Interpolants and Explicit Definitions in Horn Description Logics

Marie Fortin, Boris Konev, and Frank Wolter

University of Liverpool, UK

Abstract. We show that the vast majority of Horn Description Logics does not enjoy the Craig interpolation nor the projective Beth definability property. This is the case, for example, for $\mathcal{EL}$ with nominals, with the universal role, or with a role inclusion of the form $r \circ s \sqsubseteq s$, and for $\mathcal{ELI}$, Horn-$\mathcal{ALC}$, and Horn-$\mathcal{ALCI}$. It follows in particular that the existence of an explicit definition of a concept name cannot be reduced to subsumption checking via implicit definability. We show that nevertheless the existence of interpolants and explicit definitions can be decided in polynomial time for typical logics (such as $\mathcal{EL}^{++}$) in the $\mathcal{EL}$-family and in ExpTime for $\mathcal{ELI}$ and various extensions. It is thus not harder than subsumption, a result which is in sharp contrast to the situation in expressive DLs. We also obtain tight bounds on the size of interpolants and explicit definitions.

1 Introduction

The *projective Beth definability property* (PBDP) of a description logic (DL) $\mathcal{L}$ states that a concept or individual name is explicitly definable under an $\mathcal{L}$-ontology $\mathcal{O}$ by an $\mathcal{L}$-concept using symbols from a signature Σ if, and only if, it is implicitly definable by Σ under $\mathcal{O}$. The importance of the PBDP for DL research stems from the fact that it provides a polynomial time reduction of explicit definition existence to subsumption checking and the usefulness of having explicit definitions for numerous knowledge engineering tasks, including the equivalent rewritability of ontology-mediated queries, the construction of alignments between ontologies, the computation of referring expressions, the extraction of equivalent acyclic terminologies from ontologies, and the decomposition of ontologies [14,43,34,28,15,23,3]. The PBDP is often investigated in tandem with the *Craig interpolation property* (CIP) which states that if an $\mathcal{L}$-concept is subsumed by another $\mathcal{L}$-concept under some $\mathcal{L}$-ontology then one finds an interpolating $\mathcal{L}$-concept using the shared symbols of the two input concepts only. In fact, the CIP implies the PBDP and the interpolants obtained using the CIP can serve as explicit definitions.

Many standard Boolean DLs such as $\mathcal{ALC}$, $\mathcal{ALCI}$, $\mathcal{ALCQI}$, and their extensions with transitive roles enjoy the CIP and PBDP and sophisticated algorithms for computing interpolants and explicit definitions have been developed [15,9]. Important exceptions are the extensions of any of the above DLs with nominals and/or role hierarchies. In fact, it has recently been shown that the problem of deciding the existence of an interpolant/explicit definition becomes 2ExpTime-complete for DLs such as $\mathcal{ALC}$ with nominals and $\mathcal{ALC}$ with role hierarchies which is in sharp contrast to the ExpTime-completeness of the same problem for $\mathcal{ALC}$ itself inherited from the ExpTime-completeness of subsumption under $\mathcal{ALC}$-ontologies [2].

The aim of this paper is to determine which Horn-DLs enjoy the CIP/PBDP, investigate the complexity of deciding the existence of interpolants/explicit definitions for those that do not enjoy it, and establish bounds on the size of interpolants/explicit definitions. Rather surprisingly, it turns out that most standard Horn-DLs do not enjoy the CIP/PBDP.

Theorem 1. *$\mathcal{EL}$ with nominals, $\mathcal{EL}$ with the universal role, $\mathcal{EL}$ with a single role inclusion of the form $r \circ s \sqsubseteq s$, $\mathcal{EL}$ with role hierarchies and a single transitive role, $\mathcal{ELI}$, Horn-$\mathcal{ALC}$, and Horn-$\mathcal{ALCI}$ do not enjoy the CIP nor PBDP.*

$\mathcal{EL}$ and $\mathcal{EL}$ with role hierarchies enjoy the CIP and PBDP.

The result for $\mathcal{EL}$ with nominals has been proved in [3] already and the positive result for $\mathcal{EL}$ and $\mathcal{EL}$ with role hierarchies in [28,34]. It follows that the behaviour of Horn-DLs is entirely different from Boolean DLs: adding role hierarchies to $\mathcal{ALC}$ does not preserve the CIP/PBDP [27] but it does for $\mathcal{EL}$. On the other hand, extending $\mathcal{ALC}$ with the universal role and/or inverse roles preserves the CIP/PBDP, but it does not for $\mathcal{EL}$.

We note that in Theorem 1, the CIP and PBDP for Horn-$\mathcal{L}$ with $\mathcal{L} \in \{\mathcal{ALC}, \mathcal{ALCI}\}$ can be naturally defined in two different ways, depending on the language for interpolants/explicit definitions. If one is interested in positive existential interpolants/definitions, then one can choose $\mathcal{EL}(\mathcal{I})$-concepts, and if more expressive power is desired, one can choose general Horn-$\mathcal{L}$-concepts. We show that in fact in some cases only Horn-$\mathcal{L}$-interpolants/explicit definitions exist, but that the CIP/PBDP fails for both languages.

For $\mathcal{EL}$-ontologies with unrestricted role inclusions (RIs) $r_1 \circ \cdots \circ r_n \sqsubseteq r$ we also prove a rather general sufficient condition for the existence of explicit definitions: if $\mathcal{O}$ is an $\mathcal{EL}$-ontology with RIs such that each RI either only uses roles in Σ or no role in Σ, then any implicitly Σ-definable concept name under $\mathcal{O}$ is also explicitly Σ-definable under $\mathcal{O}$. A similar result can be shown for interpolants.

We next determine the complexity of deciding the existence of interpolants and explicit definitions and tight bounds on their size for DLs in the $\mathcal{EL}$-family of DLs.

Theorem 2. *For any DL extending $\mathcal{EL}$ with any combination of nominals, RIs, the universal role, or $\bot$, the existence of interpolants and explicit definitions is in* PTime. *If an interpolant/explicit definition exists, then there exists one of at most exponential size. This bound is optimal.*

The proof is based on a careful analysis of canonical models and the introduction and analysis of derivation trees (first used in [11] for $\mathcal{ELI}$) to estimate the size of interpolants. For DLs extending $\mathcal{ELI}$ we use tree automata and again an analysis of derivation trees to prove the following.

Theorem 3. *For $\mathcal{ELI}$ and its extension with nominals and/or the universal role deciding the existence of interpolants and explicit definitions is* ExpTime-*complete. For $\mathcal{ELI}$ with and without the universal role, if an interpolant/explicit definition exists, then there exists one of at most double exponential size. This bound is optimal.*

For $\mathcal{ELI}$ with nominals it remains open whether an interpolant/explicit definition of at most double exponential size exists, if one exists at all. Theorem 3 also holds for Horn-$\mathcal{ALC}$ and Horn-$\mathcal{ALCI}$ provided one asks for interpolants and explicit definitions in $\mathcal{ELI}$. If one admits more expressive Horn-concepts as interpolants or explicit definitions, then the decidability and complexity of existence remains open and could be attacked using the games introduced in [25]. We finally note that for all DLs considered in this paper one always finds an interpolant in the Horn fragment of first-order logic which is known to enjoy the CIP and PBDP.

2 Related Work

The CIP and PBDP have been investigated extensively. They have found applications in formal verification [40], theory combinations [17,18], and in database theory for query rewriting under views [39] and query reformulation and compilation [45,10]. Of particular relevance for this work is the investigation of interpolation and definability in modal logic in general [38] and in hybrid modal logic in particular [1,13]. Also related is work on interpolation in guarded logics [20,19,7,9,8].

The main aim of this paper is to investigate explicit definability of concept names under Horn-DL ontologies. We have therefore chosen a definition of Craig interpolation and interpolants that generalizes the projective Beth definability property and explicit definability in a natural and useful way, following [15]. There are, however, other notions of Craig interpolation that are of interest and have been investigated. Of particular importance for modularity and various other purposes is the following version: if $\mathcal{O}$ is an ontology and $C \sqsubseteq D$ an inclusion such that $\mathcal{O} \models C \sqsubseteq D$, then there exists an ontology $\mathcal{O}'$ in the shared signature of $\mathcal{O}$ and $C \sqsubseteq D$ such that $\mathcal{O}' \models C \sqsubseteq D$. This property has been considered for $\mathcal{EL}$ and various extensions in [44,28]. Currently, it is unknown whether any general results can be proved about the relationship between this version of the Craig interpolation property, the version considered in this article, and/or the projective Beth definability property.

Craig interpolation should not be confused with work on uniform interpolation, both in description logic [32,36,41,29] and in modal logic [47,30,22]. Uniform interpolants generalize Craig interpolants in the sense that a uniform interpolant is an interpolant for a fixed antecedent and any formula implied by the antecedent and sharing with it a fixed set of symbols.

Interpolant and explicit definition existence have only recently been investigated for logics that do not enjoy the CIP or PBDP. In description logic, the work of [2] on Boolean DLs with nominals and role hierarchies was discussed in the introduction. Explicit definition existence is also investigated in [3] for the case in which one is interested in the explicit definition of nominals, see also [12]. In [25], interpolant and explicit definition existence are investigated for the guarded and two-variable fragments, proving that the problems becomes harder than validity. The interpolant existence problem for linear temporal logic LTL is considered in [42].

3 Preliminaries

Let $\mathsf{N_C}$, $\mathsf{N_R}$, and $\mathsf{N_I}$ be disjoint and countably infinite sets of *concept*, *role*, and *individual names*. A *role* is a role name r or an *inverse role* r^-, with r a role name. *Nominals* take the form $\{a\}$, where a is an individual name. The *universal role* is denoted by u. $\mathcal{ELIO}_u$*-concepts* C are defined by the following syntax rule:

$$C, C' \quad ::= \quad \top \mid A \mid \{a\} \mid C \sqcap C' \mid \exists r.C$$

where A ranges over concept names, a over individual names, and r over roles (including the universal role). Fragments of $\mathcal{ELIO}_u$ are defined as usual. For example, $\mathcal{ELI}$*-concepts* are $\mathcal{ELIO}_u$-concepts without nominals and the universal role, and $\mathcal{EL}$*-concepts* are $\mathcal{ELI}$-concepts without inverse roles.

Given any of the DLs $\mathcal{L}$ introduced above, an $\mathcal{L}$*-concept inclusion ($\mathcal{L}$-CI)* takes the form $C \sqsubseteq D$ with C, D $\mathcal{L}$-concepts. An $\mathcal{L}$*-ontology* $\mathcal{O}$ is a finite set of $\mathcal{L}$-CIs.

We also consider ontologies with *role inclusions (RIs)*, expressions of the form $r_1 \circ \cdots \circ r_n \sqsubseteq r$ with $r_1, \ldots, r_n, r$ roles. An $\mathcal{EL}_u^{++}$*-ontology* is the union of an $\mathcal{ELO}_u$-ontology and a finite set of RIs that use roles names only. An $\mathcal{EL}^{++}$*-ontology* is an $\mathcal{EL}_u^{++}$-ontology without the universal role.[1] When dealing with $\mathcal{EL}_u^{++}$-ontologies it will be convenient to use the same name for the underpinning $\mathcal{ELO}_u$-concepts. Thus, $\mathcal{ELO}_u$-concepts will also be called $\mathcal{EL}_u^{++}$-concepts. The same applies to $\mathcal{EL}^{++}$-ontologies and $\mathcal{ELO}$-concepts.

We will also be considering the following more expressive Horn DLs [21], presented here in the form introduced in [37]. *Horn-*$\mathcal{ALCIO}_u$*-CIs* take the form $\neg L \sqcup R$, where L and R are concepts defined by the syntax rules

$$\begin{aligned} R, R' &::= \top \mid \bot \mid A \mid \neg A \mid \{a\} \mid \neg\{a\} \mid R \sqcap R' \mid \neg L \sqcup R \mid \exists r.R \mid \forall r.R \\ L, L' &::= \top \mid \bot \mid A \mid L \sqcap L' \mid L \sqcup L' \mid \exists r.L \end{aligned}$$

with A ranging over concept names, a over individual names, and r over roles (including the universal role). As usual, the fragment of Horn-$\mathcal{ALCIO}_u$ without nominals and the universal role is denoted by Horn-$\mathcal{ALCI}$ and Horn-$\mathcal{ALC}$ denotes the fragment of Horn-$\mathcal{ALCI}$ without inverse roles.

[1] Thus, we abuse notation slightly as the original $\mathcal{EL}^{++}$ also admits concrete domains [5] and $\bot$ (for false). We leave the addition of concrete domains for future work but show that the addition of $\bot$ does not affect our results.

Finally, we remind the reader that the Horn fragment, *Horn-FO*, of first-order logic is defined as the closure of formulas of the form

$$R(\boldsymbol{t}), \quad R_1(\boldsymbol{t}_1) \wedge \cdots \wedge R_n(\boldsymbol{t}_n) \rightarrow R(\boldsymbol{t}), \quad R_1(\boldsymbol{t}_1) \wedge \cdots \wedge R_n(\boldsymbol{t}_n) \rightarrow \bot$$

under conjunction, universal quantification, and existential quantification, where $\boldsymbol{t}_1, \ldots, \boldsymbol{t}_n, \boldsymbol{t}$ are sequences of individual variables and individual names [16].

A *signature* Σ is a set of concept, role, and individual names, uniformly referred to as *symbols*. We use $\text{sig}(X)$ to denote the set of symbols used in any syntactic object X such as a concept or an ontology. If $\mathcal{L}$ is a DL and Σ a signature, then an $\mathcal{L}(\Sigma)$*-concept* C is an $\mathcal{L}$-concept with $\text{sig}(C) \subseteq \Sigma$. The *size* $||X||$ of a syntactic object X is the number of symbols needed to write it down.

The semantics of DLs is given in terms of *interpretations* $\mathcal{I} = (\Delta^{\mathcal{I}}, \cdot^{\mathcal{I}})$, where $\Delta^{\mathcal{I}}$ is a non-empty set (the *domain*) and $\cdot^{\mathcal{I}}$ is the *interpretation function*, assigning to each $A \in \mathsf{N_C}$ a set $A^{\mathcal{I}} \subseteq \Delta^{\mathcal{I}}$, to each $r \in \mathsf{N_R}$ a relation $r^{\mathcal{I}} \subseteq \Delta^{\mathcal{I}} \times \Delta^{\mathcal{I}}$, and to each $a \in \mathsf{N_I}$ an element $a^{\mathcal{I}} \in \Delta^{\mathcal{I}}$. The interpretation $C^{\mathcal{I}} \subseteq \Delta^{\mathcal{I}}$ of a concept C in $\mathcal{I}$ is defined as usual, see [6]. An interpretation $\mathcal{I}$ *satisfies* a CI $C \sqsubseteq D$ if $C^{\mathcal{I}} \subseteq D^{\mathcal{I}}$ and an RI $r_1 \circ \cdots \circ r_n \sqsubseteq r$ if $r_1^{\mathcal{I}} \circ \cdots \circ r_n^{\mathcal{I}} \subseteq r^{\mathcal{I}}$.

We say that $\mathcal{I}$ is a *model of* an ontology if it satisfies all inclusions in it. If α is a CI or RI, we write $\mathcal{O} \models \alpha$ if all models of $\mathcal{O}$ satisfy α. We write $\mathcal{O} \models C \equiv D$ if $\mathcal{O} \models C \sqsubseteq D$ and $\mathcal{O} \models D \sqsubseteq C$.

An ontology is in *normal form* if its CIs are of the form

$$\top \sqsubseteq A, \quad A_1 \sqcap A_2 \sqsubseteq B, \quad A \sqsubseteq \{a\}, \quad \{a\} \sqsubseteq A,$$

and

$$A \sqsubseteq \exists r.B, \quad \exists r.B \sqsubseteq A$$

where A, A_1, A_2, B are concept names, r is a role or the universal role, and a is an individual name. We call an ontology $\mathcal{O}'$ a *model conservative extension* of an ontology $\mathcal{O}$ if $\mathcal{O}' \models \mathcal{O}$ and every model of $\mathcal{O}$ can be expanded to a model of $\mathcal{O}'$ by modifying the interpretation of symbols in $\text{sig}(\mathcal{O}') \setminus \text{sig}(\mathcal{O})$.

Lemma 4. *Let $\mathcal{L}$ be a DL from $\mathcal{EL}, \mathcal{ELI}, \mathcal{ELO}, \mathcal{ELIO}, \mathcal{EL}^{++}$, or an extension with the universal role, and let $\mathcal{O}$ be an $\mathcal{L}$-ontology. Then one can construct in polynomial time an $\mathcal{L}$-ontology $\mathcal{O}'$ in normal form such that $\mathcal{O}'$ is a model conservative extension of $\mathcal{O}$.*

We next introduce ABoxes as a technical tool that allows us to move from interpretations to (potentially incomplete) sets of facts and concepts. An *ABox* $\mathcal{A}$ is a (possibly infinite) set of assertions of the form $A(x)$, $r(x, y)$, $\{a\}(x)$, and $\top(x)$ with $A \in \mathsf{N_C}$, $r \in \mathsf{N_R}$, $a \in \mathsf{N_I}$, and x, y individual variables. We denote by $\text{ind}(\mathcal{A})$ the set of individual variables in $\mathcal{A}$. Note that we use individual variables instead of individual names in ABoxes as it will be convenient to distinguish the individual names used in nominals of an ontology from the individuals used in ABoxes. An ABox is *factorized* if $\{a\}(x), \{a\}(y) \in \mathcal{A}$ imply $x = y$. If Σ is a signature, then a Σ-ABox is an ABox that only uses symbols from Σ.

ABox assertions are interpreted in an interpretation $\mathcal{I}$ using a *variable assignment* v that maps individual variables to elements of $\Delta^{\mathcal{I}}$. Then $\mathcal{I}, v$ satisfies an assertion $A(x)$ if $v(x) \in A^{\mathcal{I}}$, $r(x,y)$ if $(v(x),v(y)) \in r^{\mathcal{I}}$, $\{a\}(x)$ if $a^{\mathcal{I}} = v(x)$, and $\top(x)$ is always satisfied. $\mathcal{I}, v$ satisfies an ABox if it satisfies all assertions in it. We write $\mathcal{I} \models \mathcal{A}[x \mapsto d]$ if there exists an assignment v with $v(x) = d$ such that $\mathcal{I}, v$ satisfies $\mathcal{A}$. We say that an assertion $A_0(x_0)$ *is entailed by an ontology* $\mathcal{O}$ *and ABox* $\mathcal{A}$, in symbols $\mathcal{O}, \mathcal{A} \models A_0(x_0)$, if $v(x) \in A_0^{\mathcal{I}}$ for all models $\mathcal{I}$ of $\mathcal{O}$ and assignments v such that $\mathcal{I}, v$ satisfy $\mathcal{A}$. This is the standard notion of entailment from a knowledge base consisting of an ontology and an ABox. Deciding entailment is in PTime for the DLs between $\mathcal{EL}$ and $\mathcal{EL}_u^{++}$ [5] and ExpTime-complete for the DLs between $\mathcal{ELI}$ and $\mathcal{ELIO}_u$ [6].

Every interpretation $\mathcal{I}$ defines a factorized ABox $\mathcal{A}_{\mathcal{I}}$ by identifying every $d \in \Delta^{\mathcal{I}}$ with a variable x_d and taking $A(x_d)$ if $d \in A^{\mathcal{I}}$, $r(x_c, x_d)$ if $(c,d) \in r^{\mathcal{I}}$, $\{a\}(x_d)$ if $a^{\mathcal{I}} = d$. Conversely, factorized ABoxes define interpretations in the obvious way.

3.1 Preliminaries for $\mathcal{EL}$

We associate with every ABox $\mathcal{A}$ a directed graph

$$G_{\mathcal{A}} = (\mathrm{ind}(\mathcal{A}), \bigcup_{r \in \mathsf{N_R}} \{(x,y) \mid r(x,y) \in \mathcal{A}\}).$$

Let $\mathcal{A}$ be a factorized ABox. We say that $\mathcal{A}$ is *ditree-shaped* if $G_{\mathcal{A}}$ is acyclic and $r(x,y) \in \mathcal{A}$ and $s(x,y) \in \mathcal{A}$ imply $r = s$. Let Γ be a set of individual names. Then $\mathcal{A}$ is *ditree-shaped modulo* Γ if after dropping some facts of the form $r(x,y)$ with $\{a\}(y) \in \mathcal{A}$ for some $a \in \Gamma$, it is ditree-shaped. A *pointed ABox* is a pair $\mathcal{A}, x$ with $x \in \mathrm{ind}(\mathcal{A})$. We next observe that $\mathcal{ELO}_u(\Sigma)$-concepts correspond to pointed Σ-ABoxes $\mathcal{A}, x$ such that $\mathcal{A}$ is ditree-shaped modulo $\mathsf{N_I} \cap \Sigma$. $\mathcal{ELO}(\Sigma)$-concepts correspond to *rooted* pointed Σ-ABoxes $\mathcal{A}, x$ such that $\mathcal{A}$ is tree-shaped modulo $\mathsf{N_I} \cap \Sigma$, where $\mathcal{A}, x$ is called rooted if for every $y \in \mathrm{ind}(\mathcal{A})$ there is a path from x to y in $G_{\mathcal{A}}$. The following lemma is shown in the obvious way.

Lemma 5. *For any $\mathcal{ELO}_u(\Sigma)$-concept C one can construct in polynomial time a pointed Σ-ABox $\mathcal{A}, x$ such that $\mathcal{A}$ is ditree-shaped modulo $\mathsf{N_I} \cap \Sigma$ and $d \in C^{\mathcal{I}}$ iff $\mathcal{I} \models \mathcal{A}[x \mapsto d]$, for all interpretations $\mathcal{I}$ and $d \in \Delta^{\mathcal{I}}$.*

Conversely, for any pointed Σ-ABox $\mathcal{A}, x$ such that $\mathcal{A}$ is a ditree-shaped ABox modulo Γ, one can construct in polynomial time an $\mathcal{ELO}_u(\Sigma)$-concept C such that $\Gamma = \mathsf{N_I} \cap \Sigma$ and $d \in C^{\mathcal{I}}$ iff $\mathcal{I} \models \mathcal{A}_C[x \mapsto d]$, for all interpretations $\mathcal{I}$ and $d \in \Delta^{\mathcal{I}}$.

The above also holds if one replaces $\mathcal{ELO}_u(\Sigma)$-concepts by $\mathcal{ELO}(\Sigma)$-concepts and requires the pointed ABoxes to be rooted.

We define a *canonical model* $\mathcal{I}_{\mathcal{O},A_0}$ for an $\mathcal{EL}_u^{++}$-ontology $\mathcal{O}$ in normal form and a concept name A_0 and, more generally, the canonical model $\mathcal{I}_{\mathcal{O},\mathcal{A}}$ of an $\mathcal{EL}_u^{++}$-ontology $\mathcal{O}$ in normal form and an ABox $\mathcal{A}$. They are essentially the same models as constructed in [5]. As we do not use this model do check subsumption, we give a succinct model-theoretic construction.

Assume $\mathcal{O}$ and A_0 are given and $\mathcal{O}$ is in normal form. Define an equivalence relation $\sim$ on the set of individual names a in $\mathrm{sig}(\mathcal{O})$ by setting $a \sim b$ if $\mathcal{O} \models$

$\exists u.A_0 \sqcap \{a\} \sqsubseteq \{b\}$. Let $[a] = \{b \in \text{sig}(\mathcal{O}) \mid a \sim b\}$ and set $\Delta_I = \{[a] \mid a \in \text{sig}(\mathcal{O})\}$. Say that a concept name A is *absorbed by* an individual name a if $\mathcal{O} \models \exists u.A_0 \sqcap A \sqsubseteq \{a\}$ and let Δ_C denote the set of concept names A in $\mathcal{O}$ such that $\mathcal{O} \models A_0 \sqsubseteq \exists u.A$ and A is not absorbed by any individual name.

Now let $\Delta^{\mathcal{I}_{\mathcal{O},A_0}} = \Delta_I \cup \Delta_C$ and let

$$\begin{aligned}
A^{\mathcal{I}_{\mathcal{O},A_0}} &= \{[a] \in \Delta^{\mathcal{I}_{\mathcal{O},A_0}} \mid \mathcal{O} \models \exists u.A_0 \sqcap \{a\} \sqsubseteq A\} \cup \\
&\quad\ \{B \in \Delta^{\mathcal{I}_{\mathcal{O},A_0}} \mid \mathcal{O} \models \exists u.A_0 \sqcap B \sqsubseteq A\} \\
a^{\mathcal{I}_{\mathcal{O},A_0}} &= [a] \\
r^{\mathcal{I}_{\mathcal{O},A_0}} &= \{([a],[b]) \in \Delta^{\mathcal{I}_{\mathcal{O},A_0}} \times \Delta^{\mathcal{I}_{\mathcal{O},A_0}} \mid \mathcal{O} \models \exists u.A_0 \sqcap \{a\} \sqsubseteq \exists r.\{b\}\} \cup \\
&\quad\ \{([a],B) \in \Delta^{\mathcal{I}_{\mathcal{O},A_0}} \times \Delta^{\mathcal{I}_{\mathcal{O},A_0}} \mid \mathcal{O} \models \exists u.A_0 \sqcap \{a\} \sqsubseteq \exists r.B\} \cup \\
&\quad\ \{(B,[a]) \in \Delta^{\mathcal{I}_{\mathcal{O},A_0}} \times \Delta^{\mathcal{I}_{\mathcal{O},A_0}} \mid \mathcal{O} \models \exists u.A_0 \sqcap B \sqsubseteq \exists r.\{a\}\} \cup \\
&\quad\ \{(A,B) \in \Delta^{\mathcal{I}_{\mathcal{O},A_0}} \times \Delta^{\mathcal{I}_{\mathcal{O},A_0}} \mid \mathcal{O} \models \exists u.A_0 \sqcap A \sqsubseteq \exists r.B\}
\end{aligned}$$

for every concept name $A \in \mathsf{N_C}$, $a \in \text{sig}(\mathcal{O}) \cap \mathsf{N_I}$, and $r \in \mathsf{N_R}$. We often denote the nodes $[a]$ and A by $\rho_{[a]}$ or, for simplicity, ρ_a and, respectively, ρ_A. If A_0 is absorbed by an individual a we still often denote $\rho_{[a]}$ by ρ_{A_0}.

We next define *Σ-simulations* for a signature Σ. Let $\mathcal{I}$ and $\mathcal{J}$ be interpretations. A relation S is called a Σ-simulation between $\mathcal{I}$ and $\mathcal{J}$ if the following conditions hold:

1. if $d \in A^{\mathcal{I}}$ and $(d,e) \in S$, then $e \in A^{\mathcal{J}}$, for all $A \in \Sigma$;
2. if $d = a^{\mathcal{I}}$ and $(d,e) \in S$, then $e = a^{\mathcal{J}}$, for all $a \in \Sigma$;
3. if $(d,d') \in r^{\mathcal{I}}$ and $(d,e) \in S$, then there exists e' with $(e,e') \in r^{\mathcal{J}}$ and $(d',e') \in S$, for all $r \in \Sigma$.

S is *global* if its domain is $\Delta^{\mathcal{I}}$. If $(\mathcal{I},d)$ and $(\mathcal{J},e)$ are pointed interpretations and S a Σ-simulation between $\mathcal{I}$ and $\mathcal{J}$ with $(d,e) \in S$, then we write $(\mathcal{I},d) \preceq_{\mathcal{ELO},\Sigma} (\mathcal{J},e)$ and say that $(\mathcal{I},d)$ is *$\mathcal{ELO}(\Sigma)$-simulated by* $(\mathcal{J},e)$. If S is global, then we write $(\mathcal{I},d) \preceq_{\mathcal{ELO}_u,\Sigma} (\mathcal{J},e)$ and say that $(\mathcal{I},d)$ is *$\mathcal{ELO}_u(\Sigma)$-simulated by* $(\mathcal{J},e)$.

Let $\mathcal{L} \in \{\mathcal{ELO},\mathcal{ELO}_u\}$. We write $(\mathcal{I},d) \leq_{\mathcal{L},\Sigma} (\mathcal{J},e)$ if $d \in C^{\mathcal{I}}$ implies $e \in C^{\mathcal{J}}$ for all $\mathcal{L}(\Sigma)$-concepts C. The following characterization is well known [35].

Lemma 6. *Let $\mathcal{L} \in \{\mathcal{ELO},\mathcal{ELO}_u\}$. Then*

$$(\mathcal{I},d) \preceq_{\mathcal{L},\Sigma} (\mathcal{J},e) \quad \Rightarrow \quad (\mathcal{I},d) \leq_{\mathcal{L},\Sigma} (\mathcal{J},e)$$

The converse direction holds if $\mathcal{J}$ is ω-saturated (in particular, finite).

Lemma 7. *The canonical model $\mathcal{I}_{\mathcal{O},A_0}$ is a model of $\mathcal{O}$ and for every model $\mathcal{J}$ of $\mathcal{O}$ and any $d \in \Delta^{\mathcal{J}}$ with $d \in A_0^{\mathcal{J}}$, $(\mathcal{I}_{\mathcal{O},A_0},\rho_{A_0}) \preceq_{\mathcal{L},\Sigma} (\mathcal{J},d)$, where Σ is any signature.*

Proof. We first show that $\mathcal{I}_{\mathcal{O},A_0}$ is a model of $\mathcal{O}$. It is straightforward to show that $\mathcal{I}_{\mathcal{O},A_0}$ satisfies the CIs of the form $\top \sqsubseteq A$, $A_1 \sqcap A_2 \sqsubseteq A$, $A \sqsubseteq \{a\}$, $\{a\} \sqsubseteq A$.

Assume now that $A \sqsubseteq \exists r.B \in \mathcal{O}$ and $\rho_C \in A^{\mathcal{I}_{\mathcal{O},A_0}}$ with C of the form a or A. We have $\mathcal{O} \models \exists u.C$, $\mathcal{O} \models \exists u.A_0 \sqcap C \sqsubseteq A$. Thus $\mathcal{O} \models \exists u.A_0 \sqcap C \sqsubseteq \exists r.B$. But then $(\rho_C,\rho_B) \in r^{\mathcal{I}_{\mathcal{O},A_0}}$ and $\rho_B \in B^{\mathcal{I}_{\mathcal{O},A_0}}$. Thus $\rho_C \in (\exists r,B)^{\mathcal{I}_{\mathcal{O},A_0}}$, as required.

Assume now that $\exists r.A \sqsubseteq B \in \mathcal{O}$ and $\rho_C \in (\exists r.A)^{\mathcal{I}_{\mathcal{O},A_0}}$. Then there exists ρ_D such that $(\rho_C, \rho_D) \in r^{\mathcal{I}_{\mathcal{O},A_0}}$ and $\rho_D \in A^{\mathcal{I}_{\mathcal{O},A_0}}$. Hence $\mathcal{O} \models \exists u.A_0 \sqcap C \sqsubseteq \exists r.D$ and $\mathcal{O} \models \exists u.A_0 \sqcap D \sqsubseteq A$. Thus, $\mathcal{O} \models \exists u.A_0 \sqcap C \sqsubseteq \exists r.A$. Hence since $\exists r.A \sqsubseteq B \in \mathcal{O}$, $\mathcal{O} \models \exists u.A_0 \sqcap C \sqsubseteq B$. But then $\rho_C \in B^{\mathcal{I}_{\mathcal{O},A_0}}$, as required.

Finally, assume that $r_1 \circ \cdots \circ r_n \sqsubseteq r \in \mathcal{O}$ and $(\rho_C, \rho_D) \in r_1^{\mathcal{I}_{\mathcal{O},A_0}} \circ \cdots \circ r_n^{\mathcal{I}_{\mathcal{O},A_0}}$. Then there are $\rho_{C_0}, \ldots, \rho_{C_n}$ with $(\rho_{C_i}, \rho_{C_{i+1}}) \in r_{i+1}^{\mathcal{I}_{\mathcal{O},A_0}}$ for all $i < n$, where $C_0 = C$ and $C_n = D$. We obtain $\mathcal{O} \models \exists u.A_0 \sqcap C_i \sqsubseteq \exists r_{i+1}.C_{i+1}$ for all $i < n$. Thus $\mathcal{O} \models \exists u.A_0 \sqcap C \sqsubseteq \exists r_1 \cdots \exists r_n.D$. Hence $\mathcal{O} \models \exists u.A_0 \sqcap C \sqsubseteq \exists r.D$. Hence $(\rho_C, \rho_D) \in r^{\mathcal{I}_{\mathcal{O},A_0}}$, as required.

Let $\mathcal{J}$ be a model of $\mathcal{O}$ with $A_0^{\mathcal{J}} \neq \emptyset$. Define a relation between $\Delta^{\mathcal{I}_{\mathcal{O},A_0}}$ and $\Delta^{\mathcal{J}}$ as follows: for any $\rho_C \in \Delta^{\mathcal{I}_{\mathcal{O},A_0}}$ and $d \in \Delta^{\mathcal{J}}$, let $(\rho_C, d) \in S$ if $d \in C^{\mathcal{J}}$. It is easy to see that this well-defined and that for any ρ_C there exists a $d \in \Delta^{\mathcal{J}}$ with $(\rho_C, d) \in S$. It is straightforward to show that S is a $\mathcal{ELO}_u(\Sigma)$-simulation, as required. $\square$

The following observation is a consequence of Lemma 6 and Lemma 7.

Lemma 8. *Let $\mathcal{O}$ be an $\mathcal{EL}_u^{++}$-ontology in normal form, A_0 a concept name, and C an $\mathcal{ELO}_u$-concept. Then the following conditions are equivalent:*

1. $\rho_{A_0} \in C^{\mathcal{I}_{\mathcal{O},A_0}}$;
2. $\mathcal{O} \models A_0 \sqsubseteq C$.

Next assume that $\mathcal{O}$ and an ABox $\mathcal{A}$ are given. Assume $\mathcal{O}$ is in normal form. Then one can construct in polynomial time a canonical model $\mathcal{I}_{\mathcal{O},\mathcal{A}}$ of $\mathcal{O}$ that satisfies $\mathcal{A}$ via an assignment $v_{\mathcal{O},\mathcal{A}}$. The details are straightforward, and we only give the main properties of $\mathcal{I}_{\mathcal{O},\mathcal{A}}$.

Lemma 9. *Given an $\mathcal{EL}_u^{++}$-ontology $\mathcal{O}$ in normal form and an ABox $\mathcal{A}$ one can construct in polynomial time a model $\mathcal{I}_{\mathcal{O},\mathcal{A}}$ of $\mathcal{O}$ and an assignment $v_{\mathcal{O},\mathcal{A}}$ such that for all $x \in ind(\mathcal{A})$ and all $\mathcal{ELO}_u$-concepts C the following conditions are equivalent:*

1. $v_{\mathcal{O},\mathcal{A}}(x) \in C^{\mathcal{I}_{\mathcal{O},\mathcal{A}}}$;
2. $\mathcal{O}, \mathcal{A} \models C(x)$.

3.2 Preliminaries for $\mathcal{ELI}$

For extensions of $\mathcal{ELI}$ we employ *tree-shaped ABoxes*. We associate with every ABox $\mathcal{A}$ the undirected graph

$$G^u_{\mathcal{A}} = (\text{ind}(\mathcal{A}), \bigcup_{r \in \mathsf{N_R}} \{\{x, y\} \mid r(x, y) \in \mathcal{A}\}).$$

Let $\mathcal{A}$ be factorized. We say that $\mathcal{A}$ is *tree-shaped* if $G^u_{\mathcal{A}}$ is acyclic, $r(x,y) \in \mathcal{A}$ and $s(x,y) \in \mathcal{A}$ imply $r = s$, and $r(x,y) \in \mathcal{A}$ implies $s(y,x) \notin \mathcal{A}$ for any s. $\mathcal{A}$ is *tree-shaped modulo a set Γ of individual names* if after dropping some facts $r(x,y)$ with $\{a\}(x)$ or $\{a\}(y) \in \mathcal{A}$ for some $a \in \Gamma$ it is tree-shaped. We next observe that $\mathcal{ELIO}_u(\Sigma)$-concepts correspond to pointed Σ-ABoxes $\mathcal{A}, x$ such that $\mathcal{A}$ is tree-shaped modulo $\mathsf{N_I} \cap \Sigma$. $\mathcal{ELIO}(\Sigma)$-concepts correspond to *weakly rooted* pointed Σ-ABoxes $\mathcal{A}, x$ such that $\mathcal{A}$ is tree-shaped modulo $\mathsf{N_I} \cap \Sigma$, where $\mathcal{A}, x$ is called weakly rooted if for every $y \in \mathrm{ind}(\mathcal{A})$ there is a path from x to y in $G^u_{\mathcal{A}}$. The following lemma is shown in the obvious way.

Lemma 10. *For any $\mathcal{ELIO}_u(\Sigma)$-concept C one can construct in polynomial time a pointed Σ-ABox $\mathcal{A}, x$ such that $\mathcal{A}$ is tree-shaped modulo $\mathsf{N_I} \cap \Sigma$ and $d \in C^{\mathcal{I}}$ iff $\mathcal{I} \models \mathcal{A}[x \mapsto d]$, for all interpretations $\mathcal{I}$ and $d \in \Delta^{\mathcal{I}}$.*

Conversely, for any pointed Σ-ABox $\mathcal{A}, x$ such that $\mathcal{A}$ is a tree-shaped ABox modulo Γ, one can construct in polynomial time an $\mathcal{ELIO}_u(\Sigma)$-concept C such that $\Gamma = \mathsf{N_I} \cap \Sigma$ and $d \in C^{\mathcal{I}}$ iff $\mathcal{I} \models \mathcal{A}_C[x \mapsto d]$, for all interpretations $\mathcal{I}$ and $d \in \Delta^{\mathcal{I}}$.

The above also holds if one replaces $\mathcal{ELIO}_u(\Sigma)$-concepts by $\mathcal{ELIO}(\Sigma)$-concepts and requires the pointed ABoxes to be weakly rooted.

4 Craig Interpolation Property and Projective Beth Definability Property

We introduce the Craig interpolation property (CIP) and the projective Beth definability property (PBDP) as defined in [15] and determine which DLs considered in this article enjoy the two properties. We observe that the CIP implies the PBDP, but lack a proof of the converse direction. Nevertheless, all DLs considered in this paper enjoying the PBDP also enjoy the CIP. In fact, we generally prove that a DL enjoys both properties by proving the CIP and disprove both properties by disproving the PBDP.

We set $\mathrm{sig}(\mathcal{O}, C) = \mathrm{sig}(\mathcal{O}) \cup \mathrm{sig}(C)$, for any ontology $\mathcal{O}$ and concept C. Let $\mathcal{O}_1, \mathcal{O}_2$ be $\mathcal{L}$-ontologies and let C_1, C_2 be $\mathcal{L}$-concepts. Then an $\mathcal{L}$-concept D is called an *$\mathcal{L}$-interpolant* for $C_1 \sqsubseteq C_2$ under $\mathcal{O}_1, \mathcal{O}_2$ if

- $\mathrm{sig}(D) \subseteq \mathrm{sig}(\mathcal{O}_1, C_1) \cap \mathrm{sig}(\mathcal{O}_2, C_2)$;
- $\mathcal{O}_1 \cup \mathcal{O}_2 \models C_1 \sqsubseteq D$;
- $\mathcal{O}_1 \cup \mathcal{O}_2 \models D \sqsubseteq C_2$.

Definition 11. *A DL $\mathcal{L}$ has the Craig interpolation property (CIP) if for any $\mathcal{L}$-ontologies $\mathcal{O}_1, \mathcal{O}_2$ and $\mathcal{L}$-concepts C_1, C_2 such that $\mathcal{O}_1 \cup \mathcal{O}_2 \models C_1 \sqsubseteq C_2$ there exists an $\mathcal{L}$-interpolant for $C_1 \sqsubseteq C_2$ under $\mathcal{O}_1, \mathcal{O}_2$.*

We next define the relevant definability notions. Let $\mathcal{O}$ be an ontology and A a concept name. Let $\Sigma \subseteq \mathrm{sig}(\mathcal{O})$ be a signature. An $\mathcal{L}(\Sigma)$-concept C is an *explicit $\mathcal{L}(\Sigma)$-definition of A under $\mathcal{O}$* if $\mathcal{O} \models A \equiv C$. We call A *explicitly definable in $\mathcal{L}(\Sigma)$ under $\mathcal{O}$* if there is an explicit $\mathcal{L}(\Sigma)$-definition of A under $\mathcal{O}$. The Σ-reduct $\mathcal{I}_{|\Sigma}$ of an interpretation $\mathcal{I}$ coincides with $\mathcal{I}$ except that no non-Σ symbol is interpreted in $\mathcal{I}_{|\Sigma}$. A concept A is called *implicitly definable from Σ under $\mathcal{O}$* if the Σ-reduct of any model $\mathcal{I}$ of $\mathcal{O}$ determines the set $A^{\mathcal{I}}$; in other words, if $\mathcal{I}$ and $\mathcal{J}$ are both models of $\mathcal{O}$ such that $\mathcal{I}_{|\Sigma} = \mathcal{J}_{|\Sigma}$, then $A^{\mathcal{I}} = A^{\mathcal{J}}$. It is easy to see that implicit definability can be reformulated as a standard reasoning problem as follows: a concept name A is implicitly definable from Σ under $\mathcal{O}$ iff $\mathcal{O} \cup \mathcal{O}_\Sigma \models A \equiv A'$, where $\mathcal{O}_\Sigma$ is obtained from $\mathcal{O}$ by replacing every non-Σ symbol X uniformly by a fresh symbol X'. If a concept name is explicitly definable in $\mathcal{L}(\Sigma)$ under $\mathcal{O}$, then it is implicitly definable from Σ under $\mathcal{O}$, for any language $\mathcal{L}$. A logic enjoys the projective Beth definability property if the converse implication holds as well.

Definition 12. *A DL has the projective Beth definable property (PBDP) if for any $\mathcal{L}$-ontology $\mathcal{O}$, concept name A, and signature $\Sigma \subseteq sig(\mathcal{O})$ the following holds: if A is implicitly definable from Σ under $\mathcal{O}$, then A is explicitly $\mathcal{L}(\Sigma)$-definable under $\mathcal{O}$.*

Remark 13. Observe that the CIP implies the PBDP. To see this, assume that an $\mathcal{L}$-ontology $\mathcal{O}$, concept name A and a signature Σ are given, and that A is implicitly definable from Σ under $\mathcal{O}$. Then $\mathcal{O} \cup \mathcal{O}_\Sigma \models A \equiv A'$, with $\mathcal{O}_\Sigma$ defined above. Take an $\mathcal{L}$-interpolant C for $A \sqsubseteq A'$ under $\mathcal{O}, \mathcal{O}_\Sigma$. Then C is an explicit $\mathcal{L}(\Sigma)$-definition of A under $\mathcal{O}$.

We observe that the addition of $\bot$ to any of the DLs considered in this article does not affect the CIP or BDBP.

Remark 14. Assume that $\mathcal{L}$ is any of the DLs from the $\mathcal{EL}$ or $\mathcal{ELI}$ family introduced above and let $\mathcal{L}_\bot$ denote its extension with $\bot$. Then $\mathcal{L}$ enjoys the CIP/PBDP iff $\mathcal{L}_\bot$ does. We show this for the CIP. Assume first that $C \sqsubseteq D$ and $\mathcal{O}_1, \mathcal{O}_2$ are a counterexample to the CIP of $\mathcal{L}$. Then they are also a counterexample to the CIP of $\mathcal{L}_\bot$. Conversely, assume that $C \sqsubseteq D$ and $\mathcal{O}_1, \mathcal{O}_2$ are a counterexample to the CIP of $\mathcal{L}_\bot$. We may assume that no CI in $\mathcal{O}_1 \cup \mathcal{O}_2$ uses $\bot$ in the concept on its left hand side (if it does, the CI is redundant). $\mathcal{O}_1 \cup \mathcal{O}_2 \not\models C \sqsubseteq \bot$ because otherwise $\bot$ would be an $\mathcal{L}_\bot$-interpolant. It follows that C, D do not use $\bot$. But then it is easy to see that by replacing every occurrence of $\bot$ by a fresh concept name A in $\mathcal{O}_1, \mathcal{O}_2$ one obtains a counterexample to the CIP of $\mathcal{L}$.

We now determine which of the DLs introduced in the introduction enjoy the CIP and PBDP. We begin by considering extensions of $\mathcal{EL}$ without inverse roles and then add inverse roles.

The description logic $\mathcal{EL}$ enjoys the CIP/PBDP [44]. We show in this paper that more often than not its extensions with the $\mathcal{EL}^{++}$ constructions do not enjoy the CIP/PBDP. In fact, nominals, the universal role and RIs *on their own* lead to the failure of CIP/PBDP, and explicit definitions do not exist even in $\mathcal{ELO}_u$.

Theorem 15. *The DLs $\mathcal{ELO}$, $\mathcal{EL}_u$ and $\mathcal{EL}^{++}$ do not enjoy the CIP and PBDP. For $\mathcal{EL}^{++}$, even the extensions of an $\mathcal{EL}$ ontology without nominals with (a) a single role inclusion of the form $r \circ s \sqsubseteq s$ or (b) with a single transitivity inclusion of the form $s \circ s \sqsubseteq s$ combined with role hierarchies of the form $r_1 \sqsubseteq r_2$ do not enjoy the CIP/PBDP.*

Proof. We start with showing that $\mathcal{EL}_u$ does not enjoy the PBDP. We define an $\mathcal{EL}_u$-ontology $\mathcal{O}$, signature Σ, and concept name A such that A is implicitly definable in Σ under $\mathcal{O}$ but not $\mathcal{EL}_u(\Sigma)$-explicitly definable under $\mathcal{O}$. Define $\mathcal{O}$ as the following set of CIs.

$$\begin{aligned} A &\sqsubseteq B \\ D \sqcap \exists u.A &\sqsubseteq E \\ B &\sqsubseteq \exists r.C \\ C &\sqsubseteq D \\ B \sqcap \exists r.(C \sqcap E) &\sqsubseteq A\,, \end{aligned}$$

and $\Sigma = \{B, D, E, r\}$. We have $\mathcal{O} \models A \equiv B \sqcap \forall r.(D \rightarrow E)$, so A is implicitly definable in Σ under $\mathcal{O}$. However, the models $\mathcal{I}$ and $\mathcal{I}'$ given in Fig. 1 show that A is not explicitly $\mathcal{EL}_u(\Sigma)$-definable under $\mathcal{O}$. Indeed, $\mathcal{I}$ and $\mathcal{I}'$ are both models

Fig. 1. Interpretations $\mathcal{I}$ and $\mathcal{I}'$ used to show that PBDP fails in $\mathcal{EL}_u$.

of $\mathcal{O}$, $a \in A^{\mathcal{I}}$, $a' \notin A^{\mathcal{I}'}$, and the relation

$$\{(a, a'), (b, b'), (b, b'')\}$$

is a $\mathcal{EL}_u(\Sigma)$-simulation between $\mathcal{I}$ and $\mathcal{I}'$. As $\mathcal{EL}_u$ concepts are preserved under $\mathcal{EL}_u$-simulations, for every $\mathcal{EL}_u(\Sigma)$-concept C we have $a' \in C^{\mathcal{I}'}$ if $a \in C^{\mathcal{I}}$. Hence no $\mathcal{EL}_u$ definition of A exists. As Σ contains no individual names, no $\mathcal{ELO}_u(\Sigma)$-explicit definition exists as well.

We next turn to $\mathcal{EL}$ with role inclusions. Let

$$\mathcal{O} = \{A \sqsubseteq \exists r.E, E \sqsubseteq \exists s.B, \exists s.B \sqsubseteq A, r \circ s \sqsubseteq s\}$$

and $\Sigma = \{s, E\}$. Then A is implicitly definable using Σ under $\mathcal{O}$ since

$$\mathcal{O} \models \forall x(A(x) \leftrightarrow \exists y(E(y) \wedge \forall z(s(y, z) \rightarrow s(x, z))).$$

We show that there does not exist any $\mathcal{EL}_u(\Sigma)$-explicit definition of A under $\mathcal{O}$. Interpretations $\mathcal{I}$ and $\mathcal{I}'$ given in Fig. 2 are both models of $\mathcal{O}$, $a \in A^{\mathcal{I}}$, $a' \notin A^{\mathcal{I}'}$, and the relation

$$S = \{(a, a'), (b, b'), (b, b''), (c, c'), (c, c'')\}$$

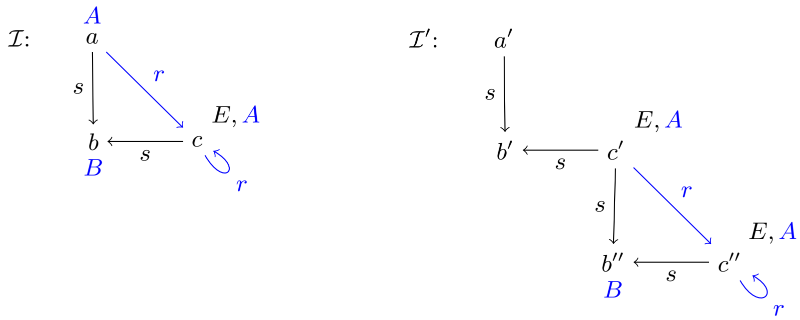


Fig. 2. Interpretations $\mathcal{I}$ and $\mathcal{I}'$ used to show that PBDP fails in $\mathcal{EL}_u^{++}$.

is a $\mathcal{EL}_u(\Sigma)$-simulation between $\mathcal{I}$ and $\mathcal{I}'$, and so no $\mathcal{EL}_u$ definition of A exists. Moreover, it can be seen that S is actually a global bisimulation relation. Global bisimulations preserve $\mathcal{ALC}_u$-concepts [31], so A does not even have an explicit $\mathcal{ALC}_u(\Sigma)$-definition.

A simple modification of the example above demonstrates that combining transitivity and role hierarchies also leads to the loss of CIP/PBDP. Let

$$\mathcal{O} = \{A \sqsubseteq \exists s.E, E \sqsubseteq \exists s_1.B, \exists s_2.B \sqsubseteq A, s_1 \sqsubseteq s, s \sqsubseteq s_2, s \circ s \sqsubseteq s\}$$

and $\Sigma = \{s_1, s_2, E\}$. Then A is implicitly definable using Σ under $\mathcal{O}$ since

$$\mathcal{O} \models \forall x(A(x) \leftrightarrow \exists y(E(y) \land \forall z(s_1(y, z) \rightarrow s_2(x, z)))).$$

Interpretations $\mathcal{I}$ and $\mathcal{I}'$ showing that A has neither $\mathcal{EL}_u(\Sigma)$- nor $\mathcal{ALC}_u(\Sigma)$-definition are given in Fig. 3.

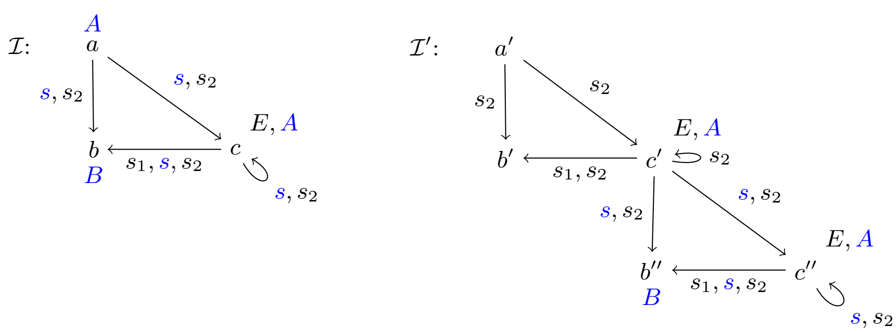


Fig. 3. Interpretations $\mathcal{I}$ and $\mathcal{I}'$ used to show that PBDP fails in $\mathcal{EL}$ extended with transitive roles and role hierarchies.

Finally, we consider nominals. We note that an example given in [3] shows that $\mathcal{ELO}$ does not enjoy the CIP/PBDP. We here give an example of an $\mathcal{ELO}$-ontology, signature Σ, and concept name A such that A is implicitly definable in Σ under $\mathcal{O}$ but not explicitly $\mathcal{ELO}_u(\Sigma)$-definable under $\mathcal{O}$. Let $\mathcal{O}$ contain the following CIs:

$$A \sqsubseteq \exists r.(E \sqcap \{c\}), \quad \top \sqsubseteq \exists s.(Q_2 \sqcap \exists s.\{c\})$$

$$\exists s.(Q_1 \sqcap Q_2 \sqcap \exists s.\{c\}) \sqsubseteq A, \quad \exists s.E \sqsubseteq Q_1$$

and let $\Sigma = \{c, s, Q_1\}$. Observe that A is implicitly definable in $\mathcal{O}$ using Σ as

$$\mathcal{O} \models A \equiv \forall s.(\exists s.\{c\} \to Q_1)$$

The same argument as above applied to the interpretations in Fig. 4 shows that there does not exist any $\mathcal{ELO}_u(\Sigma)$-explicit definition of A under $\mathcal{O}$. $\square$

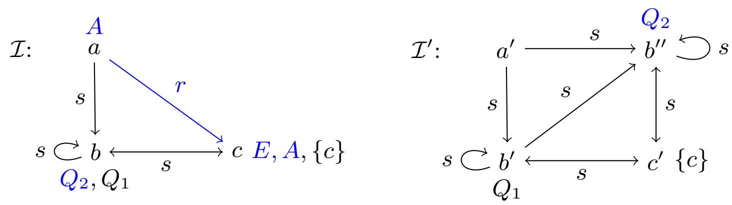


Fig. 4. Interpretations $\mathcal{I}$ and $\mathcal{I}'$ used to show that PBDP fails in $\mathcal{ELO}_u$.

Now to the positive cases. The reader might have noticed that in all examples in the proof of Theorem 15 the interpretation $\mathcal{I}$ on the left is the canonical model for $\mathcal{O}$ and A, whereas the interpretation $\mathcal{I}'$ on the right is the canonical model of $\mathcal{O}$ and the Σ-reduct of $\mathcal{I}$ seen as an ABox. Then the fact that $\rho_A \notin A^{\mathcal{I}'}$, where ρ_A is the root of $\mathcal{I}'$ (node a' in our examples), proves that an explicit definition does not exist. The converse is actually also true for $\mathcal{EL}^{++}$ and $\mathcal{EL}_u^{++}$: We prove in Section 6 that if a $\rho_A \in A^{\mathcal{I}'}$ then A is explicitly definable. This criterion can be used to show that a fragment $\mathcal{L}$ of $\mathcal{EL}^{++}/\mathcal{EL}_u^{++}$ enjoys CIP/PBDP. It suffices to show that $\rho_A \in A^{\mathcal{I}'}$ holds for any $\mathcal{L}$-ontology $\mathcal{O}$ and a concept name A that is implicitly definable under $\mathcal{O}$.

An argument along similar lines has been used in [28,33] to show that $\mathcal{EL}$ and the extension $\mathcal{ELH}$ of $\mathcal{EL}$ with *role hierarchies* (RIs of the form $r \sqsubseteq s$ with r, s role names) enjoy the CIP/PBDP. We extend this result to cover somewhat more complex role inclusions. We formulate the statement for CIP; PBDP follows from Remark 13.

A set $\mathcal{R}$ of RIs is *safe* for a signature Σ if one of the following two conditions holds:

- all inclusions in $\mathcal{R}$ are hole hierarchies, that is, inclusions of the form $r \sqsubseteq s$, or

– for each RI $r_1 \circ \cdots \circ r_n \sqsubseteq r \in \mathcal{R}$, $n \geq 1$, if $\{r_1, \ldots, r_n, r\} \cap \Sigma \neq \emptyset$ then $\{r_1, \ldots, r_n, r\} \subseteq \Sigma$.

Notice that if Σ includes all the role names then any set of RIs is safe.

Theorem 16. *Let $\mathcal{O}_1, \mathcal{O}_2$ be $\mathcal{EL}$-ontologies with RIs, C_1, C_2 be $\mathcal{EL}$-concepts, and set $\Sigma = sig(\mathcal{O}_1, C_1) \cap sig(\mathcal{O}_2, C_2)$. Assume that the set of RIs in $\mathcal{O}_1 \cup \mathcal{O}_2$ is safe for Σ and $\mathcal{O}_1 \cup \mathcal{O}_2 \models C_1 \sqsubseteq C_2$. Then an $\mathcal{EL}$-interpolant for $C_1 \sqsubseteq C_2$ under $\mathcal{O}_1$, $\mathcal{O}_2$ exists.*

We defer the proof of Theorem 16 to Section 6 and switch our attention to the inverse roles.

Theorem 17. *The DLs $\mathcal{ELI}$, $\mathcal{ELIO}$, $\mathcal{ELI}_u$ and $\mathcal{ELIO}_u$ do not enjoy the CIP and PBDP.*

Proof. We first consider $\mathcal{ELI}$. Let $\mathcal{O}$ be the ontology used in the proof of Theorem 15 to show that $\mathcal{EL}_u$ does not enjoy the PBDP. Obtain an $\mathcal{ELI}$-ontology $\mathcal{O}'$ from $\mathcal{O}$ by replacing the second CI of $\mathcal{O}$ by

$$D \sqcap \exists r^-.A \sqsubseteq E$$

Let $\Sigma = \{B, D, E, r\}$. Then A is implicitly definable from Σ under $\mathcal{O}'$ (the same explicit definition in Horn-$\mathcal{ALC}$ works), but A is not explicitly $\mathcal{ELI}(\Sigma)$-definable under $\mathcal{O}'$. This proof also works for the remaining DLs mentioned in Theorem 17 because one can show that A is also not explicitly $\mathcal{ELIO}_u(\Sigma)$-definable under $\mathcal{O}'$, and $\mathcal{ELIO}_u(\Sigma)$ is the strongest DL mentioned in Theorem 17. In fact, not only does A not have an explicit $\mathcal{ELIO}_u(\Sigma)$ definition, but no such definition exists in the positive fragment of $\mathcal{ALC}_u$. To see this, consider the interpretations given in Fig. 5 Observe that the models $\mathcal{I}, \mathcal{I}'$ show that A is not definable under $\mathcal{O}$ using

Fig. 5. Interpretations $\mathcal{I}$ and $\mathcal{I}'$ used to show that PBDP fails in $\mathcal{ELI}$.

any concept constructed from Σ using $\sqcap, \sqcup, \exists, \forall$ since for any such concept we have for $(x, x') \in \{(a, a'), (b, b'), (c, c'), (c, c'')\}$ that $x \in F^{\mathcal{I}}$ implies $x' \in F^{\mathcal{I}}$. □

Of course, interpretations $\mathcal{I}$ and $\mathcal{I}'$ given in Fig. 5 also demonstrate that concepts with implicit definitions in $\mathcal{EL}_u$ may not have explicit definitions in positive $\mathcal{ALC}_u$. What singles out Fig. 5 is that the interpretations $\mathcal{I}$ and $\mathcal{I}'$ are not the canonical models as nodes c and c' 'do not have to be there'. We need this nodes to ensure $\forall r.E$ does not distinguish a and a'.

5 Interpolant and Explicit Definition Existence

We introduce interpolant and explicit definition existence as decision problems and establish a polynomial time reduction of the latter to the former. We then show that it suffices to consider ontologies in normal form and that the addition of $\bot$ does not affect the complexity of the decision problems.

Definition 18. *Let $\mathcal{L}$ be a DL. Then* $\mathcal{L}$-interpolant existence *is the problem to decide for any $\mathcal{L}$-ontologies $\mathcal{O}_1, \mathcal{O}_2$ and $\mathcal{L}$-concepts C_1, C_2 whether there exists an $\mathcal{L}$-interpolant for $C_1 \sqsubseteq C_2$ under $\mathcal{O}_1, \mathcal{O}_2$.*

Observe that interpolant existence reduces to checking $\mathcal{O}_1 \cup \mathcal{O}_2 \models C_1 \sqsubseteq C_2$ for logics with the CIP but that this is not the case for logics without the CIP.

Definition 19. *Let $\mathcal{L}$ be a DL. Then* $\mathcal{L}$-explicit definition existence *is the problem to decide for any $\mathcal{L}$-ontology $\mathcal{O}$, signature Σ, and concept name A whether A is explicitly definable in $\mathcal{L}(\Sigma)$ under $\mathcal{O}$.*

Remark 20. There is a polynomial time reduction of $\mathcal{L}$-explicit definition existence to $\mathcal{L}$-interpolant existence. Moreover, any algorithm computing $\mathcal{L}$-interpolants also computes $\mathcal{L}$-explicit definitions and any bound of the size of $\mathcal{L}$-interpolants provides a bound on the size of $\mathcal{L}$-explicit definitions. The proof uses the same idea as Remark 13. Assume that an $\mathcal{L}$-ontology $\mathcal{O}$, concept name A, and a signature Σ are given. Let $\mathcal{O}_\Sigma$ be defined as above. Then any concept C is an $\mathcal{L}$-interpolant for $A \sqsubseteq A'$ under $\mathcal{O}, \mathcal{O}_\Sigma$ iff it is an explicit $\mathcal{L}(\Sigma)$-definition of A under $\mathcal{O}$.

We next observe that replacing the original ontologies by model conservative extensions preserves interpolants and explicit definitions. Thus, it suffices to consider ontologies in normal form and interpolants for inclusions between concept names.

Lemma 21. *Let $\mathcal{O}_1, \mathcal{O}_2$ be ontologies and C_1, C_2 concepts in any DL $\mathcal{L}$ considered in this paper. Then one can compute in polynomial time $\mathcal{L}$-ontologies $\mathcal{O}_1', \mathcal{O}_2'$ in normal form and with fresh concept names A, B such that an $\mathcal{L}$-concept C is an interpolant for $C_1 \sqsubseteq C_2$ under $\mathcal{O}_1, \mathcal{O}_2$ iff it is an interpolant for $A \sqsubseteq B$ under $\mathcal{O}_1', \mathcal{O}_2'$.*

Proof. Let $\mathcal{O}_1'$ and $\mathcal{O}_2'$ be the model conservative extensions of $\mathcal{O}_1 \cup \{A \equiv C\}$ and, respectively, $\mathcal{O}_2 \cup \{B \equiv D\}$, provided by Lemma 4. One can show that $\mathcal{O}_1'$ and $\mathcal{O}_2'$ are as required. $\square$

Observe also that an $\mathcal{L}$-concept C is an explicit $\mathcal{L}(\Sigma)$-definition of a concept name A under an $\mathcal{L}$-ontology $\mathcal{O}$ iff it is an explicit $\mathcal{L}(\Sigma)$-definition of A under the ontology $\mathcal{O}'$ in normal form given by Lemma 4, provided that Σ does not contain any symbols in $\mathrm{sig}(\mathcal{O}') \setminus \mathrm{sig}(\mathcal{O})$. We finally consider extensions with $\bot$.

Remark 22. Assume that $\mathcal{L}$ is any of the DLs from the $\mathcal{EL}$ or $\mathcal{ELI}$ family introduced above and let $\mathcal{L}_\bot$ denote its extension with $\bot$. Then there are trivial polynomial time reductions of $\mathcal{L}$-interpolant existence and $\mathcal{L}$-explicit definition existence to $\mathcal{L}_\bot$-interpolant existence and $\mathcal{L}_\bot$-explicit definition existence, respectively (see Remark 14). We establish polynomial time reductions in the converse direction modulo an oracle checking $\mathcal{O} \models C \sqsubseteq \bot$. We consider the CIP, the reduction for the PBDP is similar. The idea is the same as in Remark 14. Assume that $C \sqsubseteq D$ and $\mathcal{O}_1, \mathcal{O}_2$ are in $\mathcal{L}_\bot$. If $\mathcal{O}_1 \cup \mathcal{O}_2 \models C \sqsubseteq \bot$, then an $\mathcal{L}_\bot$-interpolant exists and we are done. Otherwise assume that no CI in $\mathcal{O}_1 \cup \mathcal{O}_2$ uses $\bot$ in the concept on its left hand side (if it does, the CI is redundant). From $\mathcal{O}_1 \cup \mathcal{O}_2 \not\models C \sqsubseteq \bot$ it follows that C do not use $\bot$. If D uses $\bot$, then $\mathcal{O}_1 \cup \mathcal{O}_2 \not\models C \sqsubseteq D$, and we are done as no $\mathcal{L}_\bot$-interpolant exists. Otherwise we have the following reduction: there exists an $\mathcal{L}_\bot$-interpolant for $C \sqsubseteq D$ under $\mathcal{O}_1 \cup \mathcal{O}_2$ iff there exists and $\mathcal{L}$-interpolant for $C \sqsubseteq D$ under $\mathcal{O}_1' \cup \mathcal{O}_2'$, where $\mathcal{O}_i'$ is obtained from $\mathcal{O}_i$ by replacing all occurrence of $\bot$ by a fresh concept name B_i, for $i = 1, 2$.

6 Interpolant and Explicit Definition Existence in $\mathcal{EL}$ and its Extensions

We show that for any extension of $\mathcal{EL}$ without inverse roles introduced above interpolant existence and explicit definition existence are decidable in PTime; moreover if an interpolant/explicit definition exists, then there exists one of at most exponential size. We note that for DLs that enjoy the CIP the complexity upper bound follows directly from the fact that subsumption is in PTime. We thus focus on those DLs that do not enjoy the CIP or PBDP.

Theorem 23. *Let $\mathcal{L} \in \{\mathcal{EL}_u, \mathcal{ELO}, \mathcal{ELO}_u, \mathcal{EL}^{++}, \mathcal{EL}_u^{++}\}$. Then $\mathcal{L}$-interpolant existence and $\mathcal{L}$-explicit definition existence are both in* PTime. *Moreover, if an interpolant or explicit definition exists, then there exists one of at most exponential size. This bound is optimal.*

We provide two proofs of the complexity upper bound. The first proof does not provide an upper bound on the size of interpolants/explicit definitions, but is more elementary. The second one provides a bound on the size of interpolants and explicit definitions[2]. Before we start with the upper bound proofs we give the straightforward but instructive proof that the exponential bound on the size of explicit definitions is optimal.

[2] Yet another way of computing interpolants is by analysing proofs in calculi that enjoy the subformula property. It has been shown in [26] that the consequence-based calculus implemented in the ELK reasoner has this property when inputs are restricted to $\mathcal{EL}$. We envisage that this approach would be most practical.

Example 24. The following example is folklore and has already been used for various succinctness arguments in DL. Let

$$\begin{aligned}\mathcal{O}_0 = \{&A \sqsubseteq M \sqcap \exists r_1.B_1 \sqcap \exists r_2.B_1\} \cup \\ &\{B_i \sqsubseteq \exists r_1.B_{i+1} \sqcap \exists r_2.B_{i+1} \mid i < n\} \cup \\ &\{B_n \sqsubseteq B, \exists r_1.B \sqcap \exists r_2.B \sqsubseteq B, B \sqcap M \sqsubseteq A\}\end{aligned}$$

and $\Sigma_0 = \{r_1, r_2, B_n\}$. Note that A triggers a marker M and a binary tree of depth n whose leafs are decorated with B_n. Conversely, if B_n is true at all leafs of a binary tree, then B is true at all nodes of the tree and B together with M entail A at its root. Let, inductively, $C_0 := B_n$ and $C_{i+1} = \exists r_1.C_i \sqcap \exists r_2.C_i$, for $0 < i < n$, and $C = M \sqcap C_n$. Then C is the smallest explicit $\mathcal{EL}(\Sigma_0)$-definition of A under $\mathcal{O}_0$.

Next let

$$\begin{aligned}\mathcal{O}_1 = \{&r_i \circ r_i \sqsubseteq r_{i+1} \mid 0 \leq i < n\} \cup \\ &\{A \sqsubseteq \exists r_0.B, B \sqsubseteq \exists r_0.B, \exists r_n.B \sqsubseteq A\}\end{aligned}$$

and $\Sigma_1 = \{r_0, B\}$. Then $\exists r_0^{2^n}.B$ is the smallest explicit $\mathcal{EL}(\Sigma_1)$-definition of A under $\mathcal{O}_1$. Observe that in the first example one enforces explicit definitions of exponential size by generating a binary tree of linear depth whereas in the second example one enforces an explicit definition of exponential size by generating a path of exponential length. The latter can only happen if role inclusions are used in the ontology. The main insight of the exponential upper bound on the size of explicit definitions in Theorem 23 is that the two examples cannot be combined to enforce a binary tree of exponential depth.

First Proof of PTime upper bound. We prove a characterization for the existence of interpolants using canonical models and simulations.

Lemma 25. *Let $\mathcal{O}_1, \mathcal{O}_2$ be $\mathcal{EL}_u^{++}$-ontologies in normal form, A, B concept names, and $\mathcal{L} \in \{\mathcal{ELO}, \mathcal{ELO}_u\}$. Let $\Sigma = sig(\mathcal{O}_1, A) \cap sig(\mathcal{O}_2, B)$. Then there does not exist an $\mathcal{L}$-interpolant for $A \sqsubseteq B$ under $\mathcal{O}_1, \mathcal{O}_2$ iff there exists a model $\mathcal{J}$ of $\mathcal{O}_1 \cup \mathcal{O}_2$ and $d \in \Delta^{\mathcal{J}}$ such that*

1. *$d \notin B^{\mathcal{J}}$;*
2. *$(\mathcal{I}_{\mathcal{O}_1 \cup \mathcal{O}_2, A}, \rho_A) \preceq_{\mathcal{L},\Sigma} (\mathcal{J}, d)$.*

Proof. Assume an $\mathcal{L}$-interpolant F exists, but there exists a model $\mathcal{J}$ of $\mathcal{O}_1 \cup \mathcal{O}_2$ and $d \in \Delta^{\mathcal{J}}$ satisfying the conditions of the lemma. As $\mathcal{O}_1 \cup \mathcal{O}_2 \models A \sqsubseteq F$, by Lemma 7, we obtain $\rho_A \in F^{\mathcal{I}_{\mathcal{O}_1 \cup \mathcal{O}_2, A}}$. By Lemma 6, $d \in F^{\mathcal{J}}$. We have derived a contradiction to the condition that $d \notin B^{\mathcal{J}}$, $\mathcal{J}$ is a model of $\mathcal{O}_1 \cup \mathcal{O}_2$, and $\mathcal{O}_1 \cup \mathcal{O}_2 \models F \sqsubseteq B$.

Assume no $\mathcal{L}$-interpolant exists. Let

$$\Gamma = \{C \in \mathcal{L}(\Sigma) \mid \rho_A \in C^{\mathcal{I}_{\mathcal{O}_1 \cup \mathcal{O}_2, A}}\}$$

By Lemma 8 and compactness, there exists a model $\mathcal{J}$ of $\mathcal{O}_1 \cup \mathcal{O}_2$ and $d \in \Delta^{\mathcal{J}}$ such that $d \in C^{\mathcal{J}}$ for all $C \in \Gamma$ but $d \notin B^{\mathcal{J}}$. We may assume that $\mathcal{J}$ is ω-saturated. Thus, by Lemma 6, $(\mathcal{I}_{\mathcal{O}_1\cup\mathcal{O}_2,A}, \rho_A) \preceq_{\mathcal{L},\Sigma} (\mathcal{J}, d)$, and $\mathcal{J}$ satisfies the conditions of the lemma. □

The characterization provided in Lemma 25 can be checked in polynomial time. Consider a fresh concept name X_d for each $d \in \Delta^{\mathcal{I}}$ for $\mathcal{I} = \mathcal{I}_{\mathcal{O}_1\cup\mathcal{O}_2,A}$. We define the $\mathcal{ELO}_u(\Sigma)$ *diagram* $\mathcal{D}(\mathcal{I})$ of $\mathcal{I}$ as the ontology consisting of the following CIs:

- $X_d \sqsubseteq A$, for every $A \in \Sigma$ and $d \in A^{\mathcal{I}}$;
- $X_{b^{\mathcal{I}}} \sqsubseteq \{b\}$, for every $b \in \Sigma$;
- $X_d \sqsubseteq \exists r.X_{d'}$, for every $r \in \Sigma$ and $(d, d') \in r^{\mathcal{I}}$;
- $X_d \sqsubseteq \exists u.X_{d'}$, for every $d, d' \in \Delta^{\mathcal{I}}$.

Denote by $\mathcal{I}_{|\Sigma}$ the Σ-reduct of the interpretation $\mathcal{I}$. Now it is straightforward to show that there exists a model $\mathcal{J}$ of $\mathcal{O}_1 \cup \mathcal{O}_2$ and $d \in \Delta^{\mathcal{J}}$ such that the conditions of Lemma 25 hold for $\mathcal{L} = \mathcal{ELO}_u$ iff $\mathcal{O}_1 \cup \mathcal{O}_2 \cup \mathcal{D}((\mathcal{I}_{\mathcal{O}_1\cup\mathcal{O}_2,A})_{|\Sigma}) \not\models X_{\rho_A} \sqsubseteq B$. The latter condition can be checked in polynomial time. If we aim at interpolants without the universal role we simply remove the CIs of the final item from the definition of $\mathcal{D}(\mathcal{I})$, denote the resulting set of inclusions by $\mathcal{D}'(\mathcal{I})$ and have that there exists a model $\mathcal{J}$ of $\mathcal{O}_1\cup\mathcal{O}_2$ and $d \in \Delta^{\mathcal{J}}$ such that the conditions of Lemma 25 hold for $\mathcal{L} = \mathcal{ELO}$ iff $\mathcal{O}_1 \cup \mathcal{O}_2 \cup \mathcal{D}'((\mathcal{I}_{\mathcal{O}_1\cup\mathcal{O}_2,A})_{|\Sigma}) \not\models X_{\rho_A} \sqsubseteq B$.

Proof of Theorem 23. The proof above does not provide a direct way to bound the size of interpolants. The following proof addresses this problem. We require the unfolding of an ABox into an ABox that is "equivalent" to an $\mathcal{ELO}_u$-concept. Let $\mathcal{A}$ be a factorized Σ-ABox and $\Gamma = \mathsf{N_I} \cap \Sigma$. The *directed unfolding of* $\mathcal{A}$ into a ditree-shaped ABox $\mathcal{A}^u$ modulo Γ is defined as follows. The individuals of $\mathcal{A}^u$ are the set of words $w = x_0 r_1 \cdots r_n x_n$ with $r_1, \ldots, r_n$ role names and $x_0, \ldots x_n \in \mathrm{ind}(\mathcal{A})$ such that $\{a\}(x_i) \notin \Gamma$ for any $i \neq 0$ and $a \in \Gamma$ and $r_{i+1}(x_i, x_{i+1}) \in \mathcal{A}$ for all $i < n$. We set $\mathrm{tail}(w) = x_n$ and define

- $A(w) \in \mathcal{A}^u$ if $A(\mathrm{tail}(w)) \in \mathcal{A}$, for $A \in \mathsf{N_C}$;
- $r(w, wrx) \in \mathcal{A}^u$ if $r(\mathrm{tail}(w), x) \in \mathcal{A}$ and $r(w, x) \in \mathcal{A}^u$ if $\{a\}(x) \in \mathcal{A}$ for some $a \in \Gamma$ and $r(\mathrm{tail}(w), x) \in \mathcal{A}$, for $r \in \mathsf{N_R}$;
- $\{a\}(x) \in \mathcal{A}^u$ if $\{a\}(x) \in \mathcal{A}$, for $a \in \Gamma$ and $x \in \mathrm{ind}(\mathcal{A})$.

Assume now that $\mathcal{O}_1, \mathcal{O}_2, A, B$ are given with $\mathcal{O}_1, \mathcal{O}_2$ in normal form. Let $\mathcal{A}^{\Sigma}_{\mathcal{O}_1\cup\mathcal{O}_2,A}$ be the Σ-reduct of the canonical model $\mathcal{I}_{\mathcal{O}_1\cup\mathcal{O}_2,A}$, regarded as an ABox. Consider the directed unfolding $\mathcal{A}^{\Sigma,u}_{\mathcal{O}_1\cup\mathcal{O}_2,A}$ of $\mathcal{A}^{\Sigma}_{\mathcal{O}_1\cup\mathcal{O}_2,A}$ modulo $\Gamma = \Sigma \cap \mathsf{N_I}$. Note that by Lemma 5 for every finite subset $\mathcal{A}$ of $\mathcal{A}^{\Sigma,u}_{\mathcal{O}_1\cup\mathcal{O}_2,A}$ containing ρ_A, the pointed ABox $\mathcal{A}, \rho_A$ corresponds to an $\mathcal{ELO}_u(\Sigma)$-concept.

Theorem 26. *Let $\mathcal{O}_1, \mathcal{O}_2$ be $\mathcal{EL}_u^{++}$-ontologies in normal form and let A, B be concept names. Let $\Sigma = sig(\mathcal{O}_1, A) \cap sig(\mathcal{O}_2, B)$. Then the following conditions are equivalent:*

1. *An $\mathcal{ELO}_u$-interpolant for $A \sqsubseteq B$ under $\mathcal{O}_1, \mathcal{O}_2$ exists;*

2. $\mathcal{O}_1 \cup \mathcal{O}_2, \mathcal{A}^{\Sigma}_{\mathcal{O}_1\cup\mathcal{O}_2,A} \models B(\rho_A)$*;*

3. there exists a subset $\mathcal{A}$ *of* $\mathcal{A}^{\Sigma,u}_{\mathcal{O}_1\cup\mathcal{O}_2,A}$ *with* $|ind(\mathcal{A})| \leq ||\mathcal{O}_1 \cup \mathcal{O}_2||^{||\mathcal{O}_1\cup\mathcal{O}_2||^5}$ *such that the* $\mathcal{ELO}_u$*-concept corresponding to* $\mathcal{A}, \rho_A$ *is an* $\mathcal{ELO}_u$*-interpolant for* $A \sqsubseteq B$ *under* $\mathcal{O}_1, \mathcal{O}_2$*.*

Observe that the equivalence of Points 1. and 2. provides another proof of the polynomial time upper bound for interpolant existence. The proof of Theorem 26 requires some preparation. In particular, we make use of derivation trees. In the context of $\mathcal{EL}$ and $\mathcal{ELI}$ derivation trees have been introduced in [11,4]. Here we extend derivation trees to deal with nominals and role inclusions.

Derivation Trees. Fix an $\mathcal{ELO}_u^{++}$-ontology $\mathcal{O}$ in normal form, an ABox $\mathcal{A}$, $x_0 \in \mathsf{ind}(\mathcal{A})$, and $A_0 \in \mathsf{N_C}$. Let $\Delta = \text{ind}(\mathcal{A}) \cup ((\mathsf{N_I} \cup \mathsf{N_C}) \cap \text{sig}(\mathcal{O}))$. A *derivation tree* for the assertion $A_0(x_0)$ in $\mathcal{A}$ w.r.t. $\mathcal{O}$ is a finite $\Delta \times (\{\top\} \cup (\mathsf{N_C} \cup \mathsf{N_I}) \cap \text{sig}(\mathcal{O}))$-labeled tree (T, V), where T is a set of nodes and $V : T \to \Delta \times (\{\top\} \cup (\mathsf{N_C} \cup \mathsf{N_I}) \cap \text{sig}(\mathcal{O}))$ the labeling function, such that

- $V(\varepsilon) = (x_0, A_0)$;
- if $V(n) = (a, C)$, then $a \in \text{ind}(\mathcal{A})$ and $C(a) \in \mathcal{A}$ or $a \in \mathsf{N_I}$ and $C = \{a\}$ or
  1. $a = C = A$ for a concept name A and n has a successor n' with $V(n') = (b, A)$ for some b; or
  2. $a = C = A$ for a concept name A and n has a successor n' such that $V(n') = (b, C')$ and $\mathcal{O} \models C' \sqsubseteq \exists r.A$ for some role name or the universal role r; or
  3. n has successors n_1, n_2 with $V(n_i) = (a, C_i)$ and $C_i \in (\mathsf{N_C} \cup \mathsf{N_I}) \cap \text{sig}(\mathcal{O})$ and $\mathcal{O} \models C_1 \sqcap C_2 \sqsubseteq C$;
  4. n has successors n_1, n_2, n_3 with $V(n_1) = (b, C)$, $V(n_2) = (a, \{c\})$, and $V(n_3) = (b, \{c\})$; or
  5. n has successors $n_2, \ldots, n_{2k-2}, n'$ and possibly additional successors from $n_{1'}, n'_2, \ldots, n'_{2k-1}, n'_{2k}$ with $k \geq 1$ such that the following holds: there is a sequence $a_1, \ldots, a_{2k}$ of elements of Δ and a sequence $r_2, r_4, \ldots, r_{2k-2}$ of role names or u such that $a = a_1$ and for every a_{2i+1} either $a_{2i+1} = a_{2i+2}$ or there is c with $V(n'_{2i+1}) = (a_{2i+1}, \{c\})$ and $V(n'_{2i+2}) = (a_{2i+2}, \{c\})$ such that for every a_{2i} with $i < k$:
     - $r_{2i}(a_{2i}, a_{2i+1}) \in \mathcal{A}$ and $V(n_{2i}) = (a_{2i}, \top)$; or
     - $a_{2i+1} \in \mathsf{N_I} \cap \text{sig}(\mathcal{O})$ and there exists $C_{2i} \in (\mathsf{N_C} \cup \mathsf{N_I}) \cap \text{sig}(\mathcal{O})$ such that $V(n_{2i}) = (a_{2i}, C_{2i})$ and $\mathcal{O} \models C_{2i} \sqsubseteq \exists r_{2i}.\{a_{2i+1}\}$; or
     - $a_{2i+1} \in \mathsf{N_C} \cap \text{sig}(\mathcal{O})$ and there exists $C_{2i} \in (\mathsf{N_C} \cup \mathsf{N_I}) \cap \text{sig}(\mathcal{O})$ such that $V(n_{2i}) = (a_{2i}, C_{2i})$ and $\mathcal{O} \models C_{2i} \sqsubseteq \exists r_{2i}.a_{2i+1}$

     and there exist $C' \in (\mathsf{N_C} \cup \mathsf{N_I}) \cap \text{sig}(\mathcal{O})$ and a role name or the universal role r such that $V(n') = (a_{2k}, C')$, $\mathcal{O} \models \exists r.C' \sqsubseteq C$, and $\mathcal{O} \models r_2 \circ r_4 \circ \ldots \circ r_{2k-2} \sqsubseteq r$.

The purpose of Conditions 1 and 2 is to establish that it follows from $\mathcal{O}$ and $\mathcal{A}$ that A is not empty: this is the case if one can show that some $x \in \text{ind}(\mathcal{A})$ is in A, some individual a in $\mathcal{O}$ is in A, or there exists a $b \in \Delta$ such that b is in C' and $\mathcal{O} \models C' \sqsubseteq \exists r.A$. The purpose of the remaining three conditions should be obvious. We note that in Condition 5, while a sequence $a_1, \dots, a_{2k}$ of exponential length in $\mathcal{O}$ and $\mathcal{A}$ might be required, the number of distinct $a_1, \dots, a_{2k}$ is bounded by $||\text{ind}(\mathcal{A})|| + ||\mathcal{O}||$ and, as the number of all possible labels does not exceed $(||\text{ind}(\mathcal{A})|| + ||\mathcal{O}||) \times ||\mathcal{O}||$, at most $(||\text{ind}(\mathcal{A})|| + ||\mathcal{O}||)^2 \times ||\mathcal{O}||$ distinct successors are needed in the derivation tree. The following example illustrates this point.

Example 27. We use the ontology from Example 24. Recall that

$$\begin{aligned}\mathcal{O} = \{&r_i \circ r_i \sqsubseteq r_{i+1} \mid 0 \leq i < n\} \cup \\ &\{A \sqsubseteq \exists r_0.B, B \sqsubseteq \exists r_0.B, \exists r_n.B \sqsubseteq A\}\end{aligned}$$

Then $\mathcal{I}_{\mathcal{O},A}$ is defined by setting

$$\begin{aligned}\Delta^{\mathcal{I}_{\mathcal{O},\mathcal{A}}} &= \{x, y\}\\ A^{\mathcal{I}_{\mathcal{O},\mathcal{A}}} &= \{x\}\\ B^{\mathcal{I}_{\mathcal{O},\mathcal{A}}} &= \{y\}\\ r_i^{\mathcal{I}_{\mathcal{O},\mathcal{A}}} &= \{(x, y), (y, y)\}, \text{ for } 0 \leq i \leq n.\end{aligned}$$

Recall that $\exists r_0^{2^n}.B$ is an explicit definition of A under $\mathcal{O}$. Consider the ABox $\mathcal{A}$ corresponding to the Σ-reduct of $\mathcal{I}_{\mathcal{O},\mathcal{A}}$. A derivation tree for $A(x)$ in $\mathcal{A}$ w.r.t. $\mathcal{O}$ is defined by setting $V(\varepsilon) = (x, A)$ and taking three successors ε with labels $(x, \top)$, $(y, \top)$, and (y, B), by applying Condition 5. Note that for the notation used in Condition 5, $a_1 = a_2 = x$ and $a_3 = \dots = a_{2^n} = y$ and $n_3 = \dots = n_{2k-2}$.

Lemma 28. *$\mathcal{O}, \mathcal{A} \models A_0(x_0)$ if and only if there is a derivation tree for $A_0(x_0)$ in $\mathcal{A}$ w.r.t $\mathcal{O}$.*

Proof. ($\Leftarrow$) is straightforward. For ($\Rightarrow$), we construct a sequence of ABoxes $\mathcal{A}_0, \mathcal{A}_1, \dots$ as follows. Define $\mathcal{A}_0$ as the union of $\mathcal{A}$ and all assertions $\{a\}(a)$ with a a nominal in $\mathcal{O}$, and let $\mathcal{A}_{i+1}$ be obtained from $\mathcal{A}_i$ by applying one of the following rules:

1. if $A(b) \in \mathcal{A}_i$, then add $A(A)$ to $\mathcal{A}_i$;
2. if $C'(b) \in \mathcal{A}_i$ and $\mathcal{O}' \models C' \sqsubseteq \exists r.A$, then add $A(A)$ to $\mathcal{A}_i$;
3. if $C_1(a), C_2(a) \in \mathcal{A}_i$ and $\mathcal{O} \models C_1 \sqcap C_2 \sqsubseteq C$, then add $C(a)$ to $\mathcal{A}_i$;
4. if $C(b), \{c\}(a), \{c\}(b) \in \mathcal{A}_i$, then add $C(a)$ to $\mathcal{A}_i$;
5. if there is a sequence $a_1, \dots, a_{2k}$ of elements of Δ and a sequence $r_2, r_4, \dots, r_{2k-2}$ of role names or u such that $a = a_1$ and for every a_{2j+1} either $a_{2j+1} = a_{2j+2}$ or there is c with $\{c\}(a_{2j+1}), \{c\}(a_{2j+2}) \in \mathcal{A}_i$ such that for every a_{2j}:
   - $r_{2j}(a_{2j}, a_{2j+1}) \in \mathcal{A}$; or
   - $a_{2j+1} \in \mathsf{N_I} \cap \text{sig}(\mathcal{O})$ and there exists $C_{2j} \in (\mathsf{N_C} \cup \mathsf{N_I}) \cap \text{sig}(\mathcal{O})$ such that $C_{2j}(a_{2j}) \in \mathcal{A}_i$ and $\mathcal{O} \models C_{2j} \sqsubseteq \exists r_{2j}.\{a_{2j+1}\}$; or
   - $a_{2j+1} \in \mathsf{N_C} \cap \text{sig}(\mathcal{O})$ and there exists $C_{2j} \in (\mathsf{N_C} \cup \mathsf{N_I}) \cap \text{sig}(\mathcal{O})$ such that $C_{2j}(a_{2j}) \in \mathcal{A}_i$ and $\mathcal{O} \models C_{2j} \sqsubseteq \exists r_{2j}.a_{2j+1}$

   and there exist $C' \in (\mathsf{N_C} \cup \mathsf{N_I}) \cap \text{sig}(\mathcal{O})$ and a role name or the universal role r such that $C'(a_{2k}) \in \mathcal{A}_i$, $\mathcal{O} \models \exists r.C' \sqsubseteq C$, and $\mathcal{O} \models r_2 \circ r_4 \circ \dots \circ r_{2k-2} \sqsubseteq r$, then add $C(a)$ to $\mathcal{A}_i$.

Note that the sequence is finite, and denote by $\mathcal{A}^*$ the final ABox.

Claim. There is a model $\mathcal{I}, v$ of $\mathcal{A}^*$ and $\mathcal{O}$ such that for all $x \in \text{ind}(\mathcal{A})$ and $A \in \mathsf{N_C}$, $v(x)^\mathcal{I} \in A^\mathcal{I}$ implies $A(x) \in \mathcal{A}^*$.

Proof of the Claim. For all $a, b \in \mathsf{ind}(\mathcal{A}^*)$, we write $a \sim b$ if $\{c\}(a), \{c\}(b) \in \mathcal{A}^*$ for some c. Notice that due to Rule 4, $a \sim b$ implies $C(a) \in \mathcal{A}^*$ if and only if $C(b) \in \mathcal{A}^*$. In particular, $\sim$ is an equivalence relation. We let $[a]$ denote the equivalence class of a. Start with an interpretation $\mathcal{I}_0$ defined by:

$$\begin{aligned}
\Delta^{\mathcal{I}_0} &= \mathsf{ind}(\mathcal{A}^*)/{\sim} \\
A^{\mathcal{I}_0} &= \{[a] \mid A(a) \in \mathcal{A}^*\} \\
a^{\mathcal{I}_0} &= \{[a]\} \\
r^{\mathcal{I}_0} &= \{([a],[b]) \mid \exists a' \in [a], b' \in [b].\ r(a',b') \in \mathcal{A}^*\}\,.
\end{aligned}$$

By definition, $\mathcal{I}_0$ satisfies all CIs in $\mathcal{O}$ that do not involve role names. We next extend $\mathcal{I}_0$ by adding pairs of the form $([a],[b])$ with $b \in \mathsf{N_C} \cup \mathsf{N_I}$ to the interpretation of role names. In detail, if $[a] \in \Delta^{\mathcal{I}_0}$ and there exist $C \in \mathsf{N_C} \cup \mathsf{N_I}$ with $a \in C^{\mathcal{I}_0}$ and $c \in \mathsf{N_I}$ with $c \in [b]$ such that $\mathcal{O} \models C \sqsubseteq \exists r.\{c\}$, then add $([a],[b])$ to $r^{\mathcal{I}_0}$. Also, if $[a] \in \Delta^{\mathcal{I}_0}$ and there exist $C \in \mathsf{N_C} \cup \mathsf{N_I}$ with $a \in C^{\mathcal{I}_0}$ and $A \in \mathsf{N_C}$ with $A \in [b]$ such that $\mathcal{O} \models C \sqsubseteq \exists r.A$, then add $([a],[b])$ to $r^{\mathcal{I}_0}$. Finally, add any pair $([a],[b])$ to $r^{\mathcal{I}_0}$ if there exists an RI $r_1 \circ \dots \circ r_n \sqsubseteq r$ that follows from $\mathcal{O}$ such that $([a],[b])$ is in relation $r_1 \circ \dots \circ r_n$ under the updated interpretations of $r_1, \dots, r_n$. This defines an interpretation $\mathcal{I}$. By Rule 2 all CIs of the form $A \sqsubseteq \exists r.B$ are satisfied in $\mathcal{I}$. By Rule 5, all CIs of the form $\exists r.B \sqsubseteq A$ are satisfied as well. By the final update of the extension of role names, all RIs are satisfied in $\mathcal{I}$. This finishes the proof of the claim.

Now suppose $\mathcal{O}, \mathcal{A} \models A_0(x_0)$. By the Claim, we have $A_0(x_0) \in \mathcal{A}^*$. Since the five rules to construct $\mathcal{A}_0, \mathcal{A}_1, \dots$ are in one-to-one correspondence with Conditions (1)–(5) from the definition of derivation trees, we can inductively construct a derivation tree for $A_0(x_0)$ in $\mathcal{A}$ w.r.t. $\mathcal{O}$. □

The following lemma states the observation that for the derivation of a fact only those assertions are relevant whose individuals occur in labels of a derivation tree.

Lemma 29. *$\mathcal{O}, \mathcal{A} \models A_0(x_0)$ and assume that (T, V) is a derivation tree for $A_0(x_0)$ in $\mathcal{A}$ w.r.t $\mathcal{O}$. Let $\mathcal{A}'$ denote the set of assertions in $\mathcal{A}$ involving only $a \in ind(\mathcal{A})$ such that (a, C) is in the range of V. Then $\mathcal{O}, \mathcal{A}' \models A_0(x_0)$.*

We next show that one can construct from a derivation tree in $\mathcal{A}$ a derivation tree in its directed unfolding. The derivation tree has the same depth but the outdegree might be exponential. We use Example 27 to illustrate this.

Example 30. Recall the ABox $\mathcal{A}$ from Example 27. Consider its unfolding $\mathcal{A}^u$ which has individuals $x, xr_0y, xr_0yr_0y, \dots$ and assertions

$$B(xr_0y), B(xr_0yr_0y), \dots \quad r_0(x, xr_0y), r_0(xr_0y, xr_0yr_0y), \dots$$

Then ε has 2^n+1 successors in the derivation tree for $A(x)$ in $\mathcal{A}^u$ labeled with:

$$(x,\top),\quad (xr_0y,\top),\cdots,(x(r_0y)^{2n-1},\top),\quad (x(r_0y)^{2n},B)\,.$$

Lemma 31. *Let $\mathcal{A}$ be a Σ-ABox and $\mathcal{O}$ an $\mathcal{EL}_u^{++}$-ontology. Let $\Gamma = \mathsf{N_I}\cap\Sigma$ and let $\mathcal{A}^u$ be the directed unfolding of $\mathcal{A}$ modulo Γ. Let (T,V) be a derivation tree for $A_0(x_0)$ in $\mathcal{A}$ w.r.t. $\mathcal{O}$. Then there exists a derivation tree (T',V') for $A_0(x_0)$ in $\mathcal{A}^u$ w.r.t. $\mathcal{O}$ of the same depth as T and such that the outdegree of T' does not exceed* $\max\{3,3k\}$*, where k is the length of the longest chain $a_1\cdots a_n$ used in Condition 5 of the construction of the derivation tree (T,V).*

Proof. Assume that (T,V) is a derivation tree for $A_0(x_0)$ in $\mathcal{A}$ w.r.t. $\mathcal{O}$. We obtain a very similar derivation tree (T',V') for $A_0(x_0)$ in $\mathcal{A}^u$ w.r.t. $\mathcal{O}$: with the exception of Condition 5, the construction is identical. For Condition 5, one potentially has to introduce "copies" of the nodes in T which correspond to the fresh individuals introduced in the unfolded ABox. In the construction of T',V' it is always the case that if the label of n in (T,V) is (a,C), then the label of copies n' of n in (T',V') takes the form (w,C) with $\mathrm{tail}(w)=a$.

In detail, we define (T',V') as follows from (T,V), starting with the root by setting $V'(\varepsilon)=V(\varepsilon)=(x_0,A_0)$.

Assume m is a copy of n and $V(n)=(a,C)$ and $V'(m)=(w,C)$. To define the successors of m and their labelings we consider the possible derivation steps for (a,C) in $\mathcal{A}$.

1. $a=C=A$ for a concept name A and n has a successor n' with $V(n')=(b,A)$ for some b: take a copy m' of n' as the only successor of m and set $V'(m')=(b,A)$.
2. $a=C=A$ for a concept name A and n has a successor n' such that $V(n')=(b,C')$ and $\mathcal{O}\models C'\sqsubseteq\exists r.A$ for some role name or the universal role r: take a copy m' of n' as the only successor of m and set $V'(m')=(b,C)$
3. n has successors n_1,n_2 with $V(n_i)=(a,C_i)$ and $C_i\in(\mathsf{N_C}\cup\mathsf{N_I})\cap\mathrm{sig}(\mathcal{O})$ and $\mathcal{O}\models C_1\sqcap C_2\sqsubseteq C$: take copies m_1,m_2 of n_1,n_2 as the successors of m and set $V'(m_i)=(w,C_i)$.
4. n has successors n_1,n_2,n_3 with $V(n_1)=(b,C)$, $V(n_2)=(a,\{c\})$, and $V(n_3)=(b,\{c\})$: take copies m_1,m_2,m_3 of n_1,n_2,n_3 as successors of m and set $V(m_1)=(b,C)$, $V(m_2)=(w,\{c\})$, and $V(m_3)=(b,\{c\})$.
5. n has successors $n_2,\dots,n_{2k-2},n'$ and possibly additional successors from $n_{1'},n'_2,\dots,n'_{2k-1},n'_{2k}$ with $k\geq 1$ such that for the sequence $a_1,\dots,a_{2k}$ of elements of Δ we have: $a=a_1$ and for every a_{2i+1} either $a_{2i+1}=a_{2i+2}$ or there is c with $V(n'_{2i+1})=(a_{2i+1},\{c\})$ and $V(n'_{2i+2})=(a_{2i+2},\{c\})$ and for every a_{2i} with $i<k$:
   - $r_{2i}(a_{2i},a_{2i+1})\in\mathcal{A}$ and $V(n_{2i})=(a_{2i},\top)$; or
   - $a_{2i+1}\in(\mathsf{N_I}\cup\mathsf{N_C})\cap\mathrm{sig}(\mathcal{O})$ and $V(n_{2i})=(a_{2i},C_{2i})$;

   and $V(n')=(a_{2k},C')$. Then take copies of $m_2,\dots,m_{2k-2},m'$ of $n_2,\dots,n_{2k-2},n'$ and copies $m'_1,m'_2,\dots,m'_{2k-1},m'_{2k}$ of $n'_1,n'_2,\dots,n'_{2k-1},n'_{2k}$ and define V' and the new sequence $b_1,\dots,b_{2k}$ as follows: set $b_1=w$ and set $b_{2i+2}=b_{2i+1}$ if $a_{2i+2}=a_{2i+1}$. If this is not the case, then assume $V(n'_{2i+1})=(a_{2i+1},\{c\})$

and $V(n'_{2i+2}) = (a_{2i+2}, \{c\})$ and let $b_{2i+2} = a_{2i+2}$, $V'(m'_{2i+1}) = (b_{2i+1}, \{c\})$, and $V(m'_{2i+2}) = (a_{2i+2}, \{c\})$. Next let

- $b_{2i+1} = b_{2i}r_{2i}a_{2i+1}$ and $V'(m_{2i}) = (b_{2i}, \top)$ if $r_{2i}(a_{2i}, a_{2i+1}) \in \mathcal{A}$ and $V(n_{2i}) = (a_{2i}, \top)$;
- $b_{2i+1} = a_{2i+1}$ and $V'(m_{2i}) = (b_{2i}, C_{2i})$ if $a_{2i+1} \in (\mathsf{N_I} \cup \mathsf{N_C}) \cap \mathrm{sig}(\mathcal{O})$ and $V(n_{2i}) = (a_{2i}, C_{2i})$.

Finally let $V'(m') = (b_{2k}, C')$.

Then (T', V') is a derivation tree for $A_0(x_0)$ in $\mathcal{A}^u$ w.r.t. $\mathcal{O}$ satisfying the conditions of the lemma. □

We now come to the proof of Theorem 26.

Proof. "1. ⇒ 2." Let F be an $\mathcal{ELO}_u$-interpolant for $A \sqsubseteq B$ under $\mathcal{O}_1, \mathcal{O}_2$. By Lemma 9, $\mathcal{I}_{\mathcal{O}_1 \cup \mathcal{O}_2, A} \models F(\rho_A)$. Thus, by definition $\mathcal{A}^{\Sigma}_{\mathcal{O}_1 \cup \mathcal{O}_2, A} \models F(\rho_A)$. Hence $\mathcal{O}_1 \cup \mathcal{O}_2, \mathcal{A}^{\Sigma}_{\mathcal{O}_1 \cup \mathcal{O}_2, A} \models B(\rho_A)$ since $\mathcal{O}_1 \cup \mathcal{O}_2 \models F \sqsubseteq B$.

"2. ⇒ 3." By Lemma 28, there exists a derivation tree (T, V) for $B(\rho_A)$ in $\mathcal{A}^{\Sigma}_{\mathcal{O}_1 \cup \mathcal{O}_2, A}$ under $\mathcal{O}_1 \cup \mathcal{O}_2$. Clearly we may assume that on any path in T each pair (a, C) occurs at most once as a label (if $V(n_1) = V(n_2)$ for n_1 below n_2 in a derivation tree, replace n_2 (with the labeled tree below it) by n_1 (and the labeled tree below n_1)). Moreover, we may assume that in Point 5 the length of the sequence $a_1, \ldots, a_{2k}$ does not exceed the maximum of 3 and

$$m := 2k^{k'(|\mathrm{ind}(\mathcal{A}^{\Sigma}_{\mathcal{O}_1 \cup \mathcal{O}_2, A})| + ||\mathcal{O}_1 \cup \mathcal{O}_2||)^2} \leq 2k^{4||\mathcal{O}_1 \cup \mathcal{O}_2||^3}$$

where k is maximal such that an RI $r_1 \circ \cdots \circ r_k \sqsubseteq r$ occurs in $\mathcal{O}_1 \cup \mathcal{O}_2$ and k' is the number of distinct role names in $\mathcal{O}_1 \cup \mathcal{O}_2$. (To show this observe that if there is some path $r_1 \circ \cdots \circ r_n$ with $(c, d) \in (r_1 \circ \cdots \circ r_n)^{\mathcal{I}}$ and $\mathcal{O} \models r_1 \circ \cdots \circ r_n \sqsubseteq r$ for the canonical model $\mathcal{I} = \mathcal{I}_{\mathcal{O}_1 \cup \mathcal{O}_2, A}$, then there exists such a path of length at most $\frac{m}{2}$.)

From Lemma 31, we thus obtain a derivation tree (T', V') for $B(\rho_A)$ in $\mathcal{A}^{\Sigma,u}_{\mathcal{O}_1 \cup \mathcal{O}_2, A}$ w.r.t. $\mathcal{O}_1 \cup \mathcal{O}_2$ such that the depth of T' does not exceed the depth of T (which is bounded by $||\mathcal{O}_1 \cup \mathcal{O}_2||^2$) and the outdegree of T' does not exceed $\max\{3, m\}$. Let $\mathcal{A}$ denote the subset of $\mathcal{A}^{\Sigma,u}_{\mathcal{O}_1 \cup \mathcal{O}_2, A}$ containing all $a \in \mathrm{ind}(\mathcal{A}^{\Sigma,u}_{\mathcal{O}_1 \cup \mathcal{O}_2, A})$ such that some (a, C) occurs in the range of V'. Then $|\mathrm{ind}(\mathcal{A})| \leq ||\mathcal{O}_1 \cup \mathcal{O}_2||^{||\mathcal{O}_1 \cup \mathcal{O}_2||^5}$ and, by Lemma 29, $\mathcal{O}_1 \cup \mathcal{O}_2, \mathcal{A} \models B(\rho_A)$, as required.

"3. ⇒ 1." is trivial. □

Let $\mathcal{A}^{\Sigma,c}_{\mathcal{O}_1 \cup \mathcal{O}_2, A}$ denote the subset of $\mathcal{A}^{\Sigma}_{\mathcal{O}_1 \cup \mathcal{O}_2, A}$ containing all assertions with individual variables that are reachable from ρ_A along a path in the directed graph defined by $\mathcal{A}^{\Sigma}_{\mathcal{O}_1 \cup \mathcal{O}_2, A}$. $\mathcal{A}^{\Sigma,c,u}_{\mathcal{O}_1 \cup \mathcal{O}_2, A}$ denotes the directed unfolding of $\mathcal{A}^{\Sigma,c}_{\mathcal{O}_1 \cup \mathcal{O}_2, A}$. Then for every finite subset $\mathcal{A}$ of $\mathcal{A}^{\Sigma,c,u}_{\mathcal{O}_1 \cup \mathcal{O}_2, A}$ such that $\mathcal{A}, \rho_A$ is rooted, the pointed ABox $\mathcal{A}, \rho_A$ corresponds to an $\mathcal{ELO}(\Sigma)$-concept C.

Theorem 32. *Let $\mathcal{O}_1, \mathcal{O}_2$ be $\mathcal{EL}^{++}_u$-ontologies in normal form and let A, B be concept names. Let $\Sigma = sig(\mathcal{O}_1, A) \cap sig(\mathcal{O}_2, B)$. Then the following conditions are equivalent:*

1. *An $\mathcal{ELO}$-interpolant for $A \sqsubseteq B$ under $\mathcal{O}_1, \mathcal{O}_2$ exists;*
2. $\mathcal{O}_1 \cup \mathcal{O}_2, \mathcal{A}^{\Sigma,c}_{\mathcal{O}_1 \cup \mathcal{O}_2, A} \models B(\rho_A)$;
3. *there exists a rooted subset $\mathcal{A}$ of $\mathcal{A}^{\Sigma,c,u}_{\mathcal{O}_1 \cup \mathcal{O}_2, A}$ containing at most $||\mathcal{O}_1 \cup \mathcal{O}_2||^{||\mathcal{O}_1 \cup \mathcal{O}_2||^5+1}$ individuals such that the $\mathcal{ELO}$-concept C corresponding to $\mathcal{A}, \rho_A$ is an $\mathcal{ELO}$-interpolant for $A \sqsubseteq B$ under $\mathcal{O}_1, \mathcal{O}_2$.*

Proof. The proof is essentially the same as that of Theorem 26, except that one possible has to add directed Σ-paths in $\mathcal{A}^{\Sigma,c,u}_{\mathcal{O}_1 \cup \mathcal{O}_2, A}$ to the subset $\mathcal{A}$ of $\mathcal{A}^{\Sigma,c,u}_{\mathcal{O}_1 \cup \mathcal{O}_2, A}$ obtained from a derivation tree in order to obtain that $\mathcal{A}, \rho_A$ is rooted. The length of these paths can be bound by $||\mathcal{O}_1 \cup \mathcal{O}_2||$. □

We conclude the section with proving Theorem 16 as a a corollary of Theorem 32. We reformulate the theorem as a proposition about normal form ontologies; the general result easily follows from the properties of normalisation.

Proposition 33. *Let $\mathcal{O}_1, \mathcal{O}_2$ be $\mathcal{EL}$-ontologies in normal form with RIs, A, B concept names, and set $\Sigma = sig(\mathcal{O}_1, A) \cap sig(\mathcal{O}_2, B)$. Assume that the set of RIs in $\mathcal{O}_1 \cup \mathcal{O}_2$ is safe for Σ and $\mathcal{O}_1 \cup \mathcal{O}_2 \models A \sqsubseteq B$. Then an $\mathcal{EL}$-interpolant for $A \sqsubseteq B$ under $\mathcal{O}_1$, $\mathcal{O}_2$ exists.*

Proof. For convenience of notation, we assume w.l.o.g. that $A \in \text{sig}(\mathcal{O}_1)$, $B \in \text{sig}(\mathcal{O}_2)$ and $\{A, B\} \cap \Sigma = \emptyset$. Suppose for a proof by contradiction that $\mathcal{O}_1 \cup \mathcal{O}_2 \models A \sqsubseteq B$ but there exists no $\mathcal{EL}$-interpolant for $A \sqsubseteq B$. Then $\mathcal{O}_1 \cup \mathcal{O}_2, \mathcal{A}^{\Sigma,c}_{\mathcal{O}_1 \cup \mathcal{O}_2, A} \not\models B(\rho_A)$. Moreover, since the language under consideration contains neither nominals nor the universal role, this strengthens to $\mathcal{O}_1 \cup \mathcal{O}_2, \mathcal{A}^{\Sigma}_{\mathcal{O}_1 \cup \mathcal{O}_2, A} \not\models B(\rho_A)$.

Let $\mathcal{J}_0$ be the canonical model of $\mathcal{O}_1 \cup \mathcal{O}_2$ and $\mathcal{A}^{\Sigma}_{\mathcal{O}_1 \cup \mathcal{O}_2, A}$. In what follows, we identify the domain of $\mathcal{I}^{\Sigma}_{\mathcal{O}_1 \cup \mathcal{O}_2, A}$ and individuals of $\mathcal{A}^{\Sigma}_{\mathcal{O}_1 \cup \mathcal{O}_2, A}$, and consider both to be subsets of the domain of $\mathcal{J}_0$. By the properties of the canonical model, we then have $\rho_A \notin B^{\mathcal{J}_0}$. Furthermore, as $\mathcal{I}_{\mathcal{O}_1 \cup \mathcal{O}_2, A}$ is a model for both $\mathcal{O}_1 \cup \mathcal{O}_2$ and $\mathcal{A}^{\Sigma}_{\mathcal{O}_1 \cup \mathcal{O}_2, A}$, there exists a $\text{sig}(\mathcal{O}_1 \cup \mathcal{O}_2)$-simulation S between $\mathcal{J}_0$ and $\mathcal{I}_{\mathcal{O}_1 \cup \mathcal{O}_2, A}$ such that $(x, x) \in S$ for all $x \in \Delta^{\mathcal{I}_{\mathcal{O}_1 \cup \mathcal{O}_2, A}}$.

Consider an interpretation $\mathcal{J}_1$ defined as follows:

$$\begin{aligned}
&\Delta^{\mathcal{J}_1} = \Delta^{\mathcal{J}_0},\\
&P^{\mathcal{J}_1} = P^{\mathcal{J}_0} \cup P^{\mathcal{I}_{\mathcal{O}_1 \cup \mathcal{O}_2, A}}, \text{for all } P \in (\text{sig}(O_1) \setminus \Sigma),\\
&P^{\mathcal{J}_1} = P^{\mathcal{J}_0}, \text{for all } P \notin (\text{sig}(O_1) \setminus \Sigma),
\end{aligned}$$

where P is a concept or role name. If $\mathcal{J}_1 \models \mathcal{O}_1 \cup \mathcal{O}_2$ we immediately derive a contradiction as we then have $\rho_A \in \mathcal{A}^{\mathcal{J}_1}$ and $\rho_A \notin B^{\mathcal{J}_1}$, contradicting $\mathcal{O}_1 \cup \mathcal{O}_2 \models A \sqsubseteq B$.

– If $\mathcal{O}_1 \cup \mathcal{O}_2$ does not contain RIs or the first safety condition holds, that is, all RIs are role hierarchies then $\mathcal{J}_1$ is indeed a model of $\mathcal{O}_1 \cup \mathcal{O}_2$.
  To see that, first we consider the RIs. Consider $r \sqsubseteq s \in \mathcal{O}_1 \cup \mathcal{O}_2$ and let $\{d_1, d_2\} \subseteq \Delta^{\mathcal{I}_{\mathcal{O}_1 \cup \mathcal{O}_2, A}}$ be such that $(d_1, d_2) \in r^{\mathcal{J}_1}$ but $(d_1, d_2) \notin r^{\mathcal{J}_0}$ (in

other cases the RI is satisfied already in $\mathcal{J}_0$.) By construction of $\mathcal{J}_1$ we have $r \in \text{sig}(\mathcal{O}_1) \setminus \Sigma$ so $r \sqsubseteq s \in \mathcal{O}_1$ and $s \notin \text{sig}(\mathcal{O}_2) \setminus \Sigma$. Since both $\mathcal{J}_0 \models r \sqsubseteq s$ and $\mathcal{I}_{\mathcal{O}_1 \cup \mathcal{O}_2, A} \models r \sqsubseteq s$ we have $(d_1, d_2) \in s^{\mathcal{J}_1}$.

We now consider CIs. As $\mathcal{J}_0$ and $\mathcal{J}_1$ are identical on all elements except $\Delta^{\mathcal{I}_{\mathcal{O}_1 \cup \mathcal{O}_2, A}}$, for all $x \in \Delta^{\mathcal{I}_{\mathcal{O}_1 \cup \mathcal{O}_2, A}}$ the relation S is a $\text{sig}(\mathcal{O}_1 \cup \mathcal{O}_2)$-simulation between $\mathcal{J}_1$ and $\mathcal{I}_{\mathcal{O}_1 \cup \mathcal{O}_2, A}$. Conversely, the embedding of $\mathcal{I}_{\mathcal{O}_1 \cup \mathcal{O}_2, A}$ into $\mathcal{J}_1$ generates a simulation, that is $(\mathcal{I}_{\mathcal{O}_1 \cup \mathcal{O}_2, A}, x) \preceq_{\mathcal{EL}, \text{sig}(\mathcal{O}_1)} (\mathcal{J}_1, x)$ for all $x \in \Delta^{\mathcal{I}_{\mathcal{O}_1 \cup \mathcal{O}_2, A}}$. By Lemma 6, for any $\text{sig}(O_1)$-$\mathcal{EL}$-concept C and for all $x \in \Delta^{\mathcal{I}_{\mathcal{O}_1 \cup \mathcal{O}_2, A}}$ we have $x \in C^{\mathcal{J}_1}$ if, and only if $x \in C^{\mathcal{I}_{\mathcal{O}_1 \cup \mathcal{O}_2, A}}$. Thus, $\mathcal{J}_1$ is a model of CIs in $\mathcal{O}_1$. By construction $\mathcal{J}_1 \models \mathcal{O}_2$.

– The interpretation $\mathcal{J}_1$ may not satisfy some RIs. We consider a sequence of interpretations $\mathcal{J}_i$ obtained by extending the interpretations of roles in $\mathcal{J}_1$ to satisfy RIs. For $i \geq 1$ we define

$$\begin{aligned}
\Delta^{\mathcal{J}_{i+1}} &= \Delta^{\mathcal{J}_i}, \\
A^{\mathcal{J}_{i+1}} &= A^{\mathcal{J}_i}, \text{for all } A \in \mathsf{N_C}, \\
r^{\mathcal{J}_{i+1}} &= r^{\mathcal{J}_i} \cup \left\{ (d_1, d_{n+1}) \;\middle|\; \begin{array}{l} r_1 \circ \cdots \circ r_n \sqsubseteq r \in \mathcal{O}_1 \cup \mathcal{O}_2 \\ \{d_1, \ldots, d_{n+1}\} \subseteq \Delta^{\mathcal{J}_i}, (d_1, d_{n+1}) \notin r^{\mathcal{J}_i} \\ (d_k, d_{k+1}) \in r_k^{\mathcal{J}_i} \text{ for all } 1 \leq k \leq n \end{array} \right\}
\end{aligned}$$

A simple inductive argument shows that the second safety condition and the fact that $(d_1, d_{n+1}) \notin r^{\mathcal{J}_i}$ imply that $\{r_1, \ldots, r_n, r\} \subseteq \text{sig}(\mathcal{O}_1)$.

Furthermore, we prove by induction that the relation S is a $\text{sig}(\mathcal{O}_1 \cup \mathcal{O}_2)$-simulation between $\mathcal{J}_{i+1}$ and $\mathcal{I}_{\mathcal{O}_1 \cup \mathcal{O}_2, A}$. For $i = 1$ this has been established above. For the induction stop it suffices to consider r-successors of d_1 in $\mathcal{J}_{i+1}$, where r is from the definition of $\mathcal{J}_{i+1}$ above. By the induction hypothesis, S is a a $\text{sig}(\mathcal{O}_1 \cup \mathcal{O}_2)$-simulation between $\mathcal{J}_i$ and $\mathcal{I}_{\mathcal{O}_1 \cup \mathcal{O}_2, A}$. Then there exist $\{v_2, \ldots, v_{n+1}\} \subseteq \Delta^{\mathcal{I}_{\mathcal{O}_1 \cup \mathcal{O}_2, A}}$ with $(d_{j+1}, v_{j+1}) \in S$ and $(v_1, v_{n+1}) \in {r_i}^{\mathcal{I}_{\mathcal{O}_1 \cup \mathcal{O}_2, A}}$ for $j \in \{1, \ldots, n\}$. As $\mathcal{I}_{\mathcal{O}_1 \cup \mathcal{O}_2, A}$ is a model of $\mathcal{O}_1$, we have $(v_1, v_{n+1}) \in r^{\mathcal{I}_{\mathcal{O}_1 \cup \mathcal{O}_2, A}}$ and $(d_{n+1}, v_{n+1}) \in S$ as required.

As $\mathcal{EL}$ canonical models defined in this paper are finite, there exists $N > 0$ such that for all $i > N$, $\mathcal{J}_i = \mathcal{J}_N$. It can be seen that $\mathcal{J}_N$ satisfies all RIs in $\mathcal{O}_1 \cup \mathcal{O}_2$ and the satisfaction of CIs is proved similarly to the case above. Then $\mathcal{J}_N$ is a model of $\mathcal{O}_1 \cup \mathcal{O}_2$ with $\rho_A \in \mathcal{A}^{\mathcal{J}_N}$ and $\rho_A \notin B^{\mathcal{J}_N}$, contradicting $\mathcal{O}_1 \cup \mathcal{O}_2 \models A \sqsubseteq B$.

□

7 Interpolant and Explicit Definition Existence in $\mathcal{ELI}$ and its Extensions

We show that for $\mathcal{ELI}$ and its extensions with nominals and universal roles, interpolant existence and explicit definition existence are decidable in ExpTime, matching the lower bound coming from subsumption. Moreover, for $\mathcal{ELI}$ and $\mathcal{ELI}_u$, if an interpolant/explicit definition exists, then there is one of at most double exponential size, and this is optimal.

Theorem 34. *Let $\mathcal{L} \in \{\mathcal{ELI}, \mathcal{ELI}_u, \mathcal{ELIO}, \mathcal{ELIO}_u\}$. Then $\mathcal{L}$-interpolant existence and $\mathcal{L}$-explicit definition existence are* ExpTime*-complete. Moreover, if $\mathcal{L} \in \{\mathcal{ELI}, \mathcal{ELI}_u\}$, then when an interpolant or explicit definition exists, there exists one of at most double exponential size. This bound is optimal.*

The section is organized as follows. We first introduce canonical models and derivation trees for $\mathcal{ELIO}_u$. We then give an ExpTime algorithm for interpolant existence, before discussing the size of interpolants.

Canonical Models. Fix an $\mathcal{ELIO}_u$-ontology $\mathcal{O}$ in normal form, and a concept name A_0. An *$\mathcal{O}$-type* is a set τ of concepts $C = A$, $C = \{a\}$ or $C = \exists u.A$ occurring in $\mathcal{O}$ (where $A \in \mathsf{N_C}$), and such that $\mathcal{O} \models \bigsqcap_{C\in\tau} C \sqsubseteq C'$ implies $C' \in \tau$ for all concepts C' of this form. We sometimes identify τ and $\bigsqcap_{C\in\tau} C$. For a role r, we write $\tau_1 \rightsquigarrow_r \tau_2$ if τ_2 is a maximal (w.r.t. inclusion) $\mathcal{O}$-type such that $\mathcal{O} \models \tau_1 \sqsubseteq \exists r.\tau_2$. We define $\tau_1 \rightsquigarrow_u \tau_2$ similarly, but additionally require that $A_1 \sqsubseteq \exists u.A_2 \in \mathcal{O}$ for some $A_1 \in \tau_1$ and $A_2 \in \tau_2$. Note that the set of all $\mathcal{O}$-types and relations $\rightsquigarrow_r$, $\rightsquigarrow_u$ can be computed in exponential time.

Let τ_{A_0} denote the set of all concepts C occurring in $\mathcal{O}$ such that $\mathcal{O} \models A_0 \sqsubseteq C$, and for all nominals a, $\tau_{A_0,a}$ the type consisting of all C such that $\mathcal{O} \models \{a\} \sqcap \exists u.A_0 \sqsubseteq C$. The *canonical model* of $\mathcal{I}_{\mathcal{O},A_0}$ of $\mathcal{O}$ and A_0 is defined as follows:

$$
\begin{aligned}
\Delta^{\mathcal{I}_{\mathcal{O},A_0}} &= \{\tau_{A_0}\} \cup \{\tau_{A_0,a} \mid a \in \mathsf{N_I} \cap \mathrm{sig}(\mathcal{O})\} \cup \\
&\qquad \{\tau_0 r_1 \tau_1 \cdots r_n \tau_n \mid n > 0, \tau_0 = \tau_{A_0}, \tau_i \rightsquigarrow_{r_{i+1}} \tau_{i+1}, \\
&\qquad\qquad \tau_n \notin \{\tau_{A_0}\} \cup \{\tau_{A_0,a} \mid a \in \mathsf{N_I} \cap \mathrm{sig}(\mathcal{O})\}\} \\
a^{\mathcal{I}_{\mathcal{O},A_0}} &= \tau_{A_0,a} \\
A^{\mathcal{I}_{\mathcal{O},A_0}} &= \{w\tau \mid w\tau \in \Delta^{\mathcal{I}_{\mathcal{O},A_0}}, A \in \tau\} \\
r^{\mathcal{I}_{\mathcal{O},A_0}} &= \{(w, wr\tau) \mid w, wr\tau \in \Delta^{\mathcal{I}_{\mathcal{O},A_0}}\} \cup \{(wr^-\tau, w) \mid w, wr^-\tau \in \Delta^{\mathcal{I}_{\mathcal{O},A_0}}\}.
\end{aligned}
$$

In particular, every type $\tau \in \Delta^{\mathcal{I}_{\mathcal{O},A_0}}$ contains the same set of concepts $\exists u.A$ as type τ_{A_0}.

We use $\mathcal{A}_{\mathcal{O},A_0}$ to denote the ABox associated with the canonical model $\mathcal{I}_{\mathcal{O},A_0}$, and $\mathcal{A}^{\Sigma}_{\mathcal{O},A_0}$ for its Σ-reduct. We let x_{A_0} denote the element of $\mathsf{ind}(\mathcal{A}_{\mathcal{O},A_0})$ corresponding to τ_{A_0}.

The following properties of canonical models can be proved in a standard way.

Lemma 35. *For all $\mathcal{ELIO}_u$-ontologies $\mathcal{O}$ in normal form and concept names A_0,*

1. *$\mathcal{I}_{\mathcal{O},A_0} \models \mathcal{O}$;*
2. *For every model $(\mathcal{I}, d)$ of $\mathcal{O}$ and A_0, there is a homomorphism from $(\mathcal{I}_{\mathcal{O},A_0}, \tau_{A_0})$ to $(\mathcal{I}, d)$;*
3. *For all concepts C, $\mathcal{O} \models A_0 \sqsubseteq C$ if and only if $\tau_{A_0} \in C^{\mathcal{I}_{\mathcal{O},A_0}}$.*

Derivation trees. Fix an $\mathcal{ELIO}_u$-ontology $\mathcal{O}$ in normal form and an ABox $\mathcal{A}$, $x_0 \in \mathsf{ind}(\mathcal{A})$ and $A_0 \in \mathsf{N_C}$. Let $\Theta_1 = \mathsf{ind}(\mathcal{A}) \cup (\mathsf{N_I} \cap \mathrm{sig}(\mathcal{O}))$, and $\Theta_2 = \mathsf{N_C} \cap$

$\mathrm{sig}(\mathcal{O}) \cup \{\{a\} \mid a \in \mathsf{N_I} \cap \mathrm{sig}(\mathcal{O})\} \cup \{\exists u.A \mid A \in \mathsf{N_C} \cap \mathrm{sig}(\mathcal{O})\}$. A *derivation tree* for the assertion $A_0(x_0)$ in $\mathcal{A}$ w.r.t. $\mathcal{O}$ is a finite $\Theta_1 \times \Theta_2$-labeled tree (T, V) such that:

- $V(\varepsilon) = (x_0, A_0)$;
- If $V(n) = (x, C)$ with $x \in \mathsf{ind}(\mathcal{A})$, then $\mathcal{O} \models \top \sqsubseteq C$ or
  1. n has successors $n_1, \ldots, n_k$, $k \geq 1$ with $V(n_i) = (a_i, C_i)$, such that $a_i = x$ or $a_i \in \mathsf{N_I} \cap \mathrm{sig}(\mathcal{O})$ for all i, and defining $C_i' = C_i$ if $a_i = x$, and $C_i' = \exists u.(\{a_i\} \sqcap C_i)$ otherwise, we have $\mathcal{O} \models C_1' \sqcap \ldots \sqcap C_k' \sqsubseteq C$; or
  2. $C = \exists u.A$ and n has a single successor n' with $V(n') = (y, \exists u.A)$; or
  3. n has a single successor n' with $V(n') = (y, A)$ such that $r(x, y) \in \mathcal{A}$ and $\mathcal{O} \models \exists r.A \sqsubseteq C$ (where r is a role name or an inverse role).
- If $V(n) = (a, C)$ with $a \in \mathsf{N_I} \cap \mathrm{sig}(\mathcal{O})$, then $C = \{a\}$ or:
  4. There exists $x \in \mathsf{ind}(\mathcal{A})$ such that n has successors $n_1, \ldots, n_k$, $k \geq 1$ with $V(n_i) = (a_i, C_i)$ and $a_i = x$ or $a_i \in \mathsf{N_I} \cap \mathrm{sig}(\mathcal{O})$ for all i, and, defining $C_i' = C_i$ if $a_i = x$, and $C_i' = \exists u.(\{a_i\} \sqcap C_i)$ otherwise, we have $\mathcal{O} \models C_1' \sqcap \ldots \sqcap C_k' \sqsubseteq \exists u.(\{a\} \sqcap C)$.

Note that a special case of rule 1 is when n has two successors labeled $(x, \{a\})$ and (a, C), and a special case of rule 4 is when n has two successors labeled $(x, \{a\})$ and (x, C).

Lemma 36. *$\mathcal{O}, \mathcal{A} \models A_0(x_0)$ if and only if there is a derivation tree for $A_0(x_0)$ in $\mathcal{A}$ w.r.t. $\mathcal{O}$.*

Proof. ($\Leftarrow$) is straightforward. For ($\Rightarrow$), we construct a sequence of ABoxes $\mathcal{A}_0, \mathcal{A}_1, \ldots$ generalized with assertions of the form $(\exists u.A)(x)$. Take $\mathcal{A}_0 = \mathcal{A} \cup \{\{a\}(x_a) \mid a \in \mathsf{N_I} \cap \mathrm{sig}(\mathcal{O})\}$ where the x_a's are fresh individual variables. Let $\mathcal{A}_{i+1}$ be obtained from $\mathcal{A}_i$ by applying one of the following rule, where C is a concept of the form $C \in \mathsf{N_C}$ or $C = \{a\}$ or $C = \exists u.A$, and $x, y \in \mathsf{ind}(\mathcal{A}_i)$:

1. if $C_1(x_1), \ldots, C_k(x_k) \in \mathcal{A}_i$, with $x_i = x$ or $x_i = x_{a_i}$ for some $a_i \in \mathsf{N_I} \cap \mathrm{sig}(\mathcal{O})$, and $\mathcal{O} \models C_1' \sqcap \ldots \sqcap C_k' \sqsubseteq C$, where $C_i' = C_i$ if $x_i = x$ and $C_i' = \exists u.(\{a_i\} \sqcap C_i)$ if $x = x_{a_i}$, then add $C(x)$;
2. if $(\exists u.A)(y) \in \mathcal{A}_i$ then add $(\exists u.A)(x)$;
3. if $r(x, y), A(y) \in \mathcal{A}_i$ and $\mathcal{O} \models \exists r.A \sqsubseteq C$, then add $C(x)$;
4. if $C_1(x_1), \ldots, C_k(x_k) \in \mathcal{A}_i$, with $x_i = x$ or $x_i = x_{a_i}$ for some $a_i \in \mathsf{N_I} \cap \mathrm{sig}(\mathcal{O})$, and $\mathcal{O} \models C_1' \sqcap \ldots \sqcap C_k' \sqsubseteq \exists u.(\{a\} \sqcap C)$, where $C_i' = C_i$ if $x_i = x$ and $C_i' = \exists u.(\{a_i\} \sqcap C_i)$ if $x = x_{a_i}$, then add $C(x_a)$.

Note that the sequence is finite, and denote by $\mathcal{A}^*$ the final ABox.

Claim. There is a model $\mathcal{I}, v$ of $\mathcal{A}^*$ and $\mathcal{O}$ such that for all $x \in \mathsf{ind}(\mathcal{A})$ and $A \in \mathsf{N_C}$, $v(x)^{\mathcal{I}} \in A^{\mathcal{I}}$ implies $A(x) \in \mathcal{A}^*$.

Proof of the Claim. For all $x, y \in \mathsf{ind}(\mathcal{A}^*)$, we write $x \sim y$ if $\{a\}(x), \{a\}(y) \in \mathcal{A}^*$ for some $a \in \mathsf{N_I} \cap \mathrm{sig}(\mathcal{O})$. Notice that if $\{a\}(x), \{a\}(y), C(x) \in \mathcal{A}^*$, then $C(x_a) \in \mathcal{A}^*$ by rule 4, and $C(y) \in \mathcal{A}^*$ by rule 1. Therefore, $x \sim y$ implies $C(x) \in \mathcal{A}^*$

if and only if $C(y) \in \mathcal{A}^*$. In particular, $\sim$ is an equivalence relation. We let $[x]$ denote the equivalence class of x. Start with an interpretation $\mathcal{I}_0$ defined by:

$$
\begin{aligned}
\Delta^{\mathcal{I}_0} &= \mathsf{ind}(\mathcal{A}^*)/{\sim} \\
A^{\mathcal{I}_0} &= \{[x] \mid A(x) \in \mathcal{A}^*\} \\
a^{\mathcal{I}_0} &= [x_a] \\
r^{\mathcal{I}_0} &= \{([x],[y]) \mid r(x,y) \in \mathcal{A}^*\}\,.
\end{aligned}
$$

Let C_x denote the conjunction of all concepts of the form $C \in \mathsf{N_C}$, $C = \{a\}$, $C = \exists u.A$, or $C = \exists u.(\{a\} \sqcap A)$ such that $\mathcal{A}^* \models C(x)$. Let $\mathcal{I}_x$ denote the canonical model for $\mathcal{O}$ and C_x rooted at $[x]$. Due to rule 1 and the universality of $\mathcal{I}_x$, for every concept name or nominal C, we have $[x] \in C^{\mathcal{I}_0}$ if and only if $[x] \in C^{\mathcal{I}_x}$. Similarly, because of rule 4, for every $a \in \mathsf{N_I} \cap \mathrm{sig}(\mathcal{O})$, $a^{\mathcal{I}_x} \in C^{\mathcal{I}_x}$ if and only if $a^{\mathcal{I}_0} \in C^{\mathcal{I}_0}$.

We can now define $\mathcal{I}$ as follows: $\Delta^{\mathcal{I}}$ is the disjoint union of $\Delta^{\mathcal{I}_0}$ and all elements in domains $\Delta^{\mathcal{I}_x} \setminus (\{[x]\} \cup \{a^{\mathcal{I}_x} \mid a \in \mathsf{N_I} \cap \mathrm{sig}(\mathcal{O})\})$. Interpretations of concept names and nominals are inherited from the $\mathcal{I}_0$ or $\mathcal{I}_x$ each element comes from. Finally, $r^{\mathcal{I}}$ is obtained by taking the union of $r^{\mathcal{I}_0}$ and all $r^{\mathcal{I}_x}$ after replacing edges to/from $a^{\mathcal{I}_x}$ with edges to/from $a^{\mathcal{I}_0}$. It is clear that for the variable assignment $v(x) = [x]$, $\mathcal{I}_0, v$ satisfies $\mathcal{A}^*$, and thus so does $\mathcal{I}, v$.

By rule 1, all concept inclusions of $\mathcal{O}$ of the form $\top \sqsubseteq A$, $A_1 \sqcap A_2 \sqsubseteq B$, $A \sqsubseteq \{a\}$ and $\{a\} \sqsubseteq A$ are satisfied by $\mathcal{I}_0$. They are also satisfied by every $\mathcal{I}_x$ (since $\mathcal{I}_x$ is a model of $\mathcal{O}$), and thus by $\mathcal{I}$. Now consider a concept inclusion $A \sqsubseteq \exists r.B \in \mathcal{O}$, where r is a role name or an inverse role. Recall that for every a and x, $a^{\mathcal{I}} \in B^{\mathcal{I}}$ if and only if $a^{\mathcal{I}_x} \in B^{\mathcal{I}_x}$. Therefore, for all $d \in \Delta^{\mathcal{I}_x}$, $d \in (\exists r.B)^{\mathcal{I}_x}$ implies $d \in (\exists r.B)^{\mathcal{I}}$. The case $A \sqsubseteq \exists u.B$ is similar. Since every $\mathcal{I}_x$ satisfies $A \sqsubseteq \exists r.B$, so does $\mathcal{I}$. Similarly, every concept inclusion $\exists r.B \sqsubseteq A \in \mathcal{O}$ is satisfied in $\mathcal{I}$: if the witness pair for $\exists r.B$ is part of $\mathcal{I}_0$, this follows from rule 3, and if not, then it is part of some $\mathcal{I}_x$, which is by definition a model of $\mathcal{O}$. For concept inclusions of the form $\exists u.B \sqsubseteq A \in \mathcal{O}$, we can observe that if there exists some $d' \in \Delta^{\mathcal{I}}$ such that $d' \in B^{\mathcal{I}}$, then $(\exists u.B)$ is in C_x for some x, i.e., by rule 2, for all x.

Finally, for all $x \in \mathsf{ind}(\mathcal{A})$ and $A \in \mathsf{N_C}$, $[x]^{\mathcal{I}} \in A^{\mathcal{I}}$ implies $[x]^{\mathcal{I}_0} \in A^{\mathcal{I}_0}$, i.e., $A(x) \in \mathcal{A}^*$. This concludes the proof of the claim.

Now suppose $\mathcal{O}, \mathcal{A} \models A_0(x_0)$. By the Claim, we have $A_0(x_0) \in \mathcal{A}^*$. Since the four rules to construct $\mathcal{A}_0, \mathcal{A}_1, \ldots$ are in one-to-one correspondence with Conditions (1)–(4) from the definition of derivation trees, we can inductively construct a derivation tree for $A_0(x_0)$ in $\mathcal{A}$ w.r.t. $\mathcal{O}$. □

A result analogous to Lemma 31 holds for $\mathcal{ELIO}_u$. The *undirected unfolding of an ABox* $\mathcal{A}$ into a tree-shaped ABox $\mathcal{A}^*$ modulo Γ is defined as follows: the individuals of $\mathcal{A}^*$ are the set of words $w = x_0 r_1 \cdots r_n x_n$ with $r_1, \ldots, r_n$ roles and $x_0, \ldots x_n \in \mathrm{ind}(\mathcal{A})$ such that $\{a\}(x_i) \notin \Gamma$ for any $i \neq 0$ and $r_{i+1}(x_i, x_{i+1}) \in \mathcal{A}$ if r_{i+1} is a role name and $r_{i+1}^-(x_{i+1}, x_i) \in \mathcal{A}$ if r_{i+1} is an inverse role, for all $i < n$. We set $\mathrm{tail}(w) = x_n$ and let

– $A(w) \in \mathcal{A}^*$ if $A(\mathrm{tail}(w)) \in \mathcal{A}$, for $A \in \mathsf{N_C}$;
– $r(w, wrx) \in \mathcal{A}^*$ if $r(\mathrm{tail}(w), x) \in \mathcal{A}$ and $r(w, x) \in \mathcal{A}^*$ if $\{a\}(x) \in \mathcal{A}$ for some $a \in \Gamma$ and $r(\mathrm{tail}(w), x) \in \mathcal{A}$, for $r \in \mathsf{N_R}$;
– $r(wrx, w) \in \mathcal{A}^*$ if $r(x, \mathrm{tail}(w)) \in \mathcal{A}$ and $r(x, w) \in \mathcal{A}^*$ if $\{a\}(x) \in \mathcal{A}$ for some $a \in \Gamma$ $r(x, \mathrm{tail}(w)) \in \mathcal{A}$, for $r \in \mathsf{N_R}$;
– $\{a\}(x) \in \mathcal{A}^*$ if $\{a\}(x) \in \mathcal{A}$, for $a \in \Gamma$ and $x \in \mathrm{ind}(\mathcal{A})$.

Lemma 37. *Let $\mathcal{A}$ be a Σ-ABox and $\mathcal{O}$ an $\mathcal{ELIO}_u$ -ontology. Let $\mathcal{A}^*$ be the directed unfolding of $\mathcal{A}$ modulo Γ. Let (T, V) be a derivation tree for $A_0(x_0)$ in $\mathcal{A}$ w.r.t. $\mathcal{O}$. Then there exists a derivation tree (T', V') for $A_0(x_0)$ in $\mathcal{A}^u$ w.r.t. $\mathcal{O}$ such that $T' = T$.*

Proof. The proof is similar to that of Lemma 31. We define (T, V') as follows from (T, V), starting with the root by setting $V'(\varepsilon) = V(\varepsilon) = (x_0, A_0)$. At each step, if $V(n) = (a, C)$ and $V'(n) = (w, C)$, then $\mathrm{tail}(w) = a$. To define the labelings of the successors of n, we consider the possible derivation steps for (a, C) in $\mathcal{A}$.

1. $a = x \in \mathsf{ind}(\mathcal{A})$, and n has successors $n_1, \ldots, n_k$, $k \geq 1$ with $V(n_i) = (a_i, C_i)$, such that $a_i = x$ or $a_i \in \mathsf{N_I} \cap \mathrm{sig}(\mathcal{O})$ for all i, and defining $C'_i = C_i$ if $a_i = x$, and $C'_i = \exists u.(\{a_i\} \sqcap C_i)$ otherwise, we have $\mathcal{O} \models C'_1 \sqcap \ldots \sqcap C'_k \sqsubseteq C$. Take $V'(n_i) = (w, C_i)$ if $x_i = x$, and $V'(n_i) = (a_i, C_i)$ if $a_i \in \mathsf{N_I} \cap \mathrm{sig}(\mathcal{O})$.
2. $C = \exists u.A$ and n has a single successor n' with $V(n') = (y, \exists u.A)$. Take $V'(n') = (y, \exists u.A)$.
3. n has a single successor n' with $V(n') = (y, A)$ such that $r(a, y) \in \mathcal{A}$ and $\mathcal{O} \models \exists r.A \sqsubseteq C$ (where r is a role name or an inverse role). Take $V'(n') = (wry, C)$.
4. $a \in \mathsf{N_I} \cap \mathrm{sig}(\mathcal{O})$ and there exists $x \in \mathsf{ind}(\mathcal{A})$ such that n has successors $n_1, \ldots, n_k$, $k \geq 1$ with $V(n_i) = (a_i, C_i)$ and $a_i = x$ or $a_i \in \mathsf{N_I} \cap \mathrm{sig}(\mathcal{O})$ for all i, and, defining $C'_i = C_i$ if $a_i = x$, and $C'_i = \exists u.(\{a_i\} \sqcap C_i)$ otherwise, we have $\mathcal{O} \models C'_1 \sqcap \ldots \sqcap C'_k \sqsubseteq \exists u.(\{a\} \sqcap C)$. Take $V'(n_i) = (w, C_i)$ if $x_i = x$, and $V'(n_i) = (a_i, C_i)$ if $a_i \in \mathsf{N_I} \cap \mathrm{sig}(\mathcal{O})$.

Then (T, V') is a derivation tree for $A_0(x_0)$ in $\mathcal{A}^*$ w.r.t. $\mathcal{O}$. □

Deciding interpolant existence for $\mathcal{ELIO}_u$ and $\mathcal{ELI}_u$. The next lemma provides a characterization of $\mathcal{ELIO}_u$ interpolant existence in terms of canonical models.

Lemma 38. *Let $\mathcal{O}_1, \mathcal{O}_2$ be $\mathcal{ELIO}_u$-ontologies in normal form, $A_1, A_2 \in \mathsf{N_C}$, and $\Sigma = \mathrm{sig}(\mathcal{O}_1, A_1) \cap \mathrm{sig}(\mathcal{O}_2, A_2)$. Then there exists an $\mathcal{ELIO}_u$-interpolant for $\mathcal{O}_1, \mathcal{O}_2, A_1, A_2$ if and only if there exists a pointed ABox $\mathcal{A}, x$ such that*

1. *$\mathcal{A}, x$ is a finite subset of $\mathcal{A}_{\mathcal{O}_1 \cup \mathcal{O}_2, A_1}, x_{A_1}$; and*
2. *$\mathcal{O}_1 \cup \mathcal{O}_2, \mathcal{A}_{|\Sigma} \models A_2(x)$.*

Proof. If D is an interpolant, then $\mathcal{A}_{\mathcal{O}_1 \cup \mathcal{O}_2, A_1} \models D(x_{A_1})$, and we obtain $\mathcal{A}$ by compactness. Conversely, suppose that we have $\mathcal{A}, x$ as above. Notice that $\mathcal{A}_{|\Sigma}$ is not necessarily tree-shaped modulo $\mathsf{N_I} \cap \Sigma$, and therefore does not necessarily directly correspond to a Σ concept. However, by Lemma 37, we also have $\mathcal{O}_1 \cup \mathcal{O}_2, \mathcal{A}^*_{|\Sigma} \models A_2(x)$, where $\mathcal{A}^*_{|\Sigma}$ is the undirected unfolding of $\mathcal{A}_{|\Sigma}$ modulo $\mathsf{N_I} \cap \Sigma$. By compactness, there exists a subset $\mathcal{A}'$ of $\mathcal{A}^*_{|\Sigma}$, that is a tree-shaped ABox modulo $\mathsf{N_I} \cap \Sigma$, such that $\mathcal{O}_1 \cup \mathcal{O}_2, \mathcal{A}' \models A_2(x)$. The associated Σ-concept is then an interpolant. □

Note that $\mathcal{A}_{\mathcal{I}_{\mathcal{O}_1 \cup \mathcal{O}_2, A_1}}$ and thus $\mathcal{A}$ are tree-shaped modulo $\mathsf{N_I} \cap \mathrm{sig}(\mathcal{O}_1, \mathcal{O}_2)$. We represent such ABoxes by trees over the alphabet

$$\Theta_{\mathcal{O}_1,\mathcal{O}_2,A_1,A_2} = 2^{\mathsf{N_C} \cap \Gamma \cup \{\{a\} \mid a \in \mathsf{N_I} \cap \Gamma\} \cup \{r, r^- \mid r \in \mathsf{N_R} \cap \Gamma\} \cup \{\exists r.\{a\}, \exists r^-.\{a\} \mid r \in \mathsf{N_R} \cap \Gamma \wedge a \in \mathsf{N_I} \cap \Gamma\}},$$

where $\Gamma = \mathrm{sig}(\mathcal{O}_1) \cup \mathrm{sig}(\mathcal{O}_2) \cup \{A_1, A_2\}$. Intuitively, the nodes of the tree correspond to the individuals of the ABox; labels $A, \{a\}, \exists r.\{a\}, \exists r^-.\{a\}$ indicate concepts that hold at the current node, while labels r or r^- are used to indicate which roles (if any) connect a node to its parent. This is defined more formally below.

A *tree* is a non-empty set $T \subseteq (\mathbb{N} \setminus \{0\})^*$ closed under prefixes and such that $n \cdot (i+1) \in T$ implies $n \cdot i \in T$ for $1 \leq i < k$. It is *k-ary* if $T \subseteq \{1, \ldots, k\}^*$. The node ε is the *root* of T. As a convention, we take $n \cdot 0 = n$ and $(n \cdot i) \cdot -1 = n$. Note that $\varepsilon \cdot -1$ is undefined. Given an alphabet Θ, a *Θ-labeled tree* is a pair (T, L) consisting of a tree T and a node-labeling function $L : T \to \Theta$.

Let $\Theta_{\mathcal{O}_1,\mathcal{O}_2,A_1,A_2}$ be defined as above. We associate with every Θ-labeled tree (T, L) the following ABox, where x_a are fresh individual variables:

$$\begin{aligned}
\mathcal{A}_{(T,L)} = {} & \{\top(x) \mid x \in T\} \cup \\
& \{\{a\}(x_a) \mid a \text{ appears in the label of some } n \in T\} \cup \\
& \{A(x) \mid x \in T \wedge A \in L(x)\} \cup \\
& \{\{a\}(x) \mid x \in T \wedge \{a\} \in L(x)\} \cup \\
& \{r(x, x \cdot i) \mid x \cdot i \in T \wedge r \in \mathsf{N_R} \wedge r \in L(x \cdot i)\} \cup \\
& \{r(x \cdot i, x) \mid x \cdot i \in T \wedge r \in \mathsf{N_R} \wedge r^- \in L(x \cdot i)\} \cup \\
& \{r(x, x_a) \mid x \in T \wedge \exists r.\{a\} \in L(x)\} \cup \\
& \{r(x_a, x) \mid x \in T \wedge \exists r^-.\{a\} \in L(x)\}\,.
\end{aligned}$$

Notice that $\mathcal{A}_{(T,L)}$ is tree-shaped modulo $\mathsf{N_I} \cap \mathrm{sig}(\mathcal{O}_1, \mathcal{O}_2)$. Conversely, for every ABox $\mathcal{A}$ that is tree-shaped modulo $\mathsf{N_I} \cap \mathrm{sig}(\mathcal{O}_1, \mathcal{O}_2)$, there exists a (not necessarily unique) tree (T, L) such that $\mathcal{A} = \mathcal{A}_{(T,L)}$. In addition, the Abox $\mathcal{A}_{\mathcal{O}_1 \cup \mathcal{O}_2, A_1}$ can be represented by a k-ary tree, where $k = |\mathcal{O}_1| + |\mathcal{O}_2| + |\mathsf{N_I} \cap \mathrm{sig}(\mathcal{O}_1, \mathcal{O}_2)|$. Therefore, we focus from now on on k-ary trees.

Our decision procedure for interpolant existence is based tree automata (whose definition is recalled below). More precisely, our goal is to prove the following:

Lemma 39. *There exists an exponential size non-deterministic tree automaton $\mathfrak{A}$ such that:*

– *For every* $(T, L) \in L(\mathfrak{A})$*,* $\mathcal{A}_{(T,L)}$ *satisfies Conditions 1 and 2 from Lemma 38;*
– *For every* $\mathcal{A}$ *satisfying Conditions 1 and 2 from Lemma 38, there exists* $(T, L) \in L(\mathfrak{A})$ *such that* $\mathcal{A} = \mathcal{A}_{(T,L)}$*.*

Moreover, $\mathfrak{A}$ *can be computed from* $\mathcal{O}_1, \mathcal{O}_2, A_1, A_2$ *in exponential time.*

The existence of an interpolant then reduces to the non-emptiness of $\mathfrak{A}$, and can be decided in exponential time.

Tree Automata. A *non-deterministic tree automaton (NTA)* over finite k-ary trees is a tuple $\mathfrak{A} = (Q, \Theta, I, \Delta)$, where Q is a set of states, Θ is the input alphabet, $I \subseteq Q$ is the set of initial states, and $\Delta \subseteq Q \times \Theta \times \bigcup_{0 \leq \ell \leq k} Q^\ell$ is the transition relation. A *run* of an NTA $\mathfrak{A} = (Q, \Theta, q_0, \Delta)$ over a k-ary input (T, L) is a Q-labeled tree (T, r) such that for all $x \in T$ with children $y_1, \ldots, y_\ell$, $(r(w), L(w), r(y_1), \ldots, r(y_\ell)) \in \Delta$. It is *accepting* if $r(\varepsilon) \in I$. The language accepted by $\mathfrak{A}$, denoted $L(\mathfrak{A})$, is the set of all finite k-ary Θ-labeled trees over which $\mathfrak{A}$ has an accepting run.

A *two-way alternating tree automaton over finite k-ary trees (2ATA)* is a tuple $\mathfrak{A} = (Q, \Theta, q_0, \delta)$ where Q is a finite set of *states*, Θ is the *input alphabet*, $q_0 \in Q$ is the *initial state*, and δ is a *transition function*. The transition function δ maps every state q and input letter $\theta \in \Theta$ to a positive Boolean formula $\delta(q, \theta)$ over the truth constants true and false and *transition atoms* of the form $(i, q) \in [k] \times Q$, where $[k] = \{-1, 0, 1, \ldots, k\}$. The semantics is given in terms of *runs*. More precisely, let (T, L) be a finite k-ary Θ-labeled tree and $\mathfrak{A} = (Q, \Theta, q_0, \delta)$ a 2ATA. An *accepting run of* $\mathfrak{A}$ *over* (T, L) is a $(T \times Q)$-labeled tree (T_r, r) such that:

1. $r(\varepsilon) = (\varepsilon, q_0)$, and
2. for all $y \in T_r$ with $r(y) = (x, q)$, there is a subset $S \subseteq [k] \times Q$ such that $S \models \delta(q, L(x))$ and for every $(i, q') \in S$, there is some successor y' of y in T_r with $r(y) = (x \cdot i, q')$.

The language accepted by $\mathfrak{A}$, denoted $L(\mathfrak{A})$, is the set of all finite k-ary Θ-labeled trees (T, L) for which there is an accepting run.

Every 2ATA $\mathfrak{A}$ can be turned into an equivalent NTA whose number of states is exponential in the number of states of $\mathfrak{A}$, and non-emptiness of a 2ATA can be decided in ExpTime [46].

In order to prove Lemma 39, we describe an NTA $\mathfrak{A}_1$ with exponentially many states for Condition 1, and a 2-ATA $\mathfrak{A}_2$ with polynomially many states for Condition 2. The automaton $\mathfrak{A}$ is then defined as the intersection of $\mathfrak{A}_1$ and the NTA equivalent to $\mathfrak{A}_2$.

Definition of $\mathfrak{A}_1$*.* We represent the canonical model for $\mathcal{O}_1 \cup \mathcal{O}_2$ and A_1 by a tree with τ_{A_1} at the root of the tree, all $\tau_{A_1,a}$ inserted directly below the root, and other elements $\tau_0 r_1 \tau_1 \cdots r_n \tau_n$ below $\tau_0 r_1 \tau_1 \cdots r_{n-1} \tau_{n-1}$. We want $\mathfrak{A}_1$ to accept all finite subsets of $\mathcal{A}_{\mathcal{O}_1 \cup \mathcal{O}_2, A_1}$, so trees obtained from this one by keeping a finite subset of nodes, and possibly removing some concepts and relations from the labels. We associate with every $\alpha \in \Theta_{\mathcal{O}_1,\mathcal{O}_2,A_1,A_2}$ a set of concepts $\tau_\alpha = (\mathsf{N_C} \cap \alpha) \cup (\mathsf{N_I} \cap \alpha) \cup \{\exists u.A \mid \mathcal{O}_1 \cup \mathcal{O}_2 \models A_1 \sqsubseteq \exists u.A\}$. We take $\mathfrak{A}_1 = (Q_1, \Theta_{\mathcal{O}_1,\mathcal{O}_2,A_1,A_2}, I_1, \Delta_1)$, where

– $Q_1 \subseteq \{q_0\} \cup \Theta_{\mathcal{O}_1,\mathcal{O}_2,A_1,A_2}$ contains $\{q_0\}$ and all $\alpha \in \Theta_{\mathcal{O}_1,\mathcal{O}_2,A_1,A_2}$ such that τ_α is an $\mathcal{O}_1 \cup \mathcal{O}_2$-type that occurs in $\mathcal{A}_{\mathcal{O}_1\cup\mathcal{O}_2,A_1}$;
– $I_1 = \{q_0\}$.
– $(q, \alpha, q_1, \dots, q_\ell) \in \Delta$ if the following conditions are satisfied, for all $1 \leq i \leq \ell$:
  • $q = q_0$ or $\alpha \subseteq q$;
  • If $r \in q_i$ then $\tau_q \rightsquigarrow_r \tau_{q_i}$;
  • If $q_i \cap \mathsf{N_R} = \emptyset$ then either $\tau_q \rightsquigarrow_u \tau_{q_i}$, or $q = q_0$ and q_i contains some nominal $\{a\}$.

Lemma 40. *For all $(T, L) \in L(\mathfrak{A}_1)$, $\mathcal{A}_{(T,L)}, \varepsilon$ is (isomorphic to) a subset of $\mathcal{A}_{\mathcal{O}_1\cup\mathcal{O}_2,A_1}, x_{A_1}$. Conversely, for every finite subset $\mathcal{A}$ of $\mathcal{A}_{\mathcal{O}_1\cup\mathcal{O}_2,A_1}$, there exists $(T, L) \in L(\mathfrak{A}_1)$ such that $\mathcal{A} = \mathcal{A}_{(T,L)}$.*

Definition of $\mathfrak{A}_2$. The construction of $\mathfrak{A}_2 = (Q_2, \Theta, q_{A_2}, \delta_2)$ relies on derivation trees. Intuitively, runs of $\mathfrak{A}_2$ on some (T, L) correspond to derivation trees for $A_0(\varepsilon)$ in $\mathcal{A}_{(T,L)}$ w.r.t. $\mathcal{O}_1 \cup \mathcal{O}_2$. The states of $\mathfrak{A}_2$ are

$$\begin{aligned} Q_2 = {} & \{q_A \mid A \in \mathsf{N_C} \cap \Gamma\} \cup \{q_{\{a\}} \mid a \in \mathsf{N_I} \cap \Gamma\} \cup \\ & \{q_{\exists r.A}, q_{\exists r^-.A} \mid r \in \mathsf{N_R} \cap \Gamma, A \in \mathsf{N_C} \cap \Gamma\} \cup \{q_{\exists u.A} \mid A \in \mathsf{N_C} \cap \Gamma\} \cup \\ & \{q_{\exists u.(\{a\}\sqcap A)} \mid a \in \mathsf{N_I} \cap \Gamma, A \in \mathsf{N_C} \cap \Gamma\} \cup \{q_{\exists u.(\{a\}\sqcap\{b\})} \mid a, b \in \mathsf{N_I} \cap \Gamma\} \cup \\ & \{q_r, q_{r^-} \mid r \in \mathsf{N_R} \cap \Sigma\} \cup \{q_{\exists r.\{a\}}, q_{\exists r^-.\{a\}} \mid a \in \mathsf{N_I} \cap \Gamma, r \in \mathsf{N_R} \cap \Gamma\}, \end{aligned}$$

where $\Gamma = \mathrm{sig}(\mathcal{O}_1) \cup \mathrm{sig}(\mathcal{O}_2)$. Intuitively, state q_C is used to check that C is entailed at the current node. States q_r and $q_{\exists r.\{a\}}$ are used to check the label of the current node. The initial state is q_{A_2}, as we are trying to construct a derivation tree for A_2 at the root.

Let us now define the transition relation. For a label $\alpha \in \Theta_{\mathcal{O}_1,\mathcal{O}_2,A_1,A_2}$, we denote by α_Σ the subset of α consisting of Σ concepts and roles. We let

$$\delta_2(q_r, \alpha) = \begin{cases} \mathsf{true} & \text{if } r \in \alpha_\Sigma \\ \mathsf{false} & \text{if } r \notin \alpha_\Sigma \end{cases} \qquad \delta_2(q_{\exists r.\{a\}}, \alpha) = \begin{cases} \mathsf{true} & \text{if } \exists r.\{a\} \in \alpha \\ \mathsf{false} & \text{if } \exists r.\{a\} \notin \alpha \end{cases}$$

$$\begin{aligned} \delta_2(q_{\exists r.A}, \alpha) = {} & (0, q_{r^-}) \wedge (-1, q_A) \vee \bigvee_{1 \leq i \leq k} (i, q_A) \wedge (i, q_r) \vee \\ & \bigvee_{a \in \mathsf{N_I} \cap \Gamma} (0, q_{\exists r.\{a\}}) \wedge (0, q_{\exists u.(\{a\}\sqcap A)}) \end{aligned}$$

$$\delta_2(q_{\exists u.C}, \alpha) = \bigvee_{\mathcal{O}_1\cup\mathcal{O}_2 \models C_1 \sqcap \dots \sqcap C_n \models \exists u.C} \ \bigwedge_{1 \leq i \leq n} (0, q_{C_i}) \vee \bigvee_{i \in \{-1,1,\dots,k\}} (i, q_{\exists u.C})$$

For $C = \{a\}$ or $C = A$, $\delta(q_C, \alpha) = \mathsf{true}$ if $C \in \alpha_\Sigma$ or $\mathcal{O}_1 \cup \mathcal{O}_2 \models \top \sqsubseteq C$, and otherwise,

$$\delta_2(q_C, \alpha) = \bigvee_{\mathcal{O}_1\cup\mathcal{O}_2 \models C_1 \sqcap \dots \sqcap C_n \models C} \ \bigwedge_{1 \leq i \leq n} (0, q_{C_i}) \vee \bigvee_{\mathcal{O}_1\cup\mathcal{O}_2 \models \exists r.B \sqsubseteq C} (0, q_{\exists r.B}).$$

Intuitively, the transitions from q_r and $q_{\exists r.\{a\}}$ check the label of the current node. Transitions from $q_{\exists r.A}$ are used to check if condition 2 can be applied in a derivation tree, that is, if from the current node x, there exists another node y such that $r(x,y)$ and $A(y)$ are in $\mathcal{A}_{(T,L)}$; the node y can be either the parent node, a child node, or a node corresponding to some nominal a. Transitions from $q_{\exists u.C}$ correspond to an application of condition 1 (if the first disjunct is used and C is a concept name or nominal), 4 (if the first disjunct is used and $C = \exists u.(\{a\} \sqcap A)$) or 3 (second disjunct). Transitions from other states q_C correspond to an application of condition 1 (first disjunct) or 2 (second disjunct).

Lemma 41. *For all concept names A, $\mathfrak{A}_2$ has a run starting from state q_A on (T,L) if and only if there exists a derivation tree for $A(\varepsilon)$ in $\mathcal{A}_{(T,L)}$ w.r.t. $\mathcal{O}_1 \cup \mathcal{O}_2$.*

Deciding interpolant existence for $\mathcal{ELIO}$ and $\mathcal{ELI}$. The decision procedure for $\mathcal{ELIO}_u$ can easily be adapted for $\mathcal{ELIO}$ and $\mathcal{ELI}$, where the universal role u cannot be used in the interpolant. The only modification needed in Lemma 38 is to require that $\mathcal{A}_{|\Sigma}$ is connected; the latter can easily be verified by a tree automaton checking that for every node x apart from the root and nominals, if x has a non-empty Σ-label then it contains a role $r \in \Sigma$, and that for every $a \in \mathsf{N_I} \cap \Sigma$, if there is a node whose label contains $\{a\}$ in the tree then there is one containing $\exists r.\{a\}$ for some r.

Size of interpolants. The NTA $\mathfrak{A}$ has exponentially many states, and thus if its language is non-empty it contains a tree (T,L) of at most exponential depth. Therefore if an ABox $\mathcal{A}$ as in Lemma 38 exists, then there is one of double exponential size. For $\mathcal{ELI}_u$ and $\mathcal{ELI}$, $\mathcal{A}$ is tree-shaped and thus $\mathcal{A}_{|\Sigma}$ corresponds directly to an interpolant of double exponential size. This is not the case for $\mathcal{ELIO}_u$, as the ABox might be tree-shaped modulo $\mathsf{N_I} \cap \mathrm{sig}(\mathcal{O}_1, \mathcal{O}_2)$ but not modulo $\mathsf{N_I} \cap \Sigma$. In that case, by Lemma 37 we have a double exponential depth derivation tree for $A_2(x)$ in $\mathcal{A}^*_{|\Sigma}$, which means it uses an at most triple exponential number of elements in $\mathcal{A}^*_{|\Sigma}$, and we obtain an interpolant of triple exponential size.

We construct an $\mathcal{ELI}$-ontology $\mathcal{O}$, signature Σ, and concept name A such that the smallest explicit $\mathcal{ELI}(\Sigma)$-definition of A under $\mathcal{O}$ is of double exponential size in $||\mathcal{O}||$. $\mathcal{O}$ is a variant of ontologies constructed in [35,41] and defined as follows. It contains $\top \sqsubseteq \exists r.\top \sqcap \exists s.\top$,

$$
\begin{aligned}
A &\sqsubseteq M \sqcap \overline{X_0} \sqcap \ldots \sqcap \overline{X_n} && \\
\exists \sigma^-.(\overline{X_i} \sqcap X_0 \sqcap \ldots \sqcap X_{i-1}) &\sqsubseteq X_i && \sigma \in \{r,s\}, i \leq n \\
\exists \sigma^-.(X_i \sqcap X_0 \sqcap \ldots \sqcap X_{i-1}) &\sqsubseteq \overline{X_i} && \sigma \in \{r,s\}, i \leq n \\
\exists \sigma^-.(\overline{X_i} \sqcap \overline{X_j}) &\sqsubseteq \overline{X_i} && \sigma \in \{r,s\}, j < i \leq n \\
\exists \sigma^-.(X_i \sqcap \overline{X_j}) &\sqsubseteq X_i && \sigma \in \{r,s\}, j < i \leq n \\
X_0 \sqcap \ldots \sqcap X_n &\sqsubseteq L &&
\end{aligned}
$$

and

$$L \sqsubseteq B, \quad \exists r.B \sqcap \exists s.B \sqsubseteq B, \quad B \sqcap M \sqsubseteq A.$$

Let $\Sigma = \{M, r, s, L\}$. Note that A triggers a marker M and a binary tree of depth 2^n using counter concept names $X_0, \ldots, X_n$ and $\overline{X_0}, \ldots, \overline{X_n}$. A concept name L is made true at the leafs. Conversely, if L is true at the leafs of a binary tree of depth 2^n then B is true at all nodes of the tree and A is entailed by M and B at its root. Define inductively

$$C_0 = L, \quad C_{k+1} = \exists r.C_k \sqcap \exists s.C_k, \quad C = C_{2^n} \sqcap M.$$

Then C is the smallest explicit $\mathcal{ELI}(\Sigma)$-definition of A under $\mathcal{O}$.

8 Expressive Horn-DLs and Horn-FO

We first observe that all our results on the complexity of interpolant and explicit definition existence also hold for more expressive Horn-DLs such as Horn-$\mathcal{ALC}$, Horn-$\mathcal{ALCI}$, and Horn-$\mathcal{ALCIO}_u$, provided that the interpolant/explicit definition is from the corresponding $\mathcal{ELI}$-fragment. This is yet another indicator that Horn-DLs behave rather differently from DLs with all Boolean operators, where for example the existence of an explicit $\mathcal{EL}(\Sigma)$-concept definition of a concept name A under an $\mathcal{ALC}$-ontology is undecidable [3].

If, however, interpolants/explicit definitions from the Horn-DL itself are admitted, the situation is different. Observe that in the counterexamples disproving the PBDP of $\mathcal{EL}_u$ and $\mathcal{ELI}$ the concept names that are implicitly but not explicitly definable (Theorem 15 for $\mathcal{EL}_u$ and Theorem 17 for $\mathcal{ELI}$) are in fact explicitly definable in Horn-$\mathcal{ALC}$. Thus, we next consider the question of whether Horn-DLs such as Horn-$\mathcal{ALC}$ and Horn-$\mathcal{ALCI}$ enjoy the CIP/PBDP for interpolants/explicit definitions in Horn-$\mathcal{ALC}$ and Horn-$\mathcal{ALCI}$, respectively. We show that this is not the case in a very strong sense by providing a Horn-$\mathcal{ALC}$-ontology $\mathcal{O}$, a signature Σ, and a concept name A such that A is implicitly Σ-definable under $\mathcal{O}$ but does not have an explicit Σ-definition even in the Horn fragment of the guarded fragment of FO introduced in [24] (and therefore not in Horn-$\mathcal{ALC}$ and Horn-$\mathcal{ALCI}$ either). The decidability and complexity of interpolant and explicit definition existence in Horn-$\mathcal{ALC}$ and Horn-$\mathcal{ALCI}$ are therefore wide open, but the games developed in [24] should provide the appropriate tool for attacking these problems.

Finally, we address the question whether for languages such as Horn-$\mathcal{ALCI}$ there exists any Horn language at all extending them in which one is guaranteed to find interpolants/explicit definitions. The answer is positive since the Horn fragment of FO contains all first-order Horn-DLs and enjoys the CIP/PBDP. It remains open, however, whether there exists a smaller such language, in particular a decidable one.

We start by formulating the complexity results if interpolants/explicit definitions in $\mathcal{ELI}$ and its extensions are considered. We remind the reader that by adding $A \sqsubseteq \bot$ to the CIs used in ontologies in normal form, one can show the following variant of Lemma 4.

Lemma 42. *Let $\mathcal{L}$ be a DL from Horn-$\mathcal{ALCI}$, Horn-$\mathcal{ALCIO}$, Horn-$\mathcal{ALCI}_u$, Horn-$\mathcal{ALCIO}_u$. Let $\mathcal{O}$ be an $\mathcal{L}$-ontology. Then one can construct in polynomial time an $\mathcal{L}$-ontology $\mathcal{O}'$ in normal form (including CIs of the form $A \sqsubseteq \bot$) such that $\mathcal{O}'$ is a model conservative extension of $\mathcal{O}$.*

We thus obtain the following complexity result from Remark 22, Lemma 21, and Theorem 34.

Theorem 43. *Let $(\mathcal{L}, \mathcal{L}')$ be (Horn-$\mathcal{ALCI}, \mathcal{ELI}$), (Horn-$\mathcal{ALCIO}, \mathcal{ELIO}$), (Horn-$\mathcal{ALCI}_u, \mathcal{ELI}_u$), or (Horn-$\mathcal{ALCIO}_u, \mathcal{ELIO}_u$). Then*

- *deciding the existence of an $\mathcal{L}'$-interpolant for an $\mathcal{L}'$-CI $C \sqsubseteq D$ under $\mathcal{L}$-ontologies $\mathcal{O}_1, \mathcal{O}_2$ is* ExpTime*-complete;*
- *deciding the existence of an explicit $\mathcal{L}'(\Sigma)$-definition of a concept name A under an $\mathcal{L}$-ontology $\mathcal{O}$ is* ExpTime*-complete.*

We now disprove the CIP and PBDP for Horn-$\mathcal{ALC}$, Horn-$\mathcal{ALCI}$ and their extensions with the universal role if Horn-concepts are admitted as interpolants or explicit definitions.

Theorem 44. *The exists a Horn-$\mathcal{ALC}$-ontology $\mathcal{O}$, a signature Σ, and a concept name A such that A is implicitly definable from Σ under $\mathcal{O}$ but not explicitly Horn-$\mathcal{ALCI}_u(\Sigma)$-definable.*

Proof. We modify the ontology from the proof of Theorem 17. In detail, let $\mathcal{O}$ contain the following CIs:

$$
\begin{aligned}
A &\sqsubseteq B\\
B &\sqsubseteq \forall r.F\\
F &\sqsubseteq \exists r_1.D_1 \sqcap \exists r_2.D_2 \sqcap \exists r_1.M \sqcap \exists r_2.M\\
A &\sqsubseteq \forall r.((F \sqcap \exists r_1.(D_1 \sqcap M) \sqcap \exists r_2.(D_2 \sqcap M)) \rightarrow E)\\
B &\sqsubseteq \exists r.C\\
C &\sqsubseteq F \sqcap \forall r_1.D_1 \sqcap \forall r_2.D_2\\
B \sqcap \exists r.(C \sqcap E) &\sqsubseteq A
\end{aligned}
$$

and let $\Sigma = \{B, D_1, D_2, E, r, r_1, r_2\}$. We note that, intuitively, the third and fourth CI should be read as

$$
\begin{aligned}
F &\sqsubseteq \exists r_1.D_1 \sqcap \exists r_2.D_2\\
A &\sqsubseteq \forall r.((F \sqcap \forall r_1.D_1 \sqcap \forall r_2.D_2) \rightarrow E)
\end{aligned}
$$

and the concept name M is introduced to achieve this in a projective way as the latter CI is not in Horn-$\mathcal{ALCI}$.

We first observe that A is implicitly definable from Σ under $\mathcal{O}$ since

$$\mathcal{O} \models A \equiv B \sqcap \forall r.(\forall r_1.D_1 \sqcap \forall r_2.D_2 \rightarrow E).$$

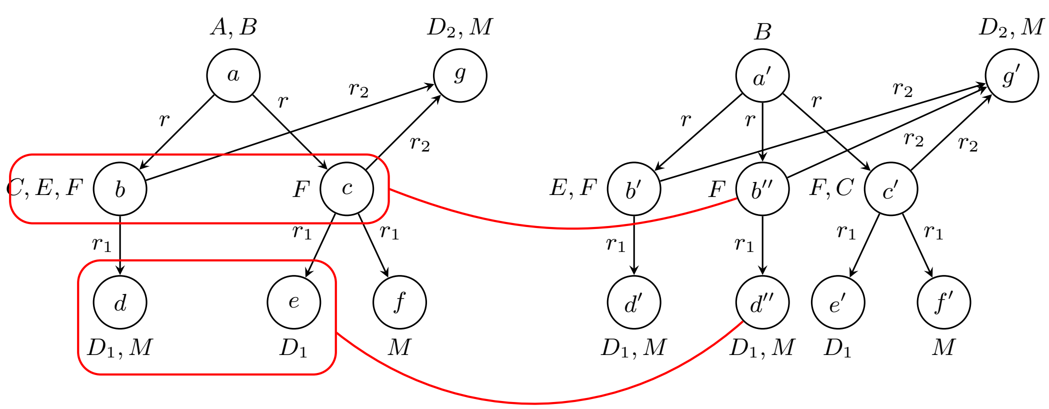
We next show that A is not explicitly Horn-$\mathcal{ALCI}_u(\Sigma)$-definable under $\mathcal{O}$. To this end consider the interpretations $\mathcal{I}$ and $\mathcal{I}'$ below. Both $\mathcal{I}$ and $\mathcal{I}'$ are models of $\mathcal{O}$, $a \in A^{\mathcal{I}}$, $a' \notin A^{\mathcal{I}'}$, but $a \in F^{\mathcal{I}}$ implies $a' \in F^{\mathcal{I}'}$ holds for every Horn-$\mathcal{ALCI}_u(\Sigma)$-concept F, and the claim follows. The latter can be proved by observing that there exists a Horn-$\mathcal{ALCI}_u(\Sigma)$-simulation between $\mathcal{I}$ and $\mathcal{I}'$ [24] containing $(\{a\}, a)$. The definition of these simulations is as follows. For any two sets X and Y and a binary relation R, we set

- $XR^{\uparrow}Y$ if for all $x \in X$ there exists $y \in Y$ with $(x, y) \in R$;
- $XR^{\downarrow}Y$ if for all $y \in Y$ there exists $x \in X$ with $(x, y) \in R$.

A relation $Z \subseteq \mathcal{P}(\Delta^{\mathcal{I}}) \times \Delta^{\mathcal{I}'}$ is a *Horn-$\mathcal{ALCI}(\Sigma)$-simulation* between $\mathcal{I}$ and $\mathcal{I}'$ if $(X, b) \in Z$ implies $X \neq \emptyset$ and the following hold:

- for any $A \in \Sigma$, if $(X, b) \in Z$ and $X \subseteq A^{\mathcal{I}}$, then $b \in A^{\mathcal{I}'}$;
- for any role r in Σ, if $(X, b) \in Z$ and $Xr^{\mathcal{I}\uparrow}Y$, then there exist $Y' \subseteq Y$ and $b' \in \Delta^{\mathcal{I}'}$ with $(b, b') \in r^{\mathcal{I}'}$ and $(Y', b') \in Z$;
- for any role r in Σ, if $(X, b) \in Z$ and $(b, b') \in r^{\mathcal{I}'}$, then there is $Y \subseteq \Delta^{\mathcal{I}}$ with $Xr^{\mathcal{I}\downarrow}Y$ and $(Y, b') \in Z$;
- if $(X, b) \in Z$, then $\mathcal{I}', b \preceq_{\mathcal{ELI},\Sigma} \mathcal{I}, a$ for every $a \in X$ (where $\preceq_{\mathcal{ELI},\Sigma}$ indicates that we have a simulation that does not only respect role names in Σ but also the inverse of role names in Σ).

We write $\mathcal{I}, X \preceq_{horn,\Sigma} \mathcal{I}', b$ if there exists a Horn-$\mathcal{ALCI}(\Sigma)$-simulation Z between $\mathcal{I}$ and $\mathcal{I}'$ such that $(X, b) \in Z$. It is shown in [24] that if $\mathcal{I}, X \preceq_{horn,\Sigma} \mathcal{I}', b$, then all Horn-$\mathcal{ALCI}(\Sigma)$-concepts true in all nodes in X are also true in b. Now one can show that the relation Z containing all pairs $(\{x\}, x')$, $(\{b, c\}, b'')$, and $(\{d, e\}, d'')$ is a Horn-$\mathcal{ALCI}(\Sigma)$-simulation between $\mathcal{I}$ and $\mathcal{I}'$, as required.



□

We note that the relation Z is even a hornGF(Σ)-simulation in the sense of [24] and that therefore one also cannot find an explicit definition of A in the Horn fragment of the guarded fragment of FO.

We finally address the question whether there exists at all a Horn language that contains the Horn-DLs considered in this paper and which enjoys the CIP and PBDP. This is indeed the case for Horn-FO, the standard Horn fragment of first-order logic. According to Exercise 6.2.6 in [16] Horn-FO has the following property.

Theorem 45. *Let φ, ψ be sentences in Horn-FO such that $\varphi \wedge \psi$ is not satisfiable. Then there exists a sentence χ in Horn-FO such that $sig(\chi) \subseteq sig(\varphi) \cap sig(\psi)$, $\varphi \models \chi$, and $\chi \wedge \psi$ is not satisfiable.*

We directly obtain the following interpolation result.

Theorem 46. *Let $\mathcal{O}_1, \mathcal{O}_2$ be Horn-$\mathcal{ALCIO}_u$-ontologies and let C_1, C_2 be Horn-$\mathcal{ALCIO}_u$-concepts such that $\mathcal{O}_1 \cup \mathcal{O}_2 \models C_1 \sqsubseteq C_2$. Then there exists a formula $\chi(x)$ in Horn-FO such that*

- *$sig(\chi) \subseteq sig(\mathcal{O}_1, C_1) \cap sig(\mathcal{O}_2, C_2)$;*
- *$\mathcal{O}_1 \models \forall x(C_1(x) \rightarrow \chi(x))$;*
- *$\mathcal{O}_2 \models \forall x(\chi(x) \rightarrow C_2(x))$.*

Proof. Take a fresh unary relation symbol $A(x)$ and a fresh individual name c. Let φ be the conjunction of all sentences in $\mathcal{O}_1 \cup \{C_1(c)\}$ and let ψ be the conjunction of all sentences in $\mathcal{O}_2 \cup \{\forall x(C_2(x) \leftrightarrow A(x)), \neg A(c)\}$. Then φ and ψ are both equivalent to sentences in Horn-FO. By definition $\varphi \wedge \psi$ is not satisfiable. Thus there exists a Horn-FO sentence χ using only c and symbols in $\mathrm{sig}(\mathcal{O}_1, C_1) \cap \mathrm{sig}(\mathcal{O}_2, C_2)$ such that $\varphi \models \chi$ and $\chi \wedge \psi$ is not satisfiable. Thus:

- $\mathcal{O}_1 \models C_1(c) \rightarrow \chi$;
- $\mathcal{O}_2 \cup \{\forall x(C_2(x) \leftrightarrow A(x))\} \models \chi \rightarrow A(c)$.

Replace c by x in $\chi, C_1(c)$, and $A(c)$. Then

- $\mathcal{O}_1 \models \forall x(C_1(x) \rightarrow \chi(x))$;
- $\mathcal{O}_2 \models \forall x(\chi(x) \rightarrow C_2(x))$,

as required. □

Applied to Horn-$\mathcal{ALCI}$ ontologies and concepts we thus always obtain an interpolant in Horn-FO and an interpolant in $\mathcal{ALCI}$ (since $\mathcal{ALCI}$ enjoys the CIP [15]). Does there exist an interpolant in the intersection of Horn-FO and $\mathcal{ALCI}$? Is it possible to give an informative syntactic description of that intersection?

9 Discussion

A few Horn-DLs remain to be investigated regarding the complexity of interpolant and explicit definition existence and the size of interpolants. For example, the extension of $\mathcal{ELI}$ with various types of role inclusions and extensions of $\mathcal{EL}$ and $\mathcal{ELI}$ with functional roles. Deciding the existence of interpolants and explicit definitions that are in Horn-$\mathcal{ALC}$ and its extensions is also an interesting open problem. The development of practical algorithms for computing interpolants and explicit definitions is another challenge. Here the characterization of interpolants and explicit definitions as finite subsets of the unfolding of the Σ-reduct of the canonical model should be an excellent starting point.

References

1. Areces, C., Blackburn, P., Marx, M.: Hybrid logics: Characterization, interpolation and complexity. J. Symb. Log. **66**(3), 977–1010 (2001)
2. Artale, A., Jung, J.C., Mazzullo, A., Ozaki, A., Wolter, F.: Living without Beth and Craig: Explicit definitions and interpolants in description logics with nominals and role hierarchies. In: Proc. of AAAI (2021)
3. Artale, A., Mazzullo, A., Ozaki, A., Wolter, F.: On free description logics with definite descriptions. In: Proc. of KR (2021)
4. Baader, F., Bienvenu, M., Lutz, C., Wolter, F.: Query and predicate emptiness in ontology-based data access. J. Artif. Intell. Res. **56**, 1–59 (2016)
5. Baader, F., Brandt, S., Lutz, C.: Pushing the $\mathcal{EL}$ envelope. In: Proc. of IJCAI. pp. 364–369 (2005)
6. Baader, F., Horrocks, I., Lutz, C., Sattler, U.: An Introduction to Description Logic. Cambridge University Press (2017)
7. Bárány, V., Benedikt, M., ten Cate, B.: Some model theory of guarded negation. J. Symb. Log. **83**(4), 1307–1344 (2018)
8. Benedikt, M., ten Cate, B., Vanden Boom, M.: Interpolation with decidable fixpoint logics. In: Proc. of LICS. pp. 378–389. IEEE Computer Society (2015)
9. Benedikt, M., ten Cate, B., Vanden Boom, M.: Effective interpolation and preservation in guarded logics. ACM Trans. Comput. Log. **17**(2), 8:1–8:46 (2016)
10. Benedikt, M., Leblay, J., ten Cate, B., Tsamoura, E.: Generating Plans from Proofs: The Interpolation-based Approach to Query Reformulation. Synthesis Lectures on Data Management, Morgan & Claypool Publishers (2016)
11. Bienvenu, M., Lutz, C., Wolter, F.: First-order rewritability of atomic queries in Horn description logics. In: Proc. of IJCAI (2013)
12. Borgida, A., Toman, D., Weddell, G.E.: On referring expressions in query answering over first order knowledge bases. In: Proc. of KR. pp. 319–328 (2016)
13. ten Cate, B.: Interpolation for extended modal languages. J. Symb. Log. **70**(1), 223–234 (2005)
14. ten Cate, B., Conradie, W., Marx, M., Venema, Y.: Definitorially complete description logics. In: Proc. of KR. pp. 79–89. AAAI Press (2006)
15. ten Cate, B., Franconi, E., Seylan, İ.: Beth definability in expressive description logics. J. Artif. Intell. Res. **48**, 347–414 (2013)
16. Chang, C., Keisler, H.J.: Model Theory. Elsevier (1998)
17. Cimatti, A., Griggio, A., Sebastiani, R.: Interpolant generation for UTVPI. In: Proc. of CADE. pp. 167–182. Springer (2009)
18. Goel, A., Krstic, S., Tinelli, C.: Ground interpolation for combined theories. In: Proc. of CADE. pp. 183–198. Springer (2009)
19. Hoogland, E., Marx, M.: Interpolation and definability in guarded fragments. Stud. Log. **70**(3), 373–409 (2002)

20. Hoogland, E., Marx, M., Otto, M.: Beth definability for the guarded fragment. In: Proc. of LPAR. pp. 273–285. Springer (1999)
21. Hustadt, U., Motik, B., Sattler, U.: Data complexity of reasoning in very expressive description logics. In: IJCAI. pp. 466–471 (2005)
22. Iemhoff, R.: Uniform interpolation and the existence of sequent calculi. Annals of Pure and Applied Logic **170**(11), 102711 (2019)
23. Jiménez-Ruiz, E., Payne, T.R., Solimando, A., Tamma, V.A.M.: Limiting logical violations in ontology alignnment through negotiation. In: Proc. of KR. pp. 217–226. AAAI Press (2016)
24. Jung, J.C., Papacchini, F., Wolter, F., Zakharyaschev, M.: Model comparison games for horn description logics. In: Proc. of LICS. pp. 1–14. IEEE (2019)
25. Jung, J.C., Wolter, F.: Living without Beth and Craig: Definitions and interpolants in the guarded and two-variable fragments. In: Proc. of LICS (2021)
26. Kazakov, Y., Krötzsch, M., Simancik, F.: The incredible ELK - from polynomial procedures to efficient reasoning with $\mathcal{EL}$ ontologies. J. Autom. Reason. **53**(1), 1–61 (2014)
27. Konev, B., Lutz, C., Walther, D., Wolter, F.: Formal properties of modularisation. In: Modular Ontologies, Lecture Notes in Computer Science, vol. 5445, pp. 25–66. Springer (2009)
28. Konev, B., Lutz, C., Ponomaryov, D.K., Wolter, F.: Decomposing description logic ontologies. In: Proc. of KR. AAAI Press (2010)
29. Koopmann, P., Schmidt, R.A.: Uniform interpolation and forgetting for ALC ontologies with aboxes. In: Proc. of AAAI. pp. 175–181. AAAI Press (2015)
30. Kowalski, T., Metcalfe, G.: Uniform interpolation and coherence. Annals of pure and applied logic **170**(7), 825–841 (2019)
31. Lutz, C., Piro, R., Wolter, F.: Description logic tboxes: Model-theoretic characterizations and rewritability. In: Walsh, T. (ed.) Proc. of IJCAI. pp. 983–988. IJCAI/AAAI (2011)
32. Lutz, C., Seylan, I., Wolter, F.: An automata-theoretic approach to uniform interpolation and approximation in the description logic EL. In: Proc. of KR. AAAI Press (2012)
33. Lutz, C., Seylan, I., Wolter, F.: Ontology-based data access with closed predicates is inherently intractable(sometimes). In: Proc. of IJCAI. pp. 1024–1030 (2013)
34. Lutz, C., Seylan, I., Wolter, F.: The data complexity of ontology-mediated queries with closed predicates. Logical Methods in Computer Science **15**(3) (2019)
35. Lutz, C., Wolter, F.: Deciding inseparability and conservative extensions in the description logic EL. J. Symb. Comput. **45**(2), 194–228 (2010)
36. Lutz, C., Wolter, F.: Foundations for uniform interpolation and forgetting in expressive description logics. In: Proc. of IJCAI. pp. 989–995. IJCAI/AAAI (2011)
37. Lutz, C., Wolter, F.: Non-uniform data complexity of query answering in description logics. In: Brewka, G., Eiter, T., McIlraith, S.A. (eds.) Proc. of KR (2012)
38. Maksimova, L., Gabbay, D.: Interpolation and Definability in Modal and Intuitionistic Logics. Clarendon Press (2005)
39. Marx, M.: Queries determined by views: Pack your views. In: Proc. of PODS. p. 2330. ACM (2007)
40. McMillan, K.L.: Interpolation and sat-based model checking. In: Proc. of CAV. pp. 1–13. Springer (2003)
41. Nikitina, N., Rudolph, S.: (Non-)succinctness of uniform interpolants of general terminologies in the description logic EL. Artif. Intell. **215**, 120–140 (2014)
42. Place, T., Zeitoun, M.: Separating regular languages with first-order logic. Log. Methods Comput. Sci. **12**(1) (2016)

43. Seylan, I., Franconi, E., de Bruijn, J.: Effective query rewriting with ontologies over dboxes. In: Proc. of IJCAI. pp. 923–925 (2009)
44. Sofronie-Stokkermans, V.: Interpolation in local theory extensions. Log. Methods Comput. Sci. **4**(4) (2008). https://doi.org/10.2168/LMCS-4(4:1)2008, `https://doi.org/10.2168/LMCS-4(4:1)2008`
45. Toman, D., Weddell, G.E.: Fundamentals of Physical Design and Query Compilation. Synthesis Lectures on Data Management, Morgan & Claypool Publishers (2011)
46. Vardi, M.Y.: Reasoning about the past with two-way automata. In: Proc. of ICALP'98. pp. 628–641 (1998)
47. Visser, A., et al.: Uniform interpolation and layered bisimulation. In: Gödel'96: Logical foundations of mathematics, computer science and physics—Kurt Gödel's legacy, pp. 139–164. Association for Symbolic Logic (1996)