_{>0.5}A_i)^{\mathcal{I},w}$ for $1 \leq i \leq 4$, there is a world v with $d \in (A_i \sqcap A_j \sqcap A_k)^{\mathcal{I},v}$ for some $1 \leq i < j < k \leq 4$: Otherwise, d would satisfy at most two of the concepts A_i in any world, so the whole set of worlds would need to have probability at least $\frac{1}{2}\sum_{i=1}^{4} p_d^{\mathcal{I}}(A_i) > \frac{1}{2} \cdot 2 = 1$, contradiction. By the TBox, it follows that $d \in D_{ijk}^{\mathcal{I},v}$.

For the second part, we show w.l.o.g. that $\mathcal{T} \not\models C \sqsubseteq D_{123}$. Like in the proof of Lemma 14, we can take $\Delta^{\mathcal{I}} = \{d\}$, and it then suffices to show that we can find a set of worlds, a probability distribution on worlds, and a choice of satisfaction or non-satisfaction of the A_i by d in these worlds such that for each i, the worlds in which d satisfies A_i have probability more than 0.5 but d does not satisfy $A_1 \sqcap A_2 \sqcap A_3$ in any world. This is easy to enforce: we take $W = \{w_1, w_2, w_3, w_4\}$ and assign probabilities $\mu(w_1) = 0.4$, $\mu(w_2) = \mu(w_3) = \mu(w_4) = 0.2$. We let d satisfy A_1 in w_2, w_3, w_4, and A_i in w_1 and w_i for $i = 2, 3, 4$; then $p_d^{\mathcal{I}}(A_i) = 0.6 > 0.5$ for all i, and $A_1 \sqcap A_2 \sqcap A_3$ has empty extension in all worlds.

A.2 Non-convexity for $\sim\, \in \{=, >\}$ and $p \geq 0.5$

For the case $\sim\, = \,>$, Lemma 15 continues to hold. In fact, the proof goes through with only slight modifications: In the first part, one can proof the claim by deriving the contradiction $pn < S \leq pn$ (instead of $pn \leq S < pn$); the model construction remains. In the case where $\sim$ is $=$, again the first part of the proof of Lemma 15 goes through, but we have to change the model construction. We add a world w_3 such that in this world d is not in the extension of any concept, i.e., $A_{ij}^{\mathcal{I},w_3} := A_i^{\mathcal{I},w_3} := B_{ij}^{\mathcal{I},w_3} := \emptyset$. Moreover, we need to modify the probability distribution μ in the following way:

$$\begin{aligned} \mu(w_{1\ell}) &:= \mu(w_{2\ell}) := \tfrac{p}{2n-1} \\ \mu(w_3) &:= 1 - \tfrac{4np}{2n-1} \end{aligned}$$

It is not hard to verify that $\mu(w_3) \geq 0$ since $p > 0.5$ and $n > \frac{1}{2(p-1)}$, and $\sum_{w \in W} \mu(w) = 1$, thus μ is a valid probability distribution. Further, we can check with the same argumentation as in the proof above that $p_d^{\mathcal{I}}(A_{i\ell}) = p$, thus $d \in C^{\mathcal{I},v}$ for every v. Finally, it is easy to check that $d \notin D_{i_0j_0}^{\mathcal{I},v}$ for any $v \in W$.

A.3 Non-convexity Implies ExpTime-Hardness of Prob-$\mathcal{EL}^{\sim p}$

We exploit non-convexity to reduce concept satisfiability in $\mathcal{ALC}$ similar to what was done in (Baader et al., 2005). Notice that the considered logics Prob-$\mathcal{EL}^{\sim p}$ have different witnesses for non-convexity. We deal here only with the case $\sim = \geq$ and $p \leq 0.5$; however, by using the appropriate non-convexity witness in Step (3) below, one gets ExpTime-hardness for any non-convex extension of $\mathcal{EL}$ appearing in this paper.

Suppose that an $\mathcal{ALC}$-TBox $\mathcal{T}$ and a concept name A_0 are given for which satisfiability is to be decided. First, we manipulate the TBox $\mathcal{T}$ as follows:

(1) Ensure that negation $\neg$ occurs in front of concept names, only: for every subconcept $\neg C$ in $\mathcal{T}$ with C compound, introduce a fresh concept name A, replace $\neg C$ with $\neg A$, and add $A \sqsubseteq C$ and $C \sqsubseteq A$ to $\mathcal{T}$.

(2) Eliminate negation: for every subconcept $\neg A$, introduce a fresh concept name $\overline{A}$, replace every occurrence of $\neg A$ with $\overline{A}$, and add $\top \sqsubseteq A \sqcup \overline{A}$ and $A \sqcap \overline{A} \sqsubseteq \bot$ to $\mathcal{T}$.

(3) Eliminate disjunction: modulo introduction of new concept names, we may assume that $\sqcup$ occurs in $\mathcal{T}$ only in the form (i) $A \sqcup B \sqsubseteq C$ and (ii) $C \sqsubseteq A \sqcup B$, where A, B, and C are concept names. The former kind of inclusion is replaced with $A \sqsubseteq C$ and $B \sqsubseteq C$. The latter one is replaced with the non-convexity witness for Prob-$\mathcal{EL}^{\geq p}$ using fresh concept names $A_1, \ldots, A_k$, B_{ij} and k as in Lemma 14.

$$\begin{aligned} A_i \sqcap A_j &\sqsubseteq P_{\geq p} B_{ij} \quad \text{for } 1 \leq i < j \leq k \\ C &\sqsubseteq P_{\geq p} A_1 \sqcap \ldots \sqcap P_{\geq p} A_k \\ P_{\geq p}(B_{12}) &\sqsubseteq A \\ P_{\geq p}(B_{ij}) &\sqsubseteq B \quad \text{for } 1 \leq i < j \leq k, ij \neq 12 \end{aligned}$$

Let $\mathcal{T}'$ be the TBox obtained by these manipulations. It is standard to prove that A_0 is satisfiable w.r.t. $\mathcal{T}$ iff A_0 is satisfiable w.r.t. $\mathcal{T}'$.

The TBox $\mathcal{T}'$ contains only the operators $\sqcap$, $\exists$, $\top$, $\bot$, and $P_{\geq p}$. Thus, it remains to deal with the concept $\bot$ being the last one not allowed by our syntax for Prob-$\mathcal{EL}^{\geq p}$. To thind end, we reduce satisfiability of A_0 w.r.t. $\mathcal{T}'$ to (non-)subsumption in Prob-$\mathcal{EL}^{\geq p}$: Introduce a fresh concept name L and obtain $\mathcal{T}''$ from $\mathcal{T}'$ by replacing every occurrence of $\bot$ with L and adding the concept inclusions $\exists r.L \sqsubseteq L$, for every role r from $\mathcal{T}'$. Then A_0 is satisfiable w.r.t. $\mathcal{T}'$ iff $\mathcal{T}'' \not\models A_0 \sqsubseteq L$.

Appendix B. Invariants in Proof of Lemma 17

The induction base for invariants 1–3 holds simply by construction; the induction base for invariants 4–6 is trivial as there is no concept C satisfying any of their premises. For the induction step, we consider separately the cases when the rule is applied to some $d \in \mathsf{Ind}(\mathcal{A})$ or to some d of the form (A, v); we start with the former. In this case, invariants 4–6 are trivially preserved except in applications of **R4**, so we focus on invariants 1–3. We need to show that the right hand sides remain consequences of $\mathcal{K}$ when the rule application at hand effectively adds new basic concepts to the ABoxes $\mathcal{A}_w$; although of course this happens in one go, we pretend for ease of notation that basic concepts are added one by one. We distinguish cases on which rule is applied.

R1 Assume $X_1 \sqcap \ldots \sqcap X_n \sqsubseteq A \in \mathcal{T}$ and $d \in X_i^{\mathcal{I},w}$ for all $1 \leq i \leq n$. Then, **R1** will add d to the extension of A in world w. We analyse the effect regarding the addition of basic concepts to some $\mathcal{A}_v$.

- $A(d)$ is added to $\mathcal{A}_w$; since this is sanctioned by the TBox, invariants 1–3 are preserved.
- $P_{>0}A(d)$ is added to $\mathcal{A}_v$ for all v (if not already present; we will omit this proviso in the further discussion). The induction hypothesis and the assumption imply that $\mathcal{K} \models P_{>0}(X_1 \sqcap \ldots \sqcap X_n)(d)$. Since $X_1 \sqcap \ldots \sqcap X_n \sqsubseteq A \in \mathcal{T}$, also $\mathcal{K} \models P_{>0}A(d)$, thus invariants 1-3 are preserved.
- $P_{=1}A(d)$ may be added to $\mathcal{A}_v$ for all v, namely if $A(d) \in \mathcal{A}_v$ for all v (after the rule application). Then in particular $A(d) \in \mathcal{A}_{v_{=1}}$. The induction hypothesis yields $\mathcal{K} \models P_{=1}A(d)$ which implies that invariants 1–3 are preserved.
- $(\exists r.A)(a)$ may be added to $\mathcal{A}_w$ for some $a \in \mathsf{Ind}(\mathcal{A})$, namely if $\mathcal{A}_w$ contains $r(a, d)$. Moreover, by assumption $X_1(d), \ldots, X_n(d) \in \mathcal{A}_w$. Preservation of invariants 1-3 is now straightforward as $(\mathcal{T}, \mathcal{A} \cup \{r(a, d), X_1(d), \ldots, X_n(d)\}) \models (\exists r.\, A)(a)$.

R2 Assume $A \sqsubseteq P_{>0}B \in \mathcal{T}$, $d \in A^{\mathcal{I},w}$, and $d \notin (P_{>0}B)^{\mathcal{I},w}$. Then, **R2** will add d to the extension of B in world $\mathsf{wit}(d, w, B)$. We again analyse which basic concepts are added to some $\mathcal{A}_v$. Throughout, we denote the original value of $\mathcal{A}_v$ before application of **R2** by $\mathcal{A}'_v$.

- $B(d)$ is added to $\mathcal{A}_v$ for $v = \mathsf{wit}(d, w, B)$. As $v \notin \{v_{\mathcal{A}}, v_{=1}\}$, only invariant 3 is affected; i.e. we have to show that $\mathcal{K} \models P_{>0}(\mathcal{A}'_v \wedge B(d))$. Observe that we can assume w.l.o.g. that $\mathcal{A}'_v = \mathcal{A}_{v_{=1}}$ since up to now any rule that we applied to some individual in world $v_{=1}$ was also applicable to this individual in v and vice versa. By the induction hypothesis, we have $\mathcal{K} \models (P_{=1}\mathcal{A}'_v) \wedge P_{>0}A(d)$, and hence $\mathcal{K} \models P_{>0}(\mathcal{A}'_v \wedge B(d))$ because $A \sqsubseteq P_{>0}B \in \mathcal{T}$. (This argument clearly extends to adding several basic concepts to $\mathcal{A}_v$ at once.)
- $P_{>0}B(d)$ is added to $\mathcal{A}_v$ for all v. This effect is handled like the previous one.

- It cannot happen that $P_{=1}B(d)$ is added to any $\mathcal{A}_v$, since we had $d \notin (P_{>0}B)^{\mathcal{I},w}$ before the rule application and $\mathcal{I}$ has more than one world.
- $(\exists r.B)(a)$ may be added to $\mathcal{A}_v$ for $v = \mathsf{wit}(d, w, B)$ and some $a \in \Delta^{\mathcal{I}}$. Since $d \in \mathsf{Ind}(\mathcal{A})$ and the interpretation of roles is affected only by **R4**, which adds only edges into domain elements of the form (B, v) to $r^{\mathcal{I},v}$, we have $a \in \mathsf{Ind}(\mathcal{A})$ and $r(a, d) \in \mathcal{A}'_v$. The argument then proceeds as for $B(d)$.

R3 Assume $A \sqsubseteq P_{=1}B \in \mathcal{T}$ and $d \in A^{\mathcal{I},w}$. **R3** adds d to the extension of B in all worlds. We again analyse which basic concepts are added to some $\mathcal{A}_v$.

- $B(d)$ is added to $\mathcal{A}_v$. By assumption, $P_{>0}A(d) \in \mathcal{A}_v$. From $A \sqsubseteq P_{=1}B \in \mathcal{T}$, we obtain $\mathcal{T} \models P_{>0}A \sqsubseteq B$. Thus, invariants 1-3 are preserved.
- $P_{>0}B(d)$ is added to $\mathcal{A}_v$. Preservation of the invariants follows from the first point.
- $P_{=1}B(d)$ is added to $\mathcal{A}_v$. Again, preservation of the invariants follows from the first point.
- $(\exists r.B)(a)$ is added to $\mathcal{A}_v$ whenever $r(a, d) \in \mathcal{A}_v$. The invariants are preserved because $(\mathcal{T}, r(a, d) \wedge A(d)) \models (\exists r.\, B)(a)$.

R4 Assume $A \sqsubseteq \exists r.B \in \mathcal{T}$ and $d \in A^{\mathcal{I},w}$. **R4** adds the pair $(d, (B, w))$ to the extension of r in world w and (B, w) to the extension of B. Thus, $\exists r.B'(d)$ is added to $\mathcal{A}_w$ whenever $(B, w) \in B'^{\mathcal{I},w}$ for some concept name B' (since (B, w) is an unnamed individual, this is the only effect regarding addition of basic concepts to some $\mathcal{A}_v$). However, as $A \sqsubseteq \exists r.B \in \mathcal{T}$ and by the inductive hypothesis (invariant 5) $\mathcal{T} \models B \sqsubseteq B'$, we also have $\mathcal{T} \models A \sqsubseteq \exists r.B'$. Since by assumption $A(d) \in \mathcal{A}_w$, the invariants 1-3 remain true. It can be shown as in the case $d \notin \mathsf{Ind}(\mathcal{A})$ that invariants 4-6 are preserved for (B, w); see below.

Let us now consider the case $d = (B, w')$. Since d is an anonymous individual, preservation of invariants 1–3 is trivial except when existential concept assertions are added to some $\mathcal{A}_w$, so we largely focus on invariants 4–6. Again, we distinguish cases on which rule is applied.

R1 Assume $X_1 \sqcap \ldots \sqcap X_n \sqsubseteq A \in \mathcal{T}$ and $d \in X_i^{\mathcal{I},w}$ for all $1 \le i \le n$. Then, **R1** will add d to the extension of A in world w. We analyse which basic concepts are made true in $\mathcal{I}$ by applying the rule.

- We have $d \in A^{\mathcal{I},w}$ after rule application. If d is inactive in w, we have $\mathcal{T} \models B \sqsubseteq P_{=1}(X_1 \sqcap \ldots \sqcap X_n)$ by the induction hypothesis. Since $X_1 \sqcap \ldots \sqcap X_n \sqsubseteq A \in \mathcal{T}$, we get that $\mathcal{T} \models B \sqsubseteq P_{=1}A$, showing that invariant 4 is preserved. The argument for invariants 5 and 6 is analogous.
- We have $d \in P_{>0}A^{\mathcal{I},v}$ for all v after rule application. Observe that for showing that invariants 4-6 are preserved, it in fact suffices to show $\mathcal{T} \models B \sqsubseteq P_{>0}A$.

  If d is inactive in w, then the induction hypothesis yields $\mathcal{T} \models B \sqsubseteq P_{=1}(X_1 \sqcap \ldots \sqcap X_n)$, thus, $\mathcal{T} \models B \sqsubseteq P_{>0}A$ since $X_1 \sqcap \ldots \sqcap X_n \sqsubseteq A \in \mathcal{T}$. Assume now that $d = (B, w')$ is **R4**-active in w, thus $w = w'$. By induction, we know that $\mathcal{T} \models B \sqsubseteq X_1 \sqcap \ldots \sqcap X_n$, thus $\mathcal{T} \models B \sqsubseteq P_{>0}A$. Assume finally that $d = (B, w')$ is

R2-active in $w \neq w'$ and $w \in \{v_E, v'_E\}$ for some concept name E. By induction, we have that $\mathcal{T} \models E \sqcap P_{>0}B \sqsubseteq X_1 \sqcap \ldots \sqcap X_n$. Observe that $(B, w') \in E^{\mathcal{I},w}$ since (B, w') is **R2**-active in w. Hence, $(B, w') \in P_{>0}E^{\mathcal{I},w'}$; since (B, w') is **R2**-active in w, and a domain element that is **R2**-active must necessarily be also **R4**-active, (B, w') is **R4**-active in w' so that the induction hypothesis yields $\mathcal{T} \models B \sqsubseteq P_{>0}E$. Thus, $\mathcal{T} \models B \sqsubseteq P_{>0}(X_1 \sqcap \ldots \sqcap X_n)$ and therefore $\mathcal{T} \models B \sqsubseteq P_{>0}A$.

– Suppose $(B, w') \in P_{=1}A^{\mathcal{I},v}$ for some v after the rule application. Since (B, w') can be **R2**-active only in v_B or v'_B and **R4**-active only in w', there is a world $\hat{w}$ such that (B, w') is inactive in $\hat{w}$ and $(B, w') \in A^{\mathcal{I},\hat{w}}$. By the induction hypothesis, $\mathcal{T} \models B \sqsubseteq P_{=1}A$, which suffices to show that invariants 4-6 are preserved.

– Suppose $e \in (\exists r.A)^{\mathcal{I},w}$ for some domain element e after the rule application. Thus, $w = w'$ and $(e, d) \in r^{\mathcal{I},w}$. The latter implies that d is **R4**-active in w, whence $d \in B^{\mathcal{I},w}$ already before the rule application. By the premise, we have that $d \in X_i^{\mathcal{I},w}$ for $1 \leq i \leq n$ so that, by induction, $\mathcal{T} \models B \sqsubseteq X_1 \sqcap \ldots \sqcap X_n$. Since $X_1 \sqcap \ldots \sqcap X_n \sqsubseteq A \in \mathcal{T}$, we obtain $\mathcal{T} \models B \sqsubseteq A$. We now distinguish cases on whether $e \in \mathsf{Ind}(\mathcal{A})$. If $e \in \mathsf{Ind}(\mathcal{A})$, then $(\exists r.B)(e) \in \mathcal{A}'_w$. Since $\mathcal{T} \models B \sqsubseteq A$, adding $(\exists r.A)(e)$ to $\mathcal{A}_w$ preserves invariants 1-3. If $e \notin \mathsf{Ind}(\mathcal{A})$, then $e = (E, v)$ for some concept name E and world v. In this case, the induction hypothesis applied to $e = (E, v)$ in w yields $\mathcal{T} \models E \sqsubseteq P_{=1}C$, $\mathcal{T} \models E \sqsubseteq C$, or $\mathcal{T} \models P_{>0}F \sqcap E \sqsubseteq C$ (for $w \in \{v_F, v'_F\}$) with $\exists r.\,B$ a basic concept in C. Clearly, $\mathcal{T} \models B \sqsubseteq A$ suffices to show that invariants 4-6 are preserved.

R2 Assume $A \sqsubseteq P_{>0}E \in \mathcal{T}$, E a concept name, $d \in A^{\mathcal{I},w}$, and $d \notin (P_{>0}E)^{\mathcal{I},w}$. Then, **R2** will add d to the extension of E in world $\hat{w} = \mathsf{wit}(d, w, E) \in \{v_E, v'_E\}$. Note that, due to the definition of wit, d has been inactive in $\hat{w}$, but after rule application is **R2**-active. We again analyse which basic concepts become satisfied by this step.

– We have $d \in E^{\mathcal{I},\hat{w}}$ after the rule application. By the induction hypothesis (noting that d was inactive in $\hat{w}$), we have $\mathcal{T} \models B \sqsubseteq P_{=1}C$, thus also $\mathcal{T} \models E \sqcap P_{>0}B \sqsubseteq C \sqcap E$.

– We have $d \in (P_{>0}E)^{\mathcal{I},v}$ for all v after the rule application. Since $d \in A^{\mathcal{I},w}$, we have $d \in (P_{>0}A)^{\mathcal{I},v}$. The case for $v = \hat{w}$ is dealt with in the previous point. For the remaining cases it suffices to note that the induction hypothesis applied to v and C as in the invariant yields $\mathcal{T} \models B \sqsubseteq P_{=1}C'$, $\mathcal{T} \models B \sqsubseteq C'$, or $\mathcal{T} \models F \sqcap P_{>0}B \sqsubseteq C'$ (for $v \in \{v_F, v'_F\}$) where C' is obtained from C by replacing $P_{>0}E$ with $P_{>0}A$. The concept inclusion $(A \sqsubseteq P_{>0}E) \in \mathcal{T}$ ensures that invariants 4–6 are preserved.

– Suppose $d \in (P_{=1}E)^{\mathcal{I},v}$ for some (and then all) v after the rule application. As in the same case for **R1**, there is some world v' such that (B, w') is inactive in v'. By the induction hypothesis, $\mathcal{T} \models B \sqsubseteq P_{=1}E$. Thus, all invariants are preserved.

– The case of $e \in (\exists r.E)^{\mathcal{I},\hat{w}}$ for some e is impossible by definition of wit and **R4**.

R3 Assume $A \sqsubseteq P_{=1}E \in \mathcal{T}$ and $d = (B, w') \in A^{\mathcal{I},w}$. **R3** adds (B, w') to the extension of E in all worlds. We analyse which basic concepts become satisfied by this step.

– We have $(B, w') \in (P_{=1}E)^{\mathcal{I},v}$ after the rule application. As in the same case for **R1**, there exists a world $\hat{w}$ such that (B, w') is inactive in $\hat{w}$. By assumption, we have $(B, w') \in P_{>0}A^{\mathcal{I},\hat{w}}$. The induction hypothesis yields $\mathcal{T} \models B \sqsubseteq P_{>0}A$; hence $\mathcal{T} \models B \sqsubseteq P_{=1}E$. Thus, all invariants remain satisfied.

– The proof for $(B, w') \in E^{\mathcal{I},v}$ and $(B, w') \in P_{>0}E^{\mathcal{I},v}$ after the rule application is a consequence of the previous item.

– Suppose $e \in (\exists r.E)^{\mathcal{I},v}$ for some e after the rule application. Thus, $(e, (B, w')) \in r^{\mathcal{I},v}$, $v = w'$, which implies that (B, w') is **R4**-active in v and $(B, w') \in (P_{>0}A)^{\mathcal{I},v}$. By the induction hypothesis, we obtain $\mathcal{T} \models B \sqsubseteq P_{>0}A$ and, since $(A \sqsubseteq P_{=1}E) \in \mathcal{T}$, also $\mathcal{T} \models B \sqsubseteq E$. The argument then proceeds as in the same case for **R1**.

R4 Assume $A \sqsubseteq \exists r.E \in \mathcal{T}$, E a concept name, and $(D, w') \in A^{\mathcal{I},w}$. Then, **R4** adds $(d, (E, w))$ to $r^{\mathcal{I},w}$ and (E, w) to $E^{\mathcal{I},w}$. We analyse which basic concepts become satisfied by this step.

– We have $(E, w) \in E^{\mathcal{I},w}$ after the rule application. Since (E, w) becomes **R4**-active in w, we have to establish invariant 5, which however is just the trivial statement $\mathcal{T} \models E \sqsubseteq E$.

– We have $(E, w) \in P_{>0}E^{\mathcal{I},v}$ for all v after the rule application. For $v = w$, it follows from the previous point that invariant 5 is preserved. So let $v \neq w$. Then (E, w) cannot be **R4**-active in v. If (E, w) is inactive in v, it suffices to observe that $\mathcal{T} \models E \sqsubseteq P_{=1}(P_{>0}E)$ to show that invariant 4 is preserved. If (E, w) is **R2**-active in v, we have to establish invariant 6, which however is just the trivial statement $\mathcal{T} \models P_{>0}E \sqsubseteq P_{>0}E$.

– Suppose $(E, w) \in P_{=1}E^{\mathcal{I},v}$ for some (and then all) v after the rule application. As in the same case for **R1**, there exists a world $\hat{w}$ such that (E, w) is inactive in $\hat{w}$. The induction hypothesis applied to this world yields $\mathcal{T} \models E \sqsubseteq P_{=1}E$. It is, then, now not hard to verify that all invariants remain preserved.

– Suppose $e \in (\exists r.B')^{\mathcal{I},v}$ for some v after the rule application. Thus, $(e, (E, w)) \in r^{\mathcal{I},v}$ and $(E, w) \in B'^{\mathcal{I},v}$. The former implies that $v = w$ and that (E, w) was **R4**-active in v already before the rule application. The induction hypothesis applied to (E, w) in $v = w$ yields $\mathcal{T} \models E \sqsubseteq B'$. The argument now proceeds as in the same case for **R1**.

Appendix C. Proof of the Claim in Lemma 19

For the sake of completeness, we repeat the claim to be proved:

Claim. $(A, w) \in D^{\mathcal{I},v}$ iff $D \in \pi((A, w), v)$ for all $D \in \mathsf{Sub}$, $(A, w) \in \Delta^{\mathcal{I}}$, and $v \in W$.

Proof of claim. We start with the "$\Leftarrow$"-direction. Let $D \in \pi((A, w), v)$. We distinguish the following cases:

- $D = P$ is a primitive concept name. Immediate by definition of $P^{\mathcal{I}}$.

- $D = P_{=1}B$. Since $D \in \pi((A, w), v)$, by definition of π one of the following cases applies:

- $D \in C_A$. Then $B \in \pi((A, w), w')$ for all w'. By definition of $B^{\mathcal{I}}$, it follows that $(A, w) \in (P_{=1}B)^{\mathcal{I},w}$ as required.
- $P_{=1}B' \in C_A$ and $D \in C_{B'}$. Due to non-applicability of rule **R2**, we then have $D \in C_A$ and can argue as in the previous case.
- $P_{>p}B' \in C_A$, and $D \in C_{B'}$. Thus, the definition of $\mathsf{cert}(\cdot)$ yields $D \in \mathsf{cert}(C_{B'})$. Since rule **R4** is not applicable, we have $D \in C_A$ and can again argue as before.

- $D = P_{>p}B$. We distinguish the same cases as above, i.e.,
  - $D \in C_A$. Then $C_B \subseteq \pi((A, w), p)$ for at least $k - 1$ distinct worlds p from $\{p_\ell \mid \ell \leq k\}$ (by Conditions 4 and 5 above) and $C_B \subseteq \pi((A, w), \delta_{Bi})$ and $C_B \subseteq \pi((A, w), \delta_{Bj})$ for distinct $i, j \in \{1, 2, 3\}$ (Conditions 6 and 7). By (P1) and the definition of $B^{\mathcal{I}}$, it follows that $(A, w) \in (P_{>p}B)^{\mathcal{I},v}$, as required.
  - $P_{=1}B' \in C_A$ and $D \in C_{B'}$. Due to non-applicability of rule **R2**, we then have $D \in C_A$ and can argue as in the previous case.
  - $P_{>p}B' \in C_A$, and $D \in C_{B'}$. By definition of $\mathsf{cert}(\cdot)$, we have $D \in \mathsf{cert}(C_{B'})$. Since rule **R4** is not applicable, we have $D \in C_A$ and can again argue as before.

- $D = \exists r.B$. By definition of $r^{\mathcal{I}}$, we have $((A, w), (B, v)) \in r^{\mathcal{I},v}$. By Condition 1 in the definition of π and by definition of $B^{\mathcal{I}}$, we have $(B, v) \in B^{\mathcal{I},v}$. The semantics then yields $(A, w) \in (\exists r.B)^{\mathcal{I},v}$.

For the "$\Rightarrow$"-direction, assume that $(A, w) \in D^{\mathcal{I},v}$. Distinguish the following cases:

- $D = P$ is a primitive concept name. Immediate by definition of $P^{\mathcal{I}}$.
- $D = P_{=1}B$. Take a $1_j \in W$ such that $w \neq 1_j$ (such a 1_j exists since there are at least two worlds of the form 1_j). Since $(A, w) \in D^{\mathcal{I},v}$, we have $(A, w) \in B^{\mathcal{I},1_j}$. By definition of $B^{\mathcal{I}}$, we thus have $C_B \subseteq \pi((A, w), 1_j)$. By definition of $\pi((A, w), 1_j)$, it follows that for every $D' \in C_B$, we have

  (i) $D' \in C_A$ with D' of the form P_*B' (by Condition 2 in the definition of π), or
  (ii) there is a $P_{=1}B' \in C_A$ with $D' \in C_{B'}$ (by Condition 3).

  Thus, $C_B \subseteq \mathsf{cert}(C_A)$ and non-applicability of the rule **R5** yields $P_{=1}B \in C_A$. By Condition 2 of π, we have $P_{=1}B \in \pi((A, w), v)$ as required.

- $D = P_{>p}B$. Since $(A, w) \in (P_{>p}B)^{\mathcal{I},v}$, (P2) yields the following cases:
  - $(A, w) \in B^{\mathcal{I},1_j}$ and $w \neq 1_j$. Then we can argue as in the previous case that $P_{=1}B \in C_A$. Thus rule **R3** yields $P_{>p}B \in C_A$ and by Condition 2 of π, we have $P_{>p}B \in \pi((A, w), v)$ as required.
  - $(A, w) \in B^{\mathcal{I},\delta_{B'j}}$ and $w \neq \delta_{B'j}$. By definition of $B^{\mathcal{I}}$, we thus have $C_B \subseteq \pi((A, w), \delta_{B'j})$. By definition of $\pi((A, w), \delta_{B'j})$, it follows that for every $D' \in C_B$, we have

    (i) $D' \in C_A$ with D' of the form P_*B'' (Condition 2 of π),
    (ii) there is a $P_{=1}B'' \in C_A$ with $D' \in C_{B''}$ (Condition 3), or

(iii) $P_{>p}B' \in C_A$ and $D' \in C_{B'}$ (Condition 6 or 7).

If exclusively (i) and (ii) apply, then $C_B \subseteq \mathsf{cert}(C_A)$; otherwise, we have $P_{>p}B' \in C_A$ and $C_B \subseteq \mathsf{cert}(C_A) \cup C_{B'}$. In the first case, non-applicability of **R5** yields $P_{=1}B \in C_A$ and **R3** yields $P_{>p}B \in C_A$. In the latter case, non-applicability of **R6** again yields $P_{>p}B \in C_A$. By Condition 2 of π, we have $P_{>p}B \in \pi((A,w),v)$ as required.

- $D = \exists r.B$. Then there is an (A',v) such that $((A,w),(A',v)) \in r^{\mathcal{I},v}$ and $(A',v) \in B^{\mathcal{I},v}$. By definition of $r^{\mathcal{I}}$, we have $\exists r.A' \in \pi((A,w),v)$, and by definition of $B^{\mathcal{I}}$, we have $C_B \subseteq \pi((A',v),v)$. Since Conditions 4–7 in the definition of π apply only to pairs $((\_,u),u') \in \Delta^{\mathcal{I}} \times W$ with $u \neq u'$, the last inclusion must be due to Conditions 1–3; that is, for every $D' \in C_B$ we have $D' \in C_{A'}$ or $P_{=1}B' \in C_{A'}$ with $D' \in C_{B'}$. In the latter case, we again obtain $D' \in C_{A'}$ by non-applicability of **R2**. Thus $C_B \subseteq C_{A'}$. To continue, we make a case distinction as follows:
  - $v = w$. Then the definition of π yields that $\exists r.A' \in C_A$ or $P_{=1}B' \in C_A$ with $\exists r.A' \in C_{B'}$. In the latter case, we again obtain $\exists r.A' \in C_A$ by non-applicability of **R2**. Since $C_B \subseteq C_{A'}$, this implies $\exists r.B \in C_A$ by non-applicability of **R1**. By definition of π, we thus have $\exists r.B \in \pi((A,w),v)$.
  - $v = 1_j$, $v \neq w$. Since $\exists r.A' \in \pi((A,w),v)$, the definition of π yields a $P_{=1}B' \in C_A$ with $\exists r.A' \in C_{B'}$. By non-applicability of **R1** and $C_B \subseteq C_{A'}$, we have $\exists r.B \in C_{B'}$. Thus, Condition 3 of π yields $\exists r.B \in \pi((A,w),v)$ as required.
  - $v = p_\ell$, $v \neq w$. Since $\exists r.A' \in \pi((A,w),v)$, the definition of π implies that there is a $P_{=1}B' \in C_A$ with $\exists r.A' \in C_{B'}$ (Condition 3) or a $P_{>p}B' \in C_A$ with $\exists r.A' \in C_{B'}$ (Conditions 4 and 5). In the former case, we can continue as in the case $v = 1_j$ above. In the latter case, non-applicability of **R1** and $C_B \subseteq C_{A'}$ yield $\exists r.B \in C_{B'}$. Thus, Conditions 4 and 5 of π yield $\exists r.B \in \pi((A,w),v)$ as required.
  - $v = \delta_{Ej}$, $v \neq w$. The reasoning is the same is in the previous case; we give it here for the sake of completeness. Since $\exists r.A' \in \pi((A,w),v)$, the definition of π implies that there is a $P_{=1}B' \in C_A$ with $\exists r.A' \in C_{B'}$ (Condition 3) or a $P_{>p}B' \in C_A$ with $\exists r.A' \in C_{B'}$ (Condition 6 and 7). In the former case, we can continue as in the case $v = 1_j$. In the latter case, non-applicability of **R1** and $C_B \subseteq C_{A'}$ yield $\exists r.B \in C_{B'}$. Thus, Conditions 6 and 7 of π yield $\exists r.B \in \pi((A,w),v)$ as required.

Appendix D. Proof for Theorem 24

Theorem 24. In Prob-$\mathcal{EL}^{01}_r$, concept subsumption is PSpace-hard even if probability restrictions $P_{=1}C$ are disallowed and only one of $\exists P_{=1}r.C$ and $\exists P_{>0}r.C$ is admitted.

Proof. We concentrate on Prob-$\mathcal{EL}^{01}$ where $\exists P_{=1}r.C$ is admitted and only sketch the modifications required for $\exists P_{>0}r.C$. The proof is by reduction of the word problem of deterministic, polynomially space-bounded Turing machines. Let $M = (Q,\Sigma,\Gamma,\delta,q_0,q_{\mathsf{acc}},q_{\mathsf{rej}})$ be such a machine, $x \in \Sigma$ an input of length n, and $m = p(n)$ the space bound of M on x. We assume w.l.o.g. that M terminates on every input, that it never attempts to move left on the left-most end of the tape, and that there are no transitions defined for q_{acc} and q_{rej}. Our aim is to construct in polynomial time a TBox $\mathcal{T}$ and concept C_0 such that $\mathcal{T} \models \top \sqsubseteq C_0$ iff M accepts x, based on the intuitions given in the proof sketch in the main part of the paper. We use the following concept and role names:

- the elements of Q are used as concept names;
- concept names $\sigma^{(i)}$ for $\sigma \in \Gamma$ and $i < m$ indicate that the content of the i-th tape cell is σ;
- concept names H_i for $i < m$ indicate that the head is on the i-th cell;
- a role name r.

The TBox $\mathcal{T}$ contains the following implications:

- Models contain an infinite r-chain of probability one:
$$\top \sqsubseteq \exists P_{=1} r.\top.$$
- At every point of the chain, there is a world that describes the initial configuration:
$$\top \sqsubseteq P_{>0}(q_0 \sqcap H_0 \sqcap x_0^{(0)} \sqcap \cdots \sqcap x_{n-1}^{(n-1)} \sqcap B^{(n)} \sqcap \cdots \sqcap B^{(m-1)})$$
where $x = x_0 \cdots x_{n-1}$ is the input and B denotes the blank symbol.
- The computation proceeds as required by M:
$$\begin{array}{rcll} \exists r.(q \sqcap H_i \sqcap \sigma^{(i)}) & \sqsubseteq & q' \sqcap H_{i-1} \sqcap \gamma^{(i)} & \text{for } 0 < i < m, \delta(q,\sigma) = (q', \gamma, L) \\ \exists r.(q \sqcap H_i \sqcap \sigma^{(i)}) & \sqsubseteq & q' \sqcap H_{i+1} \sqcap \gamma^{(i)} & \text{for } i < m-1, \delta(q,\sigma) = (q', \gamma, R) \\ \exists r.(\sigma^{(i)} \sqcap H_j) & \sqsubseteq & \sigma^{(i)} & \text{for } i, j < m, i \neq j. \end{array}$$

Finally, let C_0 be $P_{>0} q_{\mathsf{acc}}$.

Claim. $\mathcal{T} \models \top \sqsubseteq C_0$ iff M accepts x.

"$\Rightarrow$". Assume that M does not accept x. Let $c_0, \ldots, c_{k-1}$ be the (rejecting) computation of M on x, with configurations c_i represented in the obvious way as sets of concept names. Define a probabilistic interpretation $\mathcal{I} = (\Delta^{\mathcal{I}}, W, (\mathcal{I}_w)_{w \in W}, \mu)$ by putting:

- $\Delta^{\mathcal{I}} = W = \mathbb{N}$;
- $\mu(0) = 1/2$, $\mu(1) = 1/4$, $\mu(2) = 1/8$, ...
- $A^{\mathcal{I},j} = \{i \in \Delta^{\mathcal{I}} \mid j - i < k \wedge A \in c_{j-i}\}$;
- $r^{\mathcal{I},j} = \{(i, i+1) \mid i \in \Delta\}$;

It can be verified that $\mathcal{I}$ is a model of $\mathcal{T}$ with $0 \notin C_0^{\mathcal{I},0}$.

"$\Leftarrow$". Assume that M accepts x and let $c_{k-1}, \ldots, c_0$ be the (accepting) computation of M on x, again with configurations represented as sets of concept names (and for the sake of convenience named in reverse order, i.e. c_{k-1} is the starting configuration). Let $\mathcal{I}$ be a probabilistic model of $\mathcal{T}$ and let $d_0 \in \Delta^{\mathcal{I}}$, and $w_0 \in W$. By definition of $\mathcal{T}$, there is an infinite chain $d_0, d_1, \dots \in \Delta^{\mathcal{I}}$ such that $p^{\mathcal{I}}_{d_i,d_{i+1}}(r) = 1$. There must also be a world w such that $d_{k-1} \in A^{\mathcal{I},w}$ iff $A \in c_{k-1}$. It follows that for all $j < k$, $d_j \in A^{\mathcal{I},w}$ iff $A \in c_j$. Consequently, $d_0 \in q_{\mathsf{acc}}^{\mathcal{I},w}$ and thus $d_0 \in (P_{>0} q_{\mathsf{acc}})^{\mathcal{I},w_0}$.

To adapt the described reduction to Prob-$\mathcal{EL}^{01}_{r>0}$, we can simply replace $\top \sqsubseteq \exists P_{=1} r.\top$ with $\top \sqsubseteq \exists P_{>0} r.\top$ and all concepts $\exists r.C$ with $\exists P_{>0} r.C$. $\square$

Appendix E. Proof of Lemma 27

We first finish the "$\Leftarrow$"-direction, by showing that the invariants are preserved. The proof is by a straightforward induction on the number of applications of the rules in Figure 5. The induction base is trivial since $\emptyset \models A \sqsubseteq A$ and $\emptyset \models A \sqsubseteq \top$. For the induction step we start by showing soundness of the rules **R1-R6**, that is, for every set of concepts Γ we have

$$\mathcal{T} \models \sqcap \Gamma \sqsubseteq \sqcap \mathsf{cl}(\Gamma). \tag{$*$}$$

This fact is a direct consequence of the semantics for rules **R1-R5**. For **R6** assume $\exists \alpha.A \in \Gamma$ and $B \in Q(A)$. Invariant (**inv1**) implies $\mathcal{T} \models A \sqsubseteq B$, which means that we can certainly add $\exists \alpha.B$ to Γ.

Next, we analyse backtraces a little closer and prove the following claim.

Claim 1. If t is a backtrace to B, then $\mathcal{T} \models B \sqsubseteq P_{>0}(\sqcap \Gamma(t))$.

Proof of Claim 1. Let $t = S, A_1, \alpha_2, \ldots, \alpha_n, A_n$. The proof is by induction on the length n of t. For the induction base we let $n = 1$ and distinguish cases according to the form of S. Consider first the case that the backtrace starts with $S = B$, i.e., $P_{>0} B \in Q(A_1)$. From invariants (**inv1**) and (**inv2**) it follows that $\mathcal{T} \models A_1 \sqsubseteq P_{>0}(B \sqcap \sqcap Q_{\mathsf{cert}}(A_1))$. Since $\Gamma(t) = \mathsf{cl}(\{B\} \cup Q_{\mathsf{cert}}(A_1))$, by $(*)$, we obtain $\mathcal{T} \models A_1 \sqsubseteq P_{>0}(\sqcap \Gamma(t))$.

Assume now that the backtrace starts with $S = (r, B)$, i.e., $(A_1, B) \in R(P_{>0} r)$. Hence, by (**inv3**), we have $\mathcal{T} \models A_1 \sqsubseteq P_{>0}(\exists P_{>0} r.B)$, thus $\mathcal{T} \models A_1 \sqsubseteq P_{>0}(\exists r.P_{>0} B)$. From invariant (**inv2**) we get $\mathcal{T} \models A_1 \sqsubseteq P_{=1}(\sqcap Q_{\mathsf{cert}}(A_1))$ and $\mathcal{T} \models B \sqsubseteq P_{=1}(\sqcap Q_{\mathsf{cert}}(B))$. Overall, we obtain:

$$\mathcal{T} \models A_1 \sqsubseteq P_{>0}\left(\sqcap Q_{\mathsf{cert}}(A_1) \sqcap \exists r. \sqcap Q_{\mathsf{cert}}(B)\right).$$

Since $\Gamma(t) = \mathsf{cl}(Q_{\mathsf{cert}}(A_1) \cup \{\exists r.B' \mid B' \in Q_{\mathsf{cert}}(B)\})$, we can apply $(*)$ to obtain

$$\mathcal{T} \models A_1 \sqsubseteq P_{>0}(\sqcap \Gamma(t)).$$

For the induction step, let $n > 1$. By Definition 25, $(A_n, A_{n-1}) \in R(\alpha_n)$, thus, invariant (**inv3**) yields $\mathcal{T} \models A_n \sqsubseteq P_{>0}(\exists \alpha_n.A_{n-1})$. Applying the induction hypothesis, we get

$$\mathcal{T} \models A_n \sqsubseteq P_{>0}\left(\exists \alpha_n.P_{>0}(\sqcap \Gamma(t_{n-1}))\right).$$

Since $\exists\alpha_n.P_{>0}C \sqsubseteq P_{>0}\exists\alpha_n.C$ is valid for all C and (recall that α_n is a probabilistic role), we obtain

$$\mathcal{T} \models A_n \sqsubseteq P_{>0}\left(\exists\alpha_n.\bigsqcap\Gamma(t_{n-1})\right).$$

On the other hand, (**inv2**) implies $\mathcal{T} \models A_n \sqsubseteq P_{=1}(\bigsqcap Q_{\mathsf{cert}}(A_n))$. Together this yields:

$$\mathcal{T} \models A_n \sqsubseteq P_{>0}\left(\bigsqcap Q_{\mathsf{cert}}(A_n) \sqcap \exists\alpha_n.\bigsqcap\Gamma(t_{n-1})\right).$$

Since $\Gamma(t) = \mathsf{cl}(Q_{\mathsf{cert}}(A_n) \cup \{\exists\alpha_n.B \mid B \in \Gamma(t_{n-1})\})$ and by $(*)$ we finally have

$$\mathcal{T} \models A_n \sqsubseteq P_{>0}(\bigsqcap\Gamma(t)).$$

This finishes the proof of the claim.

It remains to show that the rules in Figure 5 preserve the invariants:

S1 Direct consequence of $(*)$.

S2 Since $P_{>0}B \sqsubseteq P_{=1}(P_{>0}B)$ and $P_{=1}B \sqsubseteq P_{=1}(P_{=1}B)$ are valid concept inclusions, this is a direct consequence of the semantics.

S3 $C \in Q_{\mathsf{cert}}(A)$ implies $\mathcal{T} \models A \sqsubseteq P_{=1}C$ by invariant (**inv2**), hence also $\mathcal{T} \models A \sqsubseteq C$ by the semantics.

S4 $\exists\alpha.B \in Q(A)$ implies $\mathcal{T} \models A \sqsubseteq \exists\alpha.B$ by invariant (**inv1**), thus also $\mathcal{T} \models A \sqsubseteq P_{>0}(\exists\alpha.B)$.

S5 On the one hand, $(A_1, A_2) \in R(\alpha)$ implies $\mathcal{T} \models A_1 \sqsubseteq P_{>0}(\exists\alpha.A_2)$, by (**inv3**). On the other hand, $B \in Q_{\mathsf{cert}}(A_2)$ yields $\mathcal{T} \models A_2 \sqsubseteq P_{=1}B$, by (**inv2**). By the semantics, together they imply $\mathcal{T} \models A_1 \sqsubseteq P_{=1}(\exists\alpha.B)$.

S6 Let t be a backtrace to B and $\Gamma = \Gamma(t)$ its type. By the Claim 1, $\mathcal{T} \models B \sqsubseteq P_{>0}C$ for every $C \in \Gamma$. Thus, in particular, $\mathcal{T} \models B \sqsubseteq P_{\sim p}A$ for all $P_{\sim p}A \in \Gamma$. By the semantics, $\mathcal{T} \models B \sqsubseteq P_{=1}(P_{\sim p}A)$, so $P_{\sim p}A$ can be added to $Q_{\mathsf{cert}}(B)$.

S7 Analogously to **S6**.

Now, to finish the "⇒"-direction, it remains to prove the Claim from the proof of Lemma 27. We start with showing the following important properties of the construction.

(i) for every $\sigma = (\alpha_1, w_1, A_1)\cdots(\alpha_n, w_n, A_n) \in \Delta^{\mathcal{I}}$ we have $(A_j, A_{j+1}) \in R(\hat{\alpha}_{j+1})$ for all $1 \leq j < n$;

(ii) for every $\sigma = (\alpha_1, w_1, A_1)\cdots(\alpha_n, w_n, A_n) \in \Delta^{\mathcal{I}}$, there are worlds w, v with $\pi(\sigma, w) = Q(A_n)$ and $\pi(\sigma, v) = Q_{\mathsf{cert}}(A_n)$;

(iii) $P_{\sim p}A \in \pi(\sigma, w)$ if and only if $P_{\sim p}A \in \pi(\sigma, v)$ for all $\sigma \in \Delta^{\mathcal{I}}$ and $w, v \in W$;

(iv) for all probabilistic roles α and $\sigma, \sigma' \in \Delta^{\mathcal{I}}$ with $\sigma' = \sigma \cdot (\alpha, v, B)$ and $A \in \pi(\sigma', w)$ we have $\exists\alpha.A \in \pi(\sigma, w)$.

Note that Point (i) verifies that the construction, specifically steps 2(b) and 3(b), is well-defined, i.e. that the arguments of $\Gamma(\cdot)$ are actually backtraces.

Proof of Points (i)-(iv). For the proof we assume that $\sigma = (\alpha_1, w_1, A_1) \cdots (\alpha_n, w_n, A_n)$.

We prove Point (i) by induction on the number of rule applications. The induction base is immediate by the definition of $\Delta^{\mathcal{I}_0}$. For the induction step note that new elements are introduced only in rule 1. So assume rule 1 is applied to σ, i.e., $\exists \alpha_{n+1}.A_{n+1} \in \pi_i(\sigma, w)$ for some $w \in W_i$. By construction, $\pi_i(\sigma, w)$ is closed under cl, thus, **R5** yields $\exists \hat{\alpha}_{n+1}.A_{n+1} \in \pi_i(\sigma, w)$. Now, rules **S4**, **S3**, and **S7** (depending on whether $\pi(\sigma, w)$ was defined as $Q(B)$, $Q_{\mathsf{cert}}(B)$, or $\Gamma(t)$ for some backtrace t) imply $(A_n, A_{n+1}) \in R(\hat{\alpha}_n)$, i.e., Point (i) is preserved after application of rule 1

Point (ii) can be proved by induction on the number of rule applications. The induction base is clear by the definition of $\mathcal{I}_0$. For the induction step it suffices to look at rule 1 and observe that when σ is added to $\Delta^{\mathcal{I}}$, we put $\pi(\sigma, w) = Q(A_n)$ for one world w and $\pi(\sigma, v) = Q_{\mathsf{cert}}(A_n)$ for all other worlds v (and there exist at least two worlds).

For Point (iii) we make a case distinction on $\pi(\sigma, w)$ and $\pi(\sigma, v)$. Inspection of the rules generating $\pi(\sigma, w)$ and $\pi(\sigma, v)$ shows that both are either $Q(A_n)$, $Q_{\mathsf{cert}}(A_n)$ or $\Gamma(t)$ for some backtrace t to A_n. For symmetry reasons it suffices to consider $\pi(\sigma, w)$. If $\pi(\sigma, w) = Q_{\mathsf{cert}}(A_n)$, then $P_{\sim p}A \in Q(A)$ (by rule **S3**) and $P_{\sim p}A \in \Gamma(t)$ for any backtrace t to A_n, by Definition 26. If $\pi(\sigma, w) = Q(A_n)$, then by rule **S2**, $P_{\sim p}A$ will be in $Q_{\mathsf{cert}}(A_n)$ and we proceed as before. If $\pi(\sigma, w) = \Gamma(t)$ for some backtrace t to A_n, then by **S6** we have $P_{\sim p}A \in Q_{\mathsf{cert}}(A_n)$ and, again, we can continue as before.

For proving Point (iv), we make a case distinction on whether $\pi(\sigma', w)$ was set to $Q(B)$ (via rule 1(a)), to $Q_{\mathsf{cert}}(B)$ (via rule 2(c) or 3(c)), or to $\Gamma(t)$ for some backtrace t (via rule 2(b) or 3(b)).

- Assume $\pi(\sigma', w)$ was set to $Q(B)$, i.e., rule 1 was applied to $\exists \alpha.B \in \pi(\sigma, w)$. Since $\pi(\sigma, w)$ is closed under cl and $A \in \pi(\sigma', w) = Q(B)$, by **R6** we obtain $\exists \alpha.A \in \pi(\sigma, w)$.

- Assume $\pi(\sigma', w)$ was set to $Q_{\mathsf{cert}}(B)$. By Point (i), $(A_n, B) \in R(\alpha)$ (note that α is probabilistic). Now, **S5** implies $\exists \alpha.A \in Q_{\mathsf{cert}}(A_n) \subseteq \pi(\sigma, w)$.

- Assume $\pi(\sigma', w)$ was set to $\Gamma(t)$ for some backtrace to t to B, that is, $\pi(\sigma', w)$ was assigned using step 2(b) or 3(b). Observe that $t' = t, \alpha, A_n$ is a backtrace to A_n as, by Point (i), we have $(A_n, B) \in R(\alpha)$. Moreover, due to step 2(b) (resp., 3(b)) we have $\pi(\sigma, w) = \Gamma(t')$. By definition of the type of a backtrace, we get that

$$\Gamma(t') = \mathsf{cl}(Q_{\mathsf{cert}}(A_n) \cup \{\exists \alpha.B' \mid B' \in \Gamma(t)\}).$$

  Since, by assumption, we have $A \in \pi(\sigma', w) = \Gamma(t)$, we get $\exists \alpha.A \in \Gamma(t') = \pi(\sigma, w)$.

This finishes the proof of the properties and we can show the Claim from the main part.

Claim. For all $\sigma \in \Delta^{\mathcal{I}}$, $w \in W$, and $C \in \mathsf{BC}$, we have $\sigma \in C^{\mathcal{I},w}$ iff $C \in \pi(\sigma, w)$.

Proof of the Claim. We prove the claim by induction on C (mostly just a case distinction except where composite basic concepts mention concept names). Throughout the following we assume $\sigma = (\alpha_1, w_1, A_1) \cdots (\alpha_n, w_n, A_n)$.

- $C = \top$. Then both $\sigma \in \top^{\mathcal{I},w}$ and $\top \in \pi(\sigma, w)$ for all $\sigma \in \Delta^{\mathcal{I}}$ and $w \in W$.
- $C = A \in \mathsf{CN}$. For this case, the claim holds trivially by definition of the interpretation of concept names.
- $C = P_{>0}A$. "$\Rightarrow$": Let $\sigma \in (P_{>0}A)^{\mathcal{I},w}$. Then, by the semantics, $\sigma \in A^{\mathcal{I},v}$ for some $v \in W$. The induction hypothesis implies $A \in \pi(\sigma, v)$. By **R3**, also $P_{>0}A \in \pi(\sigma, v)$, and by Point (iii), $P_{>0}A \in \pi(\sigma, w)$.

  "$\Leftarrow$": Let $P_{>0}A \in \pi(\sigma, w)$. By Point (ii), there is some world v with $\pi(\sigma, v) = Q(A_n)$. By Point (iii), $P_{>0}A \in \pi(\sigma, v)$. By construction rule 2(a), the world $v = (\sigma, A)$ is in W. By step (b) of rule 2, $A \in \pi(\sigma, v) = \Gamma(\delta(v)_1)$. The induction hypothesis yields $\sigma \in A^{\mathcal{I},v}$, thus $\sigma \in (P_{>0}A)^{\mathcal{I},w}$.
- $C = P_{=1}A$. "$\Rightarrow$": Let $\sigma \in (P_{=1}A)^{\mathcal{I},w}$, thus $\sigma \in A^{\mathcal{I},v}$ for all $v \in W$. By the induction hypothesis, $A \in \pi(\sigma, v)$ for all $v \in W$. By Point (ii), there is a world v' such that $\pi(\sigma, v') = Q_{\mathsf{cert}}(A_n)$; thus, $A \in Q_{\mathsf{cert}}(A_n)$. By **S3**, $P_{=1}A \in Q(A_n)$, and by **S2** also $P_{=1}A \in Q_{\mathsf{cert}}(A_n) = \pi(\sigma, v')$. By Point (iii) we obtain $P_{=1}A \in \pi(\sigma, w)$.

  "$\Leftarrow$": Let $P_{=1}A \in \pi(\sigma, w)$. By Point (iii), $P_{=1}A \in \pi(\sigma, v)$ for all $v \in W$. Since all $\pi(\sigma, v)$ are closed under cl, **R2** implies $A \in \pi(\sigma, v)$ for all v. By the hypothesis, $\sigma \in A^{\mathcal{I},v}$ for all $v \in W$, which, by the semantics, implies $\sigma \in (P_{=1}A)^{\mathcal{I},w}$.
- $C = \exists r.A$. "$\Rightarrow$": Assume $\sigma \in (\exists r.A)^{\mathcal{I},w}$. By the semantics, there is a $\sigma' \in \Delta^{\mathcal{I}}$ such that $\sigma' \in A^{\mathcal{I},w}$ and $(\sigma, \sigma') \in r^{\mathcal{I},w}$. By the induction hypothesis, we know that $A \in \pi(\sigma', w)$. Due to the model construction, there are three possibilities for (σ, σ') being in $r^{\mathcal{I},w}$:
  - $\sigma' = \sigma \cdot (P_{>0}r, v, B)$ and $w = (\sigma, r, B)$ for some concept name B. By construction rule 3(c), $\pi(\sigma', w) = Q_{\mathsf{cert}}(B)$ since σ' is not a prefix of σ. Hence, $A \in Q_{\mathsf{cert}}(B)$. By rule 3(b), we have that $\pi(\sigma, w) = \Gamma(\delta(w)_1) = \mathsf{cl}(Q_{\mathsf{cert}}(A_n) \cup \{\exists r.B' \mid B' \in Q_{\mathsf{cert}}(B)\})$. Since $A \in Q_{\mathsf{cert}}(B)$, $\exists r.A \in \pi(\sigma, w)$.
  - $\sigma' = \sigma \cdot (r, w, B)$ for some B. By construction, in particular rule 1, we have $\exists r.B \in \pi(\sigma, w)$ and $\pi(\sigma', w) = Q(B)$. Hence, $A \in Q(B)$. Since $\pi(\sigma, w)$ is closed under cl, rule **R6** yields $\exists r.A \in \pi(\sigma, w)$.
  - $\sigma' = \sigma \cdot (P_{=1}r, v, B)$. We apply Point (iv) to obtain $\exists P_{=1}r.A \in \pi(\sigma, w)$. Since $\pi(\sigma, w)$ is closed under cl, rule **R4** yields $\exists r.A \in \pi(\sigma, w)$.

  "$\Leftarrow$": Let $\exists r.A \in \pi(\sigma, w)$. By rule 1 of the construction, there is a domain element $\sigma' = \sigma \cdot (r, w, A)$ with $\pi(\sigma', w) = Q(A)$, thus $A \in \pi(\sigma', w)$ and, by induction, $\sigma' \in A^{\mathcal{I},w}$. By definition of the interpretation of role names, $(\sigma, \sigma') \in r^{\mathcal{I},w}$. Hence, $\sigma \in (\exists r.A)^{\mathcal{I},w}$.
- $C = \exists P_{=1}r.A$. "$\Rightarrow$": Let $\sigma \in (\exists P_{=1}r.A)^{\mathcal{I},w}$, thus there is a domain element σ' with $\sigma' \in A^{\mathcal{I},w}$ and $(\sigma, \sigma') \in r^{\mathcal{I},v}$ for all $v \in W$. By the induction hypothesis, $A \in \pi(\sigma', w)$. Consider now the worlds $0, \varepsilon \in W$: By definition of the interpretation of r, it follows from $(\sigma, \sigma') \in r^{\mathcal{I},0} \cap r^{\mathcal{I},\varepsilon}$ that $\sigma' = \sigma \cdot (P_{=1}r, v, B)$ for some world $v \in W$ and a concept name B. By Point (iv), this together with $A \in \pi(\sigma', w)$ yields $\exists P_{=1}r.A \in \pi(\sigma, w)$.

  "$\Leftarrow$": Let $\exists P_{=1}r.A \in \pi(\sigma, w)$. By rule 1 of the construction, there is a domain element $\sigma' = \sigma \cdot (P_{=1}r, w, A)$ with $\pi(\sigma', w) = Q(A)$, thus $A \in \pi(\sigma', w)$ and $\sigma' \in A^{\mathcal{I},w}$.

By definition of the interpretation of role names, $(\sigma, \sigma') \in r^{\mathcal{I},v}$ for all $v \in W$. Hence, $\sigma \in (\exists P_{=1} r.A)^{\mathcal{I},w}$.

- $C = \exists P_{>0} r.A$. "$\Rightarrow$": Let $\sigma \in (\exists P_{>0} r.A)^{\mathcal{I},w}$, thus there is a $\sigma' \in \Delta^{\mathcal{I}}$ with $\sigma' \in A^{\mathcal{I},w}$ and $(\sigma, \sigma') \in r^{\mathcal{I},v}$ for some $v \in W$. By the induction hypothesis, $A \in \pi(\sigma', w)$. Again, we distinguish the three cases of the interpretation of the roles.
  - $\sigma' = \sigma \cdot (P_{>0} r, v', B)$ and $v = (\sigma, r, B)$ for some concept name B. It follows immediately from Point (iv) that $\exists P_{>0} r.A \in \pi(\sigma, w)$.
  - $\sigma' = \sigma \cdot (r, v, B)$ for some concept name B. By construction, in particular rule 1, we have $\exists r.B \in \pi(\sigma, w)$ and $\pi(\sigma', w) = Q(B)$. Since $\pi(\sigma, w)$ is closed under cl, rule **R6** implies $\exists r.A \in \pi(\sigma, w)$. Thus, by **R5**, $\exists P_{>0} r.A \in \pi(\sigma, w)$.
  - $\sigma' = \sigma \cdot (P_{=1} r, v', B)$. Applying Point (iv) yields $\exists P_{=1} r.A \in \pi(\sigma, w)$. Since $\pi(\sigma, w)$ is closed under cl, by rule **R4**, we obtain $\exists P_{>0} r.A \in \pi(\sigma, w)$.

  "$\Leftarrow$". Let $\exists P_{>0} r.A \in \pi(\sigma, w)$. On the one hand, by rule 1 of the construction there is a domain element $\sigma' = \sigma \cdot (P_{>0} r, w, A)$ with $\pi(\sigma', w) = Q(A)$. By hypothesis we get $\sigma' \in A^{\mathcal{I},w}$. On the other hand, Point (i) implies $(A_n, A) \in R(P_{>0} r)$. Thus, by rule 3(a), the world $v = (\sigma, r, A)$ exists. By definition of the interpretation of role names $(\sigma, \sigma') \in r^{\mathcal{I},v}$ for $v = (\sigma, A, r)$. Hence $\sigma \in (\exists P_{>0} r.A)^{\mathcal{I},w}$.

References

Acar, E. (2014). Computing subjective expected utility using probabilistic description logics. In *Proceedings of the 7th european starting AI researcher symposium (STAIRS)* (pp. 21–30).

Artale, A., Kontchakov, R., Lutz, C., Wolter, F., & Zakharyaschev, M. (2007). Temporalising tractable description logics. In *Proceedings of the fourteenth international symposium on temporal representation and reasoning (time07).* IEEE Computer Society Press.

Artale, A., Lutz, C., & Toman, D. (2007). A description logic of change. In *Proceedings of the international joint conference on artificial intelligence (IJCAI)* (p. 218-223).

Baader, F. (2003). Terminological cycles in a description logic with existential restrictions. In G. Gottlob & T. Walsh (Eds.), *Proceedings of the 18th international joint conference on artificial intelligence* (pp. 325–330). Morgan Kaufmann.

Baader, F., Brandt, S., & Lutz, C. (2005). Pushing the $\mathcal{EL}$ envelope. In *Proceedings of the international joint conference on artificial intelligence (IJCAI)* (pp. 364–369).

Baader, F., Calvanese, D., McGuinness, D. L., Nardi, D., & Patel-Schneider, P. F. (Eds.). (2003). *The description logic handbook.* Cambridge University Press.

Bacchus, F., Grove, A. J., Koller, D., & Halpern, J. Y. (1992). From statistics to beliefs. In *Proceedings of the 10th national conference on artificial intelligence (AAAI)* (pp. 602–608).

Bodenreider, O., Smith, B., & Burgun, A. (2004). The ontology-epistemology divide: A case study in medical terminology. In *Proceedings of formal ontology in information systems (FOIS)* (pp. 185–195).

Borgwardt, S., Distel, F., & Peñaloza, R. (2015). The limits of decidability in fuzzy description logics with general concept inclusions. *Artif. Intell.*, *218*, 23–55.

Brázdil, T., Forejt, V., Kretínský, J., & Kucera, A. (2008). The satisfiability problem for probabilistic ctl. In *Proceedings of logic in computer science, (LICS)* (p. 391-402).

Canny, J. F. (1988). Some algebraic and geometric computations in PSPACE. In *Symposium on theory of computing (STOC)* (pp. 460–467).

Ceylan, İ. İ., & Peñaloza, R. (2014). The bayesian description logic $\mathcal{BEL}$. In *Proceedings of the 7th international joint conference on automated reasoning (IJCAR)* (pp. 480–494).

Ceylan, İ. İ., & Peñaloza, R. (2015). Probabilistic query answering in the bayesian description logic $\mathcal{BEL}$. In *Proceedings of the international conference on scalable uncertainty management (SUM)* (pp. 21–35).

Chvátal, V. (1983). *Linear programming.* W. H. Freeman.

Cozman, F. G., Polastro, R. B., Takiyama, F. I., & Revoredo, K. C. (2014). Computing inferences for relational bayesian networks based on *ALC* constructs. In *Uncertainty reasoning for the semantic web III - ISWC international workshops, URSW 2011-2013, revised selected papers* (pp. 21–40).

da Costa, P. C. G., & Laskey, K. B. (2006). PR-OWL: A framework for probabilistic ontologies. In *Proceedings of formal ontology in information systems (FOIS)* (Vol. 150, pp. 237–249).

Dalvi, N. N., & Suciu, D. (2012). The dichotomy of probabilistic inference for unions of conjunctive queries. *Journal of the ACM*, *59*(6), 30.

d'Amato, C., Fanizzi, N., & Lukasiewicz, T. (2008). Tractable reasoning with bayesian description logics. In *Proceedings of the international conference on scalable uncertainty management (SUM)* (pp. 146–159).

de Salvo Braz, R., Amir, E., & Roth, D. (2008). A survey of first-order probabilistic models. In *Innovations in bayesian networks: Theory and applications* (p. 289-317).

Ding, Z., Peng, Y., & Pan, R. (2006). BayesOWL: Uncertainty modeling in semantic web ontologies. In *Soft computing in ontologies and semantic web* (p. 3-29).

Ecke, A., Peñaloza, R., & Turhan, A. (2014). Completion-based generalization inferences for the description logic ELOR with subjective probabilities. *Int. J. Approx. Reasoning*, *55*(9), 1939–1970.

Fagin, R., Halpern, J., & Megiddo, N. (1990). A logic for reasoning about probabilities. *Inform. Comput.*, *87*, 78–128.

Fuhr, N., & Rölleke, T. (1997). A probabilistic relational algebra for the integration of information retrieval and database systems. *ACM Transactions on Information Systems*, *15*(1), 32-66.

Gabbay, D. M., Kurucz, Á., Wolter, F., & Zakharyaschev, M. (2003). *Many-dimensional modal logics: Theory and applications*. Elsevier.

Gottlob, G., Lukasiewicz, T., Martinez, M. V., & Simari, G. I. (2013). Query answering under probabilistic uncertainty in datalog+/- ontologies. *Ann. Math. Artif. Intell.*, *69*(1), 37–72.

Grove, A. J., Halpern, J. Y., & Koller, D. (1994). Random worlds and maximum entropy. *J. Artif. Intell. Res.*, *2*, 33–88.

Gutiérrez-Basulto, V., Jung, J. C., Lutz, C., & Schröder, L. (2011). A closer look at the probabilistic description logic prob-$\mathcal{EL}$. In *Proceedings of the AAAI conference on artificial intelligence (AAAI).*

Gutiérrez-Basulto, V., Jung, J. C., & Schneider, T. (2014). Lightweight description logics and branching time: A troublesome marriage. In *Proceedings of 14th international conference on principles of knowledge representation and reasoning (KR).*

Halpern, J. Y. (1990). An analysis of first-order logics of probability. *Artif. Intell*, *46*, 311–350.

Halpern, J. Y., & Rabin, M. O. (1987). A logic to reason about likelihood. *Artif. Intell.*, *32*, 379–405.

Heinsohn, J. (1994). Probabilistic description logics. In *Proceedings of the 10th conference on uncertainty in artificial intelligence (UAI)* (pp. 311–318).

Herzig, A. (2003). Modal probability, belief, and actions. *Fund. Inf.*, *57*, 323–344.

Jaeger, M. (1994). Probabilistic reasoning in terminological logics. In *Proceedings of principles of knowledge representation and reasoning (KR)* (pp. 305–316).

Jung, J. C., & Lutz, C. (2012). Ontology-based access to probabilistic data with owl-ql. In *Proceedings of the 11th international semantic web conference (ISWC)* (pp. 182–197).

Jung, J. C., Lutz, C., Goncharov, S., & Schröder, L. (2014). Monodic fragments of probabilistic first-order logic. In *Proceedings of the international colloquium on automata, languages, and programming (ICALP)* (pp. 256–267).

Kazakov, Y. (2009). Consequence-driven reasoning for horn shiq ontologies. In *Proceedings of the international joint conference on artificial intelligence (ICJAI)* (pp. 2040–2045).

Klinov, P., & Parsia, B. (2013). Pronto: A practical probabilistic description logic reasoner. In *Proceedings of the international workshop on uncertainty reasoning for the semantic web (URSW)* (pp. 59–79).

Klinov, P., Parsia, B., & Sattler, U. (2009). On correspondences between probabilistic first-order and description logics. In *Description logics* (Vol. 477). CEUR-WS.org.

Koller, D., & Halpern, J. Y. (1996). Irrelevance and conditioning in first-order probabilistic logic. In *Proceedings of the national conference on artificial intelligence (AAAI).*

Koller, D., Levy, A., & Pfeffer, A. (1997). P-Classic: A tractable probabilistic description logic. In *Proceedings of the national conference on artificial intelligence (AAAI)* (p. 390-397).

Kriegel, F. (2015). Axiomatization of general concept inclusions in probabilistic description logics. In *Proceedings of the 38th german conference on artificial intelligence (KI 2015)* (Vol. 9324, pp. 124–136).

Lukasiewicz, T. (2007). Probabilistic description logic programs. *Int. J. Approx. Reasoning*, *45*(2), 288–307.

Lukasiewicz, T. (2008). Expressive probabilistic description logics. *Artif. Intell*, *172*, 852–883.

Lukasiewicz, T., & Straccia, U. (2008). Managing uncertainty and vagueness in description logics for the semantic web. *J. Web Sem.*, *6*(4), 291-308.

Lutz, C., & Schröder, L. (2010). Probabilistic description logics for subjective uncertainty. In *Proceedings of the international conference on principles of knowledge representation and reasoning (KR).*

Lutz, C., Wolter, F., & Zakharyaschev, M. (2008). Temporal description logics: A survey. In *International symposium on temporal representation and reasoning (TIME)* (p. 3-14).

Mauá, D. D., & Cozman, F. G. (2015). Dl-lite bayesian networks: A tractable probabilistic graphical model. In *Proceedings of the 9th international conference on scalable uncertainty management (SUM)* (pp. 50–64).

Niepert, M., Noessner, J., & Stuckenschmidt, H. (2011). Log-linear description logics. In *Proceedings of the international joint conference on artificial intelligence IJCAI* (pp. 2153–2158).

Peñaloza, R., & Potyka, N. (2016). Probabilistic reasoning in the description logic *ALCP* with the principle of maximum entropy. In *Proceedings of the 10th international conference on scalable uncertainty management (SUM)* (pp. 246–259).

Pratt, V. R. (1979). Models of program logics. In *Proceedings of annual symposium on foundations of computer science (FOCS)* (pp. 115–122).

Price, C., & Spackman, K. (2000). SNOMED Clinical Terms. *Br. J. Healthcare Comput. Inf. Mgmt.*, *17*, 27–31.

Rector, A., & Horrocks, I. (1997). Experience building a large, re-usable medical ontology using a description logic with transitivity and concept inclusions. In *Proceedings of the workshop on ontological engineering, AAAI spring symposium (AAAI).*

Riguzzi, F., Bellodi, E., Lamma, E., & Zese, R. (2015). Probabilistic description logics under the distribution semantics. *Semantic Web*, *6*(5), 477–501.

Schaefer, M. (2010). Complexity of some geometric and topological problems. In *Proceedings of the 17th international conference on graph drawing (GD)* (pp. 334–344).

Schrijver, A. (1986). *Theory of linear and integer programming.* Wiley Interscience.

Schröder, L. (2007). A finite model construction for coalgebraic modal logic. *J. Log. Algebr. Prog.*, *73*, 97–110.

Sebastiani, F. (1994). A probabilistic terminological logic for modelling information retrieval. In *ACM-SIGIR conference on research and development in information retrieval, (SIGIR'94)* (pp. 122–130).

Suciu, D., Olteanu, D., Ré, C., & Koch, C. (2011). *Probabilistic databases.* Morgan & Claypool Publishers.

W3C. (2012). *Owl 2: The web ontology language.* Retrieved 2016-12-12, from `https://www.w3.org/TR/owl2-overview/`

Yelland, P. M. (2000). An alternative combination of Bayesian networks and description logics. In *Proceedings of principles of knowledge representation and reasoning (KR)* (pp. 225–233).