Statistical $\mathcal{EL}$ is ExpTime-complete

Bartosz Bednarczyk

Computational Logic Group, Technische Universität Dresden, Germany
Institute of Computer Science, University of Wrocław, Poland

Abstract

We show that the consistency problem for Statistical $\mathcal{EL}$ ontologies, defined by Peñaloza and Potyka, is ExpTime-hard. Together with existing ExpTime upper bounds, we conclude ExpTime-completeness of the logic. Our proof goes via a reduction from the consistency problem for $\mathcal{EL}$ extended with negation of atomic concepts.

1 Introduction

Description logics (DLs) [BHLS17] are a prominent family of logical formalisms tailored to knowledge representation. Nowadays, real-world problems require the ability to handle uncertain knowledge. To deal with this issue, several probabilistic extensions of description logics were proposed in the past [CLC17, Luk08, GBJLS17, PP17] Among such extensions, the authors of [PP17] proposed Statistical $\mathcal{EL}$, a statistical variant of the well-known description logic $\mathcal{EL}$ [BBL05] famous for tractability of most of its reasoning task.

In this note we establish tight complexity bounds for the consistency problem for statistical $\mathcal{EL}$, closing the complexity gaps from [PP17]. We show that in sharp contrast to its non-probabilistic version, Statistical $\mathcal{EL}$ is ExpTime-complete and hence, provably intractable. The main novelty here is the ExpTime lower bound, while the ExpTime upper bound follows from recent work by Baader and Ecke [BE17, Corollary 15] or, alternatively, from work on probabilistic $\mathcal{ALC}$ by Lutz and Schröder [LS10, Theorem 9].

2 Preliminaries

In this section, we recall the basics on description logics (DLs) $\mathcal{EL}$ and $\mathcal{EL}^{(\neg)}$. For readers unfamiliar with DLs we recommend consulting the textbook [BHLS17], especially Chapters 2.1–2.3, 5.1 and 6.1.

We fix countably-infinite disjoint sets of *concept names* $\mathbf{N_C}$ and *role names* $\mathbf{N_R}$. Starting from $\mathbf{N_C}$ and $\mathbf{N_R}$, the set $\mathbf{C}_{\mathcal{EL}}$ of $\mathcal{EL}$ *concept descriptions* (or simply $\mathcal{EL}$ concepts) [BBL05] is built using *conjunction* ($\mathrm{C} \sqcap \mathrm{D}$), *existential restriction* ($\exists r.\mathrm{C}$) and the *top concept* ($\top$), with the grammar below:

$$\mathrm{C}, \mathrm{D} \ ::= \ \top \ | \ \mathrm{A} \ | \ \mathrm{C} \sqcap \mathrm{D} \ | \ \exists r.\mathrm{C},$$

where $\mathrm{C}, \mathrm{D} \in \mathbf{C}_{\mathcal{EL}}$, $\mathrm{A} \in \mathbf{N_C}$ and $r \in \mathbf{N_R}$. An $\mathcal{EL}$ *general concept inclusion* (GCI) has the form $\mathrm{C} \sqsubseteq \mathrm{D}$ for $\mathcal{EL}$ concepts $\mathrm{C}, \mathrm{D} \in \mathbf{C}_{\mathcal{EL}}$. An $\mathcal{EL}$ *ontology* is a finite non-empty set of $\mathcal{EL}$ GCIs. The *size* of an $\mathcal{EL}$ ontology is the total number of $\top$, role names, concept names and connectives occurring in it.

Table 1: Concepts and roles in $\mathcal{EL}$.

Name	Syntax	Semantics
top	$\top$	$\Delta^{\mathcal{I}}$
atomic concept	A	$\mathrm{A}^{\mathcal{I}} \subseteq \Delta^{\mathcal{I}}$
role	r	$r^{\mathcal{I}} \subseteq \Delta^{\mathcal{I}} \times \Delta^{\mathcal{I}}$
concept intersection	$\mathrm{C} \sqcap \mathrm{D}$	$\mathrm{C}^{\mathcal{I}} \cap \mathrm{D}^{\mathcal{I}}$
existential restriction	$\exists r.\mathrm{C}$	$\{\mathrm{d} \mid \exists \mathrm{e}.(\mathrm{d}, \mathrm{e}) \in r^{\mathcal{I}} \wedge \mathrm{e} \in \mathrm{C}^{\mathcal{I}}\}$

The semantics of $\mathcal{EL}$ is defined via *interpretations* $\mathcal{I} = (\Delta^{\mathcal{I}}, \cdot^{\mathcal{I}})$ composed of a *finite* non-empty set $\Delta^{\mathcal{I}}$ called the *domain* of $\mathcal{I}$ and an *interpretation function* $\cdot^{\mathcal{I}}$ mapping concept names to subsets of $\Delta^{\mathcal{I}}$, and role names to subsets of $\Delta^{\mathcal{I}} \times \Delta^{\mathcal{I}}$. This mapping is extended to *concepts*, *roles* (*cf.* Table 1) and finally used to define satisfaction of GCIs, namely $\mathcal{I} \models \mathrm{C} \sqsubseteq \mathrm{D}$ iff $\mathrm{C}^{\mathcal{I}} \subseteq \mathrm{D}^{\mathcal{I}}$. We say that an interpretation $\mathcal{I}$ *satisfies* an ontology $\mathcal{O}$ (or $\mathcal{I}$ is a model of $\mathcal{O}$, written: $\mathcal{I} \models \mathcal{O}$) if it satisfies all GCIs from $\mathcal{O}$. An ontology is *consistent* if it has a model and inconsistent otherwise. In the *consistency problem* for $\mathcal{EL}$ we ask if an input $\mathcal{EL}$ ontology is consistent. Note that the consistency problem for $\mathcal{EL}$ is trivial, *i.e.* every $\mathcal{EL}$ ontology is consistent.

2.1 $\mathcal{EL}$ with atomic negation

The next definitions concern $\mathcal{EL}^{(\neg)}$, the extension of $\mathcal{EL}$ with negation of atomic concepts. More precisely, the set $\mathbf{C}_{\mathcal{EL}^{(\neg)}}$ of $\mathcal{EL}^{(\neg)}$ *concepts* is defined by a slight extension of the BNF grammar for $\mathcal{EL}$:

$$\mathrm{C}, \mathrm{D} \ ::= \ \top \ | \ \mathrm{A} \ | \ \bar{\mathrm{A}} \ | \ \mathrm{C} \sqcap \mathrm{D} \ | \ \exists r.\mathrm{C},$$

where $\mathrm{C}, \mathrm{D} \in \mathbf{C}_{\mathcal{EL}^{(\neg)}}$, $\mathrm{A} \in \mathbf{N_C}$ and $r \in \mathbf{N_R}$. The semantics of $\mathcal{EL}^{(\neg)}$ concepts is defined as in Table 1 with the exception that the concepts of the form $\bar{\mathrm{A}}$ have the semantics $\bar{\mathrm{A}}^{\mathcal{I}} = \Delta^{\mathcal{I}} \setminus \mathrm{A}^{\mathcal{I}}$. The notions of GCIs, ontologies and the consistency problem are lifted to $\mathcal{EL}^{(\neg)}$ in an obvious way. We stress that in the presence of negation the consistency problem for $\mathcal{EL}^{(\neg)}$ is no longer trivial and actually is ExpTime-complete [BBL05, Theorem 6].[1]

Proposition 2.1. *The consistency problem for $\mathcal{EL}^{(\neg)}$ ontologies is* ExpTime*-hard.*

2.2 Statistical $\mathcal{EL}$

Statistical $\mathcal{EL}$, abbreviated as $\mathcal{SEL}$, is a probabilistic DL introduced recently by Peñaloza and Potyka [PP17, Section 4] to reason about statistical properties over finite domains. Statistical $\mathcal{EL}$ ontologies are composed of *probabilistic conditionals* of the form $(\mathrm{C} \mid \mathrm{D})\,[k,l]$, where C, D are $\mathcal{EL}$ concepts from $\mathbf{C}_{\mathcal{EL}}$ and $k, l \in \mathbb{Q}$ are rational numbers satisfying $0 \le k \le l \le 1$. The size of $\mathcal{SEL}$ ontologies is defined as in $\mathcal{EL}$ except that the numbers in probabilistic conditionals also contribute to the size and are measured in binary.

We say that an interpretation $\mathcal{I}$ satisfies a probabilistic conditional $(\mathrm{C} \mid \mathrm{D})\,[k,l]$ if:

$$\text{either } \mathrm{D}^{\mathcal{I}} = \emptyset \text{ or } k \le \frac{|(\mathrm{C} \sqcap \mathrm{D})^{\mathcal{I}}|}{|\mathrm{D}^{\mathcal{I}}|} \le l.$$

Note that usual $\mathcal{EL}$ GCIs $\mathrm{C} \sqsubseteq \mathrm{D}$ are equivalent to $(\mathrm{D} \mid \mathrm{C})\,[1,1]$ (*cf.* [PP17, Proposition 4]). Hence, each $\mathcal{EL}$ ontology can be seen as an $\mathcal{SEL}$ ontology and we can freely use GCIs in place of probabilistic conditionals.

3 Main result

After introducing all the required definitions, we are ready to prove the main result of this note, namely:

Theorem 3.1. *The consistency problem for $\mathcal{SEL}$ is* ExpTime*-complete.*

The ExpTime upper bound follows from [BE17, Corollary 15] or [LS10, Theorem 9], hence we focus on the lower bound only. Let $\mathcal{O}$ be an arbitrary $\mathcal{EL}^{(\neg)}$ ontology. With $\mathbf{C}_{\mathcal{O}}$ we denote the set of all concept names that appear (possibly under negation) in $\mathcal{O}$. We next design an $\mathcal{SEL}$ ontology $\mathcal{O}_{red}$ such that $\mathcal{O}_{red}$ is consistent iff $\mathcal{O}$ is and that $\mathcal{O}_{red}$ is only polynomially larger than $\mathcal{O}$. It will be composed of two $\mathcal{SEL}$ ontologies, $\mathcal{O}_{tr}$ and $\mathcal{O}_{corr}$, responsible respectively for "translating" $\mathcal{O}$ into $\mathcal{SEL}$ and for guaranteeing the correctness of the translation.

The main idea of the encoding is as follows. We first produce for each concept name A from $\mathbf{C}_{\mathcal{O}}$ two fresh, different, concepts $\mathrm{A}_+, \mathrm{A}_- \notin \mathbf{C}_{\mathcal{O}}$ intuitively intended to contain, respectively, all members of A and from its complement. Due to the lack of negation, we clearly are not able to fully formalise the above intuition, but the best we can do is to enforce, with the ontology $\mathcal{O}_{corr}$, that these concepts are interpreted as disjoint sets and each of them contains exactly half of the domain. This is sufficient for our purposes, since with fresh, pair-wise different, concepts $\mathrm{Real}, \mathrm{Real}_+, \mathrm{Real}_- \notin \mathbf{C}_{\mathcal{O}}$ we can separate the "real" model of $\mathcal{O}$ from the auxiliary parts required for the encoding. Finally, in the "translation" ontology $\mathcal{O}_{tr}$ we state that the restriction of a model of $\mathcal{O}_{red}$ to Real_+ satisfies $\mathcal{O}$. The translation simply changes all occurrences of A (resp. $\bar{\mathrm{A}}$) into A_+ (resp. A_-) and employs Real_+ to relativise concepts.

We start with the definition of $\mathcal{O}_{corr}$.

$$\mathcal{O}_{corr} := \{(\mathrm{A}_+ \mid \top)\,[0.5, 0.5]\,, (\mathrm{A}_- \mid \top)\,[0.5, 0.5]\,, (\mathrm{A}_+ \mid \mathrm{A}_-)\,[0, 0] \ \mid \ \mathrm{A} \in \{\mathrm{Real}\} \cup \mathbf{C}_{\mathcal{O}}\}$$

By unfolding the definition of probabilistic conditionals we immediately conclude the following facts.

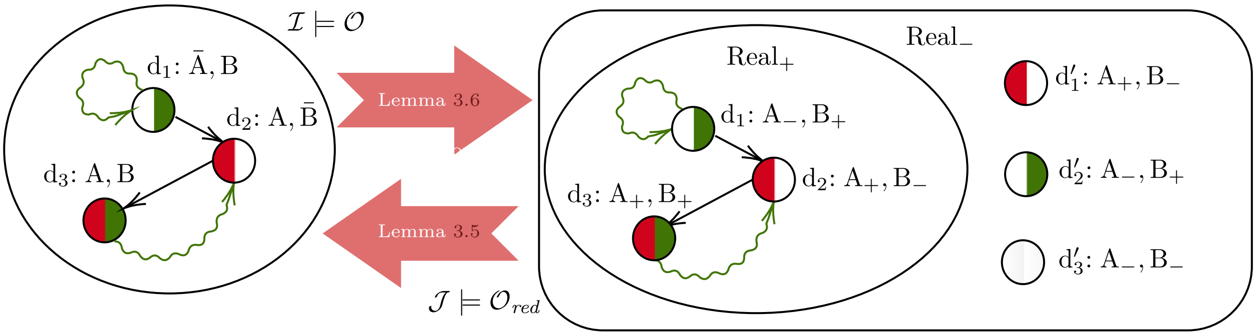
Fact 3.2. *For any concept name* A *we have that* $\mathcal{I} \models (\mathrm{A} \mid \top)\,[0.5, 0.5]$ *iff* $|\Delta^{\mathcal{I}}|$ *is even and* $|\mathrm{A}^{\mathcal{I}}| = \frac{1}{2}|\Delta^{\mathcal{I}}|$.

Fact 3.3. *For any different concept names* A, B *we have that* $\mathcal{I} \models (\mathrm{A} \mid \mathrm{B})\,[0, 0]$ *iff* $\mathrm{A}^{\mathcal{I}} \cap \mathrm{B}^{\mathcal{I}} = \emptyset$.

[1] In the setting of [BBL05] the domains of interpretations might have unrestricted (*i.e.* not necessarily finite) sizes, but the result is also applicable to our scenario since $\mathcal{EL}^{(\neg)}$ has finite model property, which follows from [BHLS17, Corr. 3.17].

Now we focus on the "translating" ontology $\mathcal{O}_{tr}$. Let $\mathfrak{tr}$ be a translation function defined by $\mathfrak{tr}(\top) = \mathrm{Real}_+$, $\mathfrak{tr}(\mathrm{A}) = \mathrm{A}_+ \sqcap \mathrm{Real}_+$ and $\mathfrak{tr}(\bar{\mathrm{A}}) = \mathrm{A}_- \sqcap \mathrm{Real}_+$ for all concept names $\mathrm{A} \in \mathbf{N_C}$ as well as $\mathfrak{tr}(\mathrm{C} \sqcap \mathrm{D}) = \mathfrak{tr}(\mathrm{C}) \sqcap \mathfrak{tr}(\mathrm{D})$ and $\mathfrak{tr}(\exists r.\mathrm{C}) = \mathrm{Real}_+ \sqcap \exists r.(\mathfrak{tr}(\mathrm{C}) \sqcap \mathrm{Real}_+)$ for complex concepts. The ontology $\mathcal{O}_{tr}$ is obtained by replacing each GCI $\mathrm{C} \sqsubseteq \mathrm{D}$ from $\mathcal{O}$ with $\mathfrak{tr}(\mathrm{C}) \sqsubseteq \mathfrak{tr}(\mathrm{D})$. Finally, we put $\mathcal{O}_{red} := \mathcal{O}_{corr} \cup \mathcal{O}_{tr}$.

Note that the size of $\mathcal{O}_{red}$ is polynomial in $|\mathcal{O}|$. For more intuitions, consult the picture below.



3.1 Correctness of the reduction

Let us start with an auxiliary notion of interpretations that are *good-for-encoding*. We say that $\mathcal{J}$ is good-for-encoding if for all concept names $\mathrm{A} \in \mathbf{C}_{\mathcal{O}}$ it satisfies $\mathrm{A}^{\mathcal{J}} = \mathrm{A}_+^{\mathcal{J}}$ and $\mathrm{A}_-^{\mathcal{J}} = \Delta^{\mathcal{J}} \setminus \mathrm{A}^{\mathcal{J}}$. The following lemma relates the translation function $\mathfrak{tr}$, good-for-encoding interpretations and their submodels.

Lemma 3.4 (Agreement lemma). *Let $\mathcal{J}$ be good-for-encoding and let $\mathcal{I}$ be its induced subinterpretation with domain $\mathrm{Real}_+^{\mathcal{J}}$. Then all $\mathcal{EL}^{(\neg)}$ concepts* C *employing* only *concept names from $\mathbf{C}_{\mathcal{O}}$ satisfy $\mathrm{C}^{\mathcal{I}} = \mathfrak{tr}(\mathrm{C})^{\mathcal{J}}$. Moreover for such concepts* C, D *we have: $\mathcal{J} \models \mathfrak{tr}(\mathrm{C}) \sqsubseteq \mathfrak{tr}(\mathrm{D})$ iff $\mathcal{I} \models \mathrm{C} \sqsubseteq \mathrm{D}$.*

Proof. We proceed by inductively on the shape of concepts C. The cases for $\mathrm{C} = \top, \mathrm{A}$ or $\bar{\mathrm{A}}$ for $\mathrm{A} \in \mathbf{N_C}$ follow immediately from the definition of $\mathfrak{tr}$ and the assumptions $\mathrm{A}^{\mathcal{J}} = \mathrm{A}_+^{\mathcal{J}}$ and $\mathrm{A}_-^{\mathcal{J}} = \Delta^{\mathcal{J}} \setminus \mathrm{A}^{\mathcal{J}}$. The case of $\mathrm{C} = \mathrm{D} \sqcap \mathrm{E}$ follows from the fact that $\mathfrak{tr}$ is homomorphic for $\sqcap$. Hence, the only interesting case is when $\mathrm{C} = \exists r.\mathrm{D}$. Assuming $\mathrm{D}^{\mathcal{I}} = \mathfrak{tr}(\mathrm{D})^{\mathcal{J}}$ we will show two inclusions.

- For the first inclusion, take $\mathrm{d} \in (\exists r.\mathrm{D})^{\mathcal{I}}$. Thus $\mathrm{d} \in \Delta^{\mathcal{I}} (= \mathrm{Real}_+^{\mathcal{J}})$ and there exists an $\mathrm{e} \in \Delta^{\mathcal{I}}$ satisfying both $(\mathrm{d}, \mathrm{e}) \in r^{\mathcal{I}}$ and $\mathrm{e} \in \mathrm{D}^{\mathcal{I}}$. Note that $\mathrm{e} \in \Delta^{\mathcal{I}}$ implies $\mathrm{e} \in \mathrm{Real}_+^{\mathcal{J}}$. Moreover, by the equality $\mathrm{D}^{\mathcal{I}} = \mathfrak{tr}(\mathrm{D})^{\mathcal{J}}$ we have $\mathrm{e} \in (\mathrm{Real}_+ \sqcap \mathfrak{tr}(\mathrm{D}))^{\mathcal{J}}$. Since $r^{\mathcal{I}} \subseteq r^{\mathcal{J}}$ we infer $\mathrm{d} \in (\exists r.(\mathrm{Real}_+ \sqcap \mathfrak{tr}(\mathrm{D})))^{\mathcal{J}}$, but because $\mathrm{d} \in \mathrm{Real}_+^{\mathcal{J}}$ we can conclude that $\mathrm{d} \in (\mathrm{Real}_+ \sqcap \exists r.(\mathrm{Real}_+ \sqcap \mathfrak{tr}(\mathrm{D})))^{\mathcal{J}} = \mathfrak{tr}(\exists r.\mathrm{D})^{\mathcal{J}}$.
- For the opposite direction take $\mathrm{d} \in \mathfrak{tr}(\exists r.\mathrm{D})^{\mathcal{J}} = (\mathrm{Real}_+ \sqcap \exists r.(\mathrm{Real}_+ \sqcap \mathfrak{tr}(\mathrm{D})))^{\mathcal{J}}$. It implies that $\mathrm{d} \in \mathrm{Real}_+^{\mathcal{J}} (= \Delta^{\mathcal{I}})$ as well as that there is an $\mathrm{e} \in \mathrm{Real}_+^{\mathcal{J}} (= \Delta^{\mathcal{I}})$ witnessing $(\mathrm{d}, \mathrm{e}) \in r^{\mathcal{J}}$ and $\mathrm{e} \in \mathfrak{tr}(\mathrm{D})^{\mathcal{J}} (= \mathrm{D}^{\mathcal{I}})$. Since both d, e belong to $\Delta^{\mathcal{I}}$ we infer that $(\mathrm{d}, \mathrm{e}) \in r^{\mathcal{I}}$ and hence $\mathrm{d} \in (\exists r.\mathrm{D})^{\mathcal{I}}$.

For the last statement of the lemma: to show that $\mathcal{J} \models \mathfrak{tr}(\mathrm{C}) \sqsubseteq \mathfrak{tr}(\mathrm{D})$ iff $\mathcal{I} \models \mathrm{C} \sqsubseteq \mathrm{D}$ hold, it suffices to invoke $\mathrm{C}^{\mathcal{I}} = \mathfrak{tr}(\mathrm{C})^{\mathcal{J}}$ and $\mathrm{D}^{\mathcal{I}} = \mathfrak{tr}(\mathrm{D})^{\mathcal{J}}$ to see that the inclusions $\mathrm{C}^{\mathcal{I}} \subseteq \mathrm{D}^{\mathcal{I}}$ and $\mathfrak{tr}(\mathrm{C})^{\mathcal{J}} \subseteq \mathfrak{tr}(\mathrm{D})^{\mathcal{J}}$ are equivalent. □

The agreement lemma can now be used to show that the consistency of $\mathcal{O}_{red}$ implies the consistency of $\mathcal{O}$.

Lemma 3.5. *If $\mathcal{O}_{red}$ is consistent then so is $\mathcal{O}$.*

Proof. Let $\mathcal{J}$ be a model of $\mathcal{O}_{red}$ with $\mathrm{A}^{\mathcal{J}} := \mathrm{A}_+^{\mathcal{J}}$ (since A does not appear in $\mathcal{O}_{red}$ this can be assumed w.l.o.g.) for all concept names A from $\mathbf{C}_{\mathcal{O}}$. By the satisfaction of $\mathcal{O}_{corr}$ we know that $\mathrm{A}_+^{\mathcal{J}}$ and $\mathrm{A}_-^{\mathcal{J}}$ are disjoint and thus $\mathcal{J}$ is good-for-encoding. Hence, take $\mathcal{I}$ to be its induced subinterpretation with domain $\mathrm{Real}_+^{\mathcal{J}}$. By applying Lemma 3.4 we know that for each GCI $\mathrm{C} \sqsubseteq \mathrm{D}$ from $\mathcal{O}$ the satisfaction of $\mathcal{J} \models \mathfrak{tr}(\mathrm{C}) \sqsubseteq \mathfrak{tr}(\mathrm{D})$ implies $\mathcal{I} \models \mathrm{C} \sqsubseteq \mathrm{D}$. Thus we get $\mathcal{I} \models \mathcal{O}$, which implies that $\mathcal{O}$ is consistent. □

We next show that the consistency of $\mathcal{O}$ implies the consistency of $\mathcal{O}_{red}$. In the proof we basically take a model of $\mathcal{O}$, duplicate each domain element and define the memberships of fresh concepts introduced by $\mathcal{O}_{red}$. Such concepts are defined in such a way that if an element from a model $\mathcal{I}$ of $\mathcal{O}$ is a member of $\mathrm{A}^{\mathcal{I}}$ then the corresponding element in a constructed model $\mathcal{J}$ for $\mathcal{O}_{red}$ is a member of $\mathrm{A}_+^{\mathcal{J}}$ while its copy belongs to $\mathrm{A}_-^{\mathcal{J}}$. In this way, the total number of elements in every concept is always equal to half of the domain.

Lemma 3.6. *If $\mathcal{O}$ is consistent then so is $\mathcal{O}_{red}$.*

Proof. Let $\mathcal{I} \models \mathcal{O}$ and let $\Delta^{\mathcal{I}} = \{\mathrm{d}_1, \mathrm{d}_2, \ldots, \mathrm{d}_n\}$. We define an interpretation $\mathcal{J}$ as follows:

1. $\Delta^{\mathcal{J}} := \{d_1, d_1', d_2, d_2', \ldots, d_n, d_n'\}$.
2. For all concept names $A \in \mathbf{C}_{\mathcal{O}}$ we put
   - $A_+^{\mathcal{J}} := A^{\mathcal{J}} := \{d_i \mid d_i \in A^{\mathcal{I}}\} \cup \{d_i' \mid d_i \not\in A^{\mathcal{I}}\}$ and $A_-^{\mathcal{J}} := \{d_i \mid d_i \not\in A^{\mathcal{I}}\} \cup \{d_i' \mid d_i \in A^{\mathcal{I}}\}$,
   - $\mathrm{Real}_+^{\mathcal{J}} := \Delta^{\mathcal{I}}$ and $\mathrm{Real}_-^{\mathcal{J}} := \Delta^{\mathcal{J}} \setminus \Delta^{\mathcal{I}}$,

   and for all other concept names B we put $B^{\mathcal{I}} := \Delta^{\mathcal{I}}$.
3. For each role name r we put $r^{\mathcal{J}} := r^{\mathcal{I}}$.

We first show $\mathcal{J} \models \mathcal{O}_{corr}$. To this end, take any name $A \in \{\mathrm{Real}\} \cup \mathbf{C}_{\mathcal{O}}$. We prove $\mathcal{J} \models (A_+ \mid A_-)[0,0]$, which by Fact 3.3 is equivalent to showing disjointness of $A_+^{\mathcal{J}}$ and $A_-^{\mathcal{J}}$. Assume the contrary, *i.e.* that there is a domain element $d \in A_+^{\mathcal{J}} \cap A_-^{\mathcal{J}}$. If $d = d_i$ for some index i then, by Item 2, it means that $d_i \in A^{\mathcal{I}}$ and $d_i \not\in A^{\mathcal{I}}$ at the same time, which is clearly not possible. The case when $d = d_i'$ for some index i is treated similarly. Next, by invoking Item 2 of the definition of $\mathcal{J}$ we can perform some basic calculations:

$$|A_+^{\mathcal{J}}| = |\{d_i \mid d_i \in A^{\mathcal{I}}\}| + |\{d_i' \mid d_i \not\in A^{\mathcal{I}}\}| = |\{d_i \mid d_i \in A^{\mathcal{I}}\}| + |\{d_i \mid d_i \not\in A^{\mathcal{I}}\}| = |\Delta^{\mathcal{I}}| = 0.5 \cdot |\Delta^{\mathcal{J}}|.$$

$$|A_-^{\mathcal{J}}| = |\{d_i \mid d_i \not\in A^{\mathcal{I}}\}| + |\{d_i' \mid d_i \in A^{\mathcal{I}}\}| = |\{d_i' \mid d_i \not\in A^{\mathcal{I}}\}| + |\{d_i' \mid d_i \in A^{\mathcal{I}}\}| = |\{d_1', \ldots, d_n'\}| = 0.5 \cdot |\Delta^{\mathcal{J}}|.$$

Hence by Fact 3.2 we get $\mathcal{J} \models (A_+ \mid \top)[0.5, 0.5]$ and $\mathcal{J} \models (A_- \mid \top)[0.5, 0.5]$, finishing the proof of $\mathcal{J} \models \mathcal{O}_{corr}$.

To prove $\mathcal{J} \models \mathcal{O}_{tr}$ we take any GCI $\mathfrak{tr}(C) \sqsubseteq \mathfrak{tr}(D)$ from $\mathcal{O}_{tr}$. From $\mathcal{I} \models \mathcal{O}$, we know that $\mathcal{I} \models C \sqsubseteq D$. Note that $\mathcal{J}$ is good-for-encoding, hence by Lemma 3.4 it follows that $\mathcal{J} \models \mathfrak{tr}(C) \sqsubseteq \mathfrak{tr}(D)$. Thus $\mathcal{J} \models \mathcal{O}_{tr}$. ☐

Lemma 3.6 and Lemma 3.5 show that the presented reduction is correct. Since our reduction is polynomial for every $\mathcal{EL}^{(\neg)}$ ontology $\mathcal{O}$, from Proposition 2.1 we can conclude the main theorem of this note.

Theorem 3.7. *The consistency problem for Statistical $\mathcal{EL}$ is* ExpTime*-hard and remains* ExpTime*-hard even if the only numbers used in probabilistic conditionals are* $0, 0.5$ *and* 1.

With the already mentioned ExpTime upper bound we conclude Theorem 3.1.

4 Conclusions

We have proved that the consistency problem for Statistical $\mathcal{EL}$ ontologies is ExpTime-complete. While the upper bound was derived from the works of Baader and Ecke [BE17] or Lutz and Schröder on (the extensions of) $\mathcal{ALC}$ with cardinality constraints, the main contribution of the paper is the lower bound. Our proof went via a reduction from the consistency problem for $\mathcal{EL}^{(\neg)}$ ontologies and heavily relied on the fact that probabilistic conditionals can express that exactly half of the domain elements belong to a certain concept.

An interesting direction for future work is to consider extensions of other well-known decidable fragments of first-order logic with probabilistic conditionals. Promising candidates are the guarded fragment [ANvB98], the guarded negation fragment [BtCS15], the two-variable logic [GKV97] and tamed fragments of existential rules [MT14]. Another idea is to study query answering [OS12] in the presence of probabilistic conditionals. Some initial results were obtained in [BBR20].

Acknowledgements

This work was supported by the Polish Ministry of Science and Higher Education program “Diamentowy Grant” no. DI2017 006447. The author would like to thank Ania Karykowska and Elisa Böhl for reading the draft of this note as well as Piotr Witkowski for serving as a rubber duck while “rubber duck debugging” the proof ideas. He also grateful for many grammar corrections and improvement ideas given by anonymous IPL reviewers.

References

[ANvB98] Hajnal Andréka, István Németi, and Johan van Benthem. Modal Languages and Bounded Fragments of Predicate Logic. *J. Philos. Log.*, 1998.

[BBL05] Franz Baader, Sebastian Brandt, and Carsten Lutz. Pushing the EL Envelope. In *IJCAI*, 2005.

[BBR20] Franz Baader, Bartosz Bednarczyk, and Sebastian Rudolph. Satisfiability and Query Answering in Description Logics with Global and Local Cardinality Constraints. In *ECAI*, 2020.

[BE17] Franz Baader and Andreas Ecke. Extending the Description Logic ALC with More Expressive Cardinality Constraints on Concepts. In *GCAI*, 2017.

[BHLS17] Franz Baader, Ian Horrocks, Carsten Lutz, and Ulrike Sattler. *An Introduction to Description Logic*. Cambridge University Press, 2017.

[BtCS15] Vince Bárány, Balder ten Cate, and Luc Segoufin. Guarded Negation. *J. ACM*, 2015.

[CLC17] Rommel N. Carvalho, Kathryn B. Laskey, and Paulo C. G. Costa. PR-OWL - a language for defining probabilistic ontologies. *Int. J. Approx. Reason.*, 2017.

[GBJLS17] Víctor Gutiérrez-Basulto, Jean Christoph Jung, Carsten Lutz, and Lutz Schröder. Probabilistic Description Logics for Subjective Uncertainty. *J. Artif. Intell. Res.*, 2017.

[GKV97] Erich Grädel, Phokion G. Kolaitis, and Moshe Y. Vardi. On the decision problem for two-variable first-order logic. *Bull. Symb. Log.*, 1997.

[LS10] Carsten Lutz and Lutz Schröder. Probabilistic Description Logics for Subjective Uncertainty. 2010.

[Luk08] Thomas Lukasiewicz. Expressive probabilistic description logics. *Artif. Intell.*, 2008.

[MT14] Marie-Laure Mugnier and Michaël Thomazo. An Introduction to Ontology-Based Query Answering with Existential Rules. In *Reasoning Web Summer School*, 2014.

[OS12] Magdalena Ortiz and Mantas Simkus. Reasoning and Query Answering in Description Logics. In *Reasoning Web Summer School*, 2012.

[PP17] Rafael Peñaloza and Nico Potyka. Towards Statistical Reasoning in Description Logics over Finite Domains. In *SUM*, 2017.