

Interpolants and Explicit Definitions in Extensions of the Description Logic \mathcal{EL}

Marie Fortin , Boris Konev , Frank Wolter

University of Liverpool

{mfortin, konev, wolter}@liverpool.ac.uk

Abstract

We show that the vast majority of extensions of the description logic \mathcal{EL} do not enjoy the Craig interpolation nor the projective Beth definability property. This is the case, for example, for \mathcal{EL} with nominals, \mathcal{EL} with the universal role, \mathcal{EL} with a role inclusion of the form $r \circ s \sqsubseteq s$, and for \mathcal{ELI} . It follows in particular that the existence of an explicit definition of a concept or individual 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 standard tractable extensions of \mathcal{EL} (such as \mathcal{EL}^{++}) and in EXPTIME for \mathcal{ELI} and various extensions. It follows that these existence problems are not harder than subsumption which is in sharp contrast to the situation for expressive DLs. We also obtain tight bounds for the size of interpolants and explicit definitions and the complexity of computing them: single exponential for tractable standard extensions of \mathcal{EL} and double exponential for \mathcal{ELI} and extensions. We close with a discussion of Horn-DLs such as Horn- \mathcal{ALCI} .

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 of concept, role, and individual names if, and only if, it is implicitly definable using Σ under \mathcal{O} . The importance of the PBDP for DL research stems from the fact that it provides a polynomial time reduction of the problem to decide the existence of an explicit definition to the well understood problem of subsumption checking. The existence of explicit definitions is important for numerous knowledge engineering tasks and applications of description logic ontologies, for example, the extraction of equivalent acyclic TBoxes from ontologies (ten Cate et al. 2006; ten Cate, Franconi, and Seylan 2013), the computation of referring expressions (or definite descriptions) for individuals (Artale et al. 2021b), the equivalent rewriting of ontology-mediated queries into concepts (Seylan, Franconi, and de Bruijn 2009; Lutz, Seylan, and Wolter 2019; Toman and Weddell 2021), the construction of alignments between ontologies (Geleta, Payne, and Tamma 2016), and the decomposition of ontologies (Konev et al. 2010).

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} , and \mathcal{ALCQI} enjoy the CIP and PBDP and sophisticated algorithms for computing interpolants and explicit definitions have been developed (ten Cate, Franconi, and Seylan 2013). 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 \mathcal{ALCO} (\mathcal{ALC} with nominals) and for \mathcal{ALCH} (\mathcal{ALC} with role hierarchies). This result 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 (Artale et al. 2021a).

Our aim in this article is threefold: (1) determine which members of the \mathcal{EL} -family of DLs enjoy the CIP/PBDP; (2) investigate the complexity of deciding the existence of interpolants/explicit definitions for those that do not enjoy it; and (3) establish tight bounds on the size of interpolants/explicit definitions and the complexity of computing them.

In what follows we discuss our main results. It has been shown in (Konev et al. 2010; Lutz, Seylan, and Wolter 2019) already that \mathcal{EL} and \mathcal{EL} with role hierarchies enjoy the CIP and PBDP. Rather surprisingly, it turns out that none of the remaining standard DLs in the \mathcal{EL} -family enjoy the CIP nor the PBDP.

Theorem 1. *The following DLs do not enjoy the CIP nor PBDP:*

1. \mathcal{EL} with the universal role,
2. \mathcal{EL} with nominals,
3. \mathcal{EL} with a single role inclusion $r \circ s \sqsubseteq s$,
4. \mathcal{EL} with role hierarchies and a transitive role,
5. the extension \mathcal{ELI} of \mathcal{EL} with inverse roles.

In Points 2 to 5, the CIP/PBDP also fails if the universal role can occur in interpolants/explicit definitions.

Theorem 1 also has interesting consequences that are not explicitly stated. For instance, it follows that neither the DL \mathcal{EL}^{++} introduced in (Baader, Brandt, and Lutz 2005) nor the extension of \mathcal{ELI} with any combination of nominals, role hierarchies, or transitive roles enjoy the CIP/PBDP. With the exception of the failure of the CIP/PBDP for \mathcal{EL} with nominals (without the universal role in interpolants/explicit definitions) (Artale et al. 2021b), our results are new.

It follows from Theorem 1 that the behaviour of extensions of \mathcal{EL} is fundamentally different from extensions of \mathcal{ALC} : adding role hierarchies to \mathcal{ALC} does not preserve the CIP/PBDP (Konev et al. 2009) but it does for \mathcal{EL} ; on the other hand, adding the universal role or inverse roles to \mathcal{ALC} preserves the CIP/PBDP (ten Cate, Franconi, and Seylan 2013) but it does not for \mathcal{EL} .

Theorem 1 leaves open the behaviour of a few natural DLs between \mathcal{EL} and its extension with arbitrary role inclusions. For instance, what happens if one only adds transitive roles or, more generally, role inclusions using a single role name only? To cover these cases we show a general result that implies that these DLs enjoy the CIP and PBDP. In particular, it follows that in Point 4 of Theorem 1 the combination of role hierarchies with a transitive role is necessary for failure of the CIP/PBDP.

We next discuss our main result about tractable extensions of \mathcal{EL} .

Theorem 2. *For \mathcal{EL} and any extension with any combination of nominals, role inclusions, the universal role, or \perp , 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 that can be computed in exponential time. This bound is optimal.*

It follows that for tractable extensions of \mathcal{EL} the complexity of deciding the existence of interpolants and explicit definitions does not depend on the CIP/PBDP, in sharp contrast to the behaviour of \mathcal{ALCO} and \mathcal{ALCH} . Moreover, the proof shows how interpolants and explicit definitions can be computed from the canonical models introduced in (Baader, Brandt, and Lutz 2005), if they exist. It applies derivation trees (first introduced in (Bienvenu, Lutz, and Wolter 2013) for DLs without nominals and role hierarchies) to estimate the size of interpolants and provide an exponential time algorithm for computing them.

Theorem 3. *For \mathcal{ELI} and any extension with any combination of nominals, the universal role, or \perp , the existence of interpolants and explicit definitions is EXPTIME-complete. If an interpolant/explicit definition exists, then there exists one of at most double exponential size that can be computed in double exponential time. This bound is optimal.*

The proof of Theorem 3 shows how an interpolant or explicit definition can be extracted from a (potentially infinite) tree-shaped canonical model. The EXPTIME complexity bound is proved using an encoding as an emptiness problem for tree automata that also uses derivation trees. It does not seem possible to obtain tight bounds on the size of interpolants using derivation trees; instead we generalize transfer sequences for this purpose (also first introduced in (Bienvenu, Lutz, and Wolter 2013)).

In the final section, we consider expressive Horn-DLs such as Horn- \mathcal{ALCI} . We first observe that Theorem 3 also holds for Horn- \mathcal{ALCI} and extensions with nominals and the universal role, provided one asks for interpolants and explicit definitions in \mathcal{ELI} (and extensions with nominals and the universal role, respectively). If one admits expressive Horn-concepts as interpolants or explicit definitions, then sometimes interpolants and explicit definitions exist that previously did not exist. We show that nevertheless the CIP/PBDP also fail in this case for DLs including Horn- \mathcal{ALC} , \mathcal{ELI} , and Horn- \mathcal{ALCI} .

Detailed proofs are given in the arxiv version of this article: <https://arxiv.org/abs/2202.07186>.

2 Related Work

The CIP and PBDP have been investigated extensively in databases, with applications to query rewriting under views and query compilation (Toman and Weddell 2011; Benedikt et al. 2016). The computation of explicit definitions under Horn ontologies can be seen as an instance of query reformulation under constraints (Deutsch, Popa, and Tannen 2006) which has been a major research topic for many years. The Chase and Backchase approach that is central to this research closely resembles our use of canonical models. We do not assume, however, that the chase terminates. In (Benedikt et al. 2016; 2017), it is shown that the reformulation of CQs into CQs under tgds can be reduced to entailment using Lyndon interpolation of first-order logic. By linking reformulation into CQs and definability using concepts, this approach can potentially be used to obtain alternative proofs of complexity upper bounds for the existence of interpolants and explicit definitions in our languages. Also relevant is the investigation of interpolation in basic modal logic (Maksimova and Gabbay 2005) and hybrid modal logic (Areces, Blackburn, and Marx 2001; ten Cate 2005).

The main aim of this article is to investigate explicit definability of concept and individual names under ontologies. We have therefore chosen a definition of the CIP and interpolants that generalizes the projective Beth definability property and explicit definability in a natural and useful way, following (ten Cate, Franconi, and Seylan 2013). There are, however, other notions of Craig interpolation that are of interest. 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 (Sofronie-Stokkermans 2008; Konev et al. 2010). Currently, it is unknown whether there exists any interesting relationship between this version of the CIP and the version we investigate in this article.

Craig interpolants should not be confused with uniform interpolants (or forgetting) (Lutz, Seylan, and Wolter 2012; Lutz and Wolter 2011; Nikitina and Rudolph 2014; Koopmann and Schmidt 2015). 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. Extending work on Boolean DLs we discussed already, it is shown that they become harder than validity also in the guarded and two-variable fragment (Jung and Wolter 2021). The interpolant existence problem for linear temporal logic LTL is considered in (Place and Zeitoun 2016). In the context of referring expressions, explicit definition existence is investigated in (Artale et al. 2021b), see also (Borgida, Toman, and Weddell 2016).

3 Preliminaries

Let N_C , N_R , and 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:

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

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 $r_1 \circ \dots \circ r_n \sqsubseteq r$ with r_1, \dots, r_n, r role names. An \mathcal{ELRO}_u -ontology with RIs is called an \mathcal{ELRO}_u -ontology. A set of RIs is a *role hierarchy* if all its RIs are of the form $r \sqsubseteq s$ with r, s role names.

A *signature* Σ is a set of concept, role, and individual names, uniformly referred to as (*non-logical*) *symbols*. We follow common practice and do not regard the universal role u as a non-logical symbol as its interpretation is fixed. 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 $A \in N_C$ a set $A^{\mathcal{I}} \subseteq \Delta^{\mathcal{I}}$, to each $r \in N_R$ a relation $r^{\mathcal{I}} \subseteq \Delta^{\mathcal{I}} \times \Delta^{\mathcal{I}}$, and to each $a \in N_I$ 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 (Baader et al. 2017). An interpretation \mathcal{I} satisfies a CI $C \sqsubseteq D$ if $C^{\mathcal{I}} \subseteq D^{\mathcal{I}}$ and an RI $r_1 \circ \dots \circ r_n \sqsubseteq r$ if $r_1^{\mathcal{I}} \circ \dots \circ r_n^{\mathcal{I}} \subseteq r^{\mathcal{I}}$. We say that \mathcal{I} is a *model* of an ontology \mathcal{O} 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. It is well known that for any \mathcal{ELIO}_u -ontology \mathcal{O} with or without RIs one can construct in polynomial time a conservative extension \mathcal{O}' using the same constructors as \mathcal{O} that is in normal form.

$\mathcal{L}(\Sigma)$ -concepts can be characterized using $\mathcal{L}(\Sigma)$ -simulations which we define next. Let \mathcal{I} and \mathcal{J} be interpretations. A relation $S \subseteq \Delta^{\mathcal{I}} \times \Delta^{\mathcal{J}}$ is called an $\mathcal{ELIO}(\Sigma)$ -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 $A \in N_C \cap \Sigma$;
2. if $d = a^{\mathcal{I}}$ and $(d, e) \in S$, then $e = a^{\mathcal{J}}$, for all $a \in N_I \cap \Sigma$;
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 $r \in N_R \cap \Sigma$.

S is called an $\mathcal{ELIO}_u(\Sigma)$ -simulation if $\Delta^{\mathcal{I}}$ is the domain of S and an $\mathcal{ELIO}(\Sigma)$ -simulation if Condition 3 also holds for inverse roles from Σ . Condition 2 is dropped if \mathcal{L} does not use nominals. We write $(\mathcal{I}, d) \preceq_{\mathcal{L}, \Sigma} (\mathcal{J}, e)$ if there exists an $\mathcal{L}(\Sigma)$ -simulation S between \mathcal{I} and \mathcal{J} with $(d, e) \in S$. 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 (Lutz and Wolter 2010; Lutz, Piro, and Wolter 2011).

Lemma 1. *Let $\mathcal{L} \in \{\mathcal{EL}, \mathcal{EL}_u, \mathcal{ELO}, \mathcal{ELO}_u, \mathcal{ELI}, \mathcal{ELI}_u\}$. Then $(\mathcal{I}, d) \preceq_{\mathcal{L}, \Sigma} (\mathcal{J}, e)$ implies $(\mathcal{I}, d) \leq_{\mathcal{L}, \Sigma} (\mathcal{J}, e)$. The converse direction holds if \mathcal{J} is finite.*

4 Craig Interpolation Property and Projective Beth Definability Property

We introduce the Craig interpolation property (CIP) as defined in (ten Cate, Franconi, and Seylan 2013) and the projective Beth definability property (PBDP) and prove Theorem 1 from the introduction to this article. 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.

Set $\text{sig}(\mathcal{O}, C) = \text{sig}(\mathcal{O}) \cup \text{sig}(C)$, 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} -interpolant¹ for $C_1 \sqsubseteq C_2$ under $\mathcal{O}_1, \mathcal{O}_2$ if

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

Definition 1. *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 $\mathcal{O}_1, \mathcal{O}_2$.*

¹Important variations of this definition are to drop \mathcal{O}_2 in Point 2 and \mathcal{O}_1 in Point 3, respectively, or to consider only one ontology $\mathcal{O} = \mathcal{O}_1 = \mathcal{O}_2$ and regard the signature Σ of the interpolant as an input given independently from \mathcal{O}, C_1, C_2 . This has an effect on the CIP, but our results on interpolant computation and existence are not affected.

We next define the relevant definability notions. Let \mathcal{O} be an ontology and A a concept name. Let $\Sigma \subseteq \text{sig}(\mathcal{O})$ 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 symbol that is not in Σ is interpreted in $\mathcal{I}_{|\Sigma}$. A concept A is called *implicitly definable using 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 $A \notin \Sigma$ is implicitly definable using Σ under \mathcal{O} iff $\mathcal{O} \cup \mathcal{O}_{\Sigma} \models A \equiv A'$, where \mathcal{O}_{Σ} is obtained from \mathcal{O} by replacing every symbol X not in Σ (including A) uniformly by a fresh symbol X' .

Definition 2. A DL \mathcal{L} has the *projective Beth definable property (PBDP)* if for any \mathcal{L} -ontology \mathcal{O} , concept name A , and signature $\Sigma \subseteq \text{sig}(\mathcal{O})$ the following holds: if A is implicitly definable using Σ under \mathcal{O} , then A is explicitly $\mathcal{L}(\Sigma)$ -definable under \mathcal{O} .

Remark 1. 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}_{Σ} 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} .

Remark 2. The PBDP implies that implicitly definable nominals are explicitly definable and that, more generally, every implicitly definable concept C is explicitly definable. This can be shown by adding $A \equiv C$ to the ontology for a fresh concept name A and asking for an explicit definition of A in the extended ontology.

Remark 3. The CIP and PBDP are invariant under adding \perp (interpreted as the empty set) to the languages introduced above. The straightforward proof is given in the appendix of the full version.

We next prove that the majority of tractable extensions of \mathcal{EL} does not enjoy the CIP nor PBDP.

Theorem 1. *The following DLs do not enjoy the CIP nor PBDP:*

1. \mathcal{EL} with the universal role,
2. \mathcal{EL} with nominals,
3. \mathcal{EL} with a single role inclusion $r \circ s \sqsubseteq s$,
4. \mathcal{EL} with role hierarchies and a transitive role,
5. \mathcal{EL} with inverse roles.

In Points 2 to 5, the CIP/PBDP also fails if the universal role can occur in interpolants/explicit definitions.

Proof. We first show that \mathcal{EL}_u does not enjoy the PBDP. Point 1 then follows using Remark 1. We define an \mathcal{EL}_u -ontology \mathcal{O}_u , signature Σ , and concept name A such that A is implicitly definable using Σ under \mathcal{O}_u but not $\mathcal{EL}_u(\Sigma)$ -explicitly definable under \mathcal{O}_u . Define \mathcal{O}_u as the following set of CIs:

$$A \sqsubseteq B, \quad D \sqcap \exists u.A \sqsubseteq E, \quad B \sqsubseteq \exists r.C$$

$C \sqsubseteq D, \quad B \sqcap \exists r.(C \sqcap E) \sqsubseteq A$,
and let $\Sigma = \{B, D, E, r\}$. We have $\mathcal{O}_u \models A \equiv B \sqcap \forall r.(D \rightarrow E)$,² so A is implicitly definable using Σ under \mathcal{O}_u . The interpretations \mathcal{I} and \mathcal{I}' given in Figure 1 show that A is not explicitly $\mathcal{EL}_u(\Sigma)$ -definable under \mathcal{O}_u . Indeed, \mathcal{I}

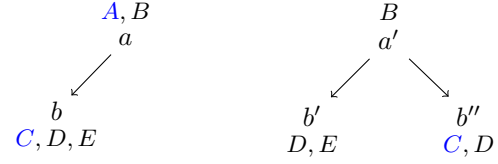


Figure 1: Interpretations \mathcal{I} (left) and \mathcal{I}' (right) used for \mathcal{O}_u .

and \mathcal{I}' are both models of \mathcal{O}_u , $a \in A^{\mathcal{I}}$, $a' \notin A^{\mathcal{I}'}$, and the relation $\{(a, a'), (b, b')\}$ is a $\mathcal{EL}_u(\Sigma)$ -simulation between \mathcal{I} and \mathcal{I}' . As $\mathcal{EL}_u(\Sigma)$ -concepts are preserved under $\mathcal{EL}_u(\Sigma)$ -simulations (Lemma 1), if $\mathcal{O}_u \models A \equiv F$ for some $\mathcal{EL}_u(\Sigma)$ -concept F , then from $a \in A^{\mathcal{I}}$ we obtain $a \in F^{\mathcal{I}}$. This implies $a' \in F^{\mathcal{I}'}$, and so $a' \in A^{\mathcal{I}'}$. As $a' \notin A^{\mathcal{I}'}$, we obtain a contradiction.

We next prove Point 2. An example from (Artale et al. 2021b) shows that $\mathcal{EL}\mathcal{O}$ does not enjoy the CIP/PBDP. Here we show that $\mathcal{EL}\mathcal{O}$ does not enjoy the CIP/PBDP, even if interpolants/explicit definitions are from $\mathcal{EL}\mathcal{O}_u$. Let \mathcal{O}_n 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 using Σ under \mathcal{O}_n as $\mathcal{O}_n \models A \equiv \forall s.(\exists s.\{c\} \rightarrow Q_1)$. The relation $\{(a, a'), (b, b'), (c, c')\}$ is an $\mathcal{EL}\mathcal{O}_u(\Sigma)$ -simulation between the interpretations \mathcal{I} and \mathcal{I}' defined in Figure 2. Now we can apply the same argument as in Point 1 to show that A is not explicitly $\mathcal{EL}\mathcal{O}_u(\Sigma)$ -definable under \mathcal{O}_n .

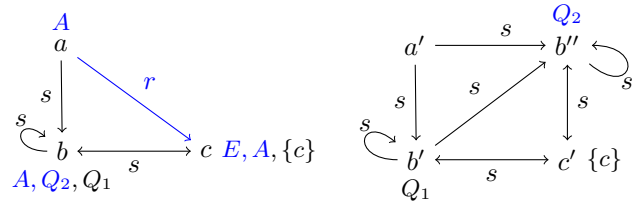


Figure 2: Interpretations \mathcal{I} (left) and \mathcal{I}' (right) used for \mathcal{O}_n .

For Point 3, let \mathcal{O}_r contain

$$A \sqsubseteq \exists r.E, \quad E \sqsubseteq \exists s.B, \quad \exists s.B \sqsubseteq A, \quad r \circ s \sqsubseteq s,$$

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

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

We show that there does not exist any $\mathcal{EL}_u(\Sigma)$ -explicit definition of A under \mathcal{O}_r . The interpretations \mathcal{I} and \mathcal{I}' given

²Here and in what follows we use standard \mathcal{ALC} syntax and semantics and set $C \rightarrow D := \neg C \sqcup D$ (Baader et al. 2017).

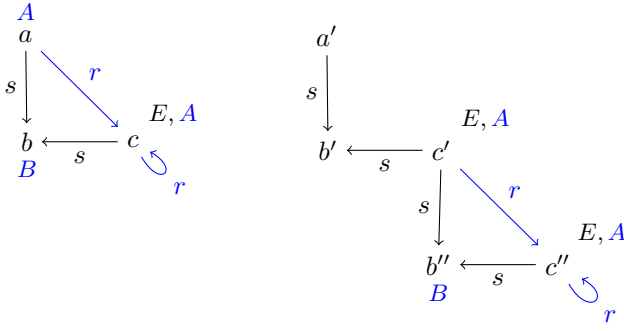


Figure 3: Interpretations \mathcal{I} (left) and \mathcal{I}' (right) used for \mathcal{O}_r .

in Figure 3 are both models of \mathcal{O}_r , $a \in A^{\mathcal{I}}$, $a' \notin A^{\mathcal{I}'}$, and the relation $\{(a, a'), (b, b'), (c, c')\}$ is an $\mathcal{EL}_u(\Sigma)$ -simulation between \mathcal{I} and \mathcal{I}' . One can now show in the same way as in Point 1 that no $\mathcal{EL}_u(\Sigma)$ -definition of A under \mathcal{O}_r exists.

Point 4 is shown in the appendix of the full version using a modification of the ontology used for Point 3.

To prove Point 5, obtain an \mathcal{ELI} -ontology \mathcal{O}_i from \mathcal{O}_u defined above by replacing the second CI of \mathcal{O}_u by $D \sqcap \exists r^-.A \sqsubseteq E$. Let, as before, $\Sigma = \{B, D, E, r\}$. Then A is implicitly definable from Σ under \mathcal{O}_i (the same explicit definition works), but A is not explicitly $\mathcal{ELI}_u(\Sigma)$ -definable under \mathcal{O}_i (the same interpretations \mathcal{I} and \mathcal{I}' work). \square

We next discuss a general positive result on interpolation and explicit definition existence that shows that Theorem 1 is essentially optimal. A set \mathcal{R} of RIs is *safe* for a signature Σ if for each RI $r_1 \circ \dots \circ r_n \sqsubseteq r \in \mathcal{R}$, $n \geq 1$, if $\{r_1, \dots, r_n, r\} \cap \Sigma \neq \emptyset$ then $\{r_1, \dots, r_n, r\} \subseteq \Sigma$.

Theorem 4. *Let $\mathcal{O}_1, \mathcal{O}_2$ be \mathcal{EL} -ontologies with RIs, C_1, C_2 be \mathcal{EL} -concepts, and set $\Sigma = \text{sig}(\mathcal{O}_1, C_1) \cap \text{sig}(\mathcal{O}_2, C_2)$. 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.*

The proof technique is based on simulations and similar to (Konev et al. 2010; Lutz, Seylan, and Wolter 2019). Theorem 4 has a few interesting consequences. For instance, \mathcal{EL} with transitive roles enjoys both the CIP and PBDP since transitivity is expressed by the role inclusion $r \circ r \sqsubseteq r$ which is safe for any signature (as it only uses a single role name).

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 \perp does not affect the complexity of the decision problems.

Definition 3. *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 4. *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 4. 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 on the size of \mathcal{L} -interpolants provides a bound on the size of \mathcal{L} -explicit definitions. The proof is similar to the proof of Remark 1.

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

Lemma 2. *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 normal form conservative extensions of $\mathcal{O}_1 \cup \{A \equiv C\}$ and, respectively, $\mathcal{O}_2 \cup \{B \equiv D\}$, computed in polynomial time. One can show that \mathcal{O}'_1 and \mathcal{O}'_2 are as required. \square

Remark 5. Assume that \mathcal{L} is any of the DLs introduced above and let \mathcal{L}_\perp denote its extension with \perp . Then \mathcal{L} -interpolant existence and \mathcal{L} -explicit definition existence can be reduced in polynomial time to \mathcal{L}_\perp -interpolant existence and \mathcal{L}_\perp -explicit definition existence, respectively. The converse direction also holds modulo an oracle deciding whether $\mathcal{O} \models C \sqsubseteq \perp$.

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

The aim of this section is to analyse interpolants and explicit definitions for extensions of \mathcal{EL} with any combination of nominals, role inclusions, or the universal role. We show the following result from the introduction.

Theorem 2. *For \mathcal{EL} and any extension with any combination of nominals, role inclusions, the universal role, or \perp , 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 that can be computed in exponential time. This bound is optimal.*

Before we start with a sketch of the proof we give instructive examples showing that the exponential bound on the size of explicit definitions is optimal.

Example 1. Variants of the following example have already been used for various succinctness arguments in DL. Let

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

and $\Sigma_0 = \{r_1, r_2, B_n, M\}$. A triggers a marker M and a binary tree of depth n whose leaves are decorated with B_n . Conversely, if B_n is true at all leaves of a binary tree of depth n , 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}_b . Next let

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

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}_p .

Observe that using \mathcal{O}_b one enforces explicit definitions of exponential size by generating a binary tree of linear depth whereas using \mathcal{O}_p this is achieved by generating a path of exponential length. The latter can only happen if role inclusions are used in the ontology. One insight provided by the exponential upper bound on the size of explicit definitions in Theorem 2 is that the two examples cannot be combined to enforce a binary tree of exponential depth.

To continue with the proof we 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 $A \in \mathbf{N}_C$, $r \in \mathbf{N}_R$, $a \in \mathbf{N}_I$, and x, y individual variables (we call individuals used in ABoxes variables to distinguish them from individual names used in nominals). We denote by $\text{ind}(\mathcal{A})$ the set of individual variables in \mathcal{A} . A Σ -ABox is an ABox using symbols from Σ only. Models of ABoxes are defined as usual. We do not make the unique name assumption.

Every interpretation \mathcal{I} defines an 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, ABoxes \mathcal{A} define interpretations in the obvious way (by identifying variables x, y if $\{a\}(x), \{a\}(y) \in \mathcal{A}$). We associate with every ABox \mathcal{A} a directed graph $G_{\mathcal{A}} = (\text{ind}(\mathcal{A}), \bigcup_{r \in \mathbf{N}_R} \{(x, y) \mid r(x, y) \in \mathcal{A}\})$. 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* in the sense that $G_{\mathcal{A}}$ is acyclic and $r(x, y) \in \mathcal{A}$ and $s(x, y) \in \mathcal{A}$ imply $r = s$. A *pointed ABox* is a pair \mathcal{A}, x with $x \in \text{ind}(\mathcal{A})$. Then $\mathcal{EL}\mathcal{O}_u(\Sigma)$ -concepts correspond to pointed Σ -ABoxes \mathcal{A}, x such that \mathcal{A} is ditree-shaped modulo $\mathbf{N}_I \cap \Sigma$ and $\mathcal{EL}\mathcal{O}(\Sigma)$ -concepts correspond to *rooted* pointed Σ -ABoxes \mathcal{A}, x such that \mathcal{A} is ditree-shaped modulo $\mathbf{N}_I \cap \Sigma$, where \mathcal{A}, x is called *rooted* if for every $y \in \text{ind}(\mathcal{A})$ there is a path from x to y in $G_{\mathcal{A}}$. We write $\mathcal{O}, \mathcal{A} \models C(x)$ if $x^{\mathcal{I}} \in C^{\mathcal{I}}$ for every model \mathcal{I} of \mathcal{O} and \mathcal{A} .

Given an $\mathcal{EL}\mathcal{R}\mathcal{O}_u$ -ontology \mathcal{O} in normal form and a concept name A , one can construct in polynomial time the *canonical model* $\mathcal{I}_{\mathcal{O}, A}$ of \mathcal{O} and A using the approach introduced in (Baader, Brandt, and Lutz 2005). More generally, the canonical model $\mathcal{I}_{\mathcal{O}, \mathcal{A}}$ for an ABox \mathcal{A} and ontology \mathcal{O} can be constructed in polynomial time and is a model of both \mathcal{O} and \mathcal{A} such that for any $\mathcal{EL}\mathcal{O}_u$ -concept C using symbols from \mathcal{O} only and any $x \in \text{ind}(\mathcal{A})$,

(\dagger) $\mathcal{O}, \mathcal{A} \models C(x)$ iff $x \in C^{\mathcal{I}_{\mathcal{O}, \mathcal{A}}}$,

details are given in the appendix of the full version. We let $\mathcal{I}_{\mathcal{O}, A} = \mathcal{I}_{\mathcal{O}, \mathcal{A}}$ with $\mathcal{A} = \{A(\rho_A)\}$. Note that in (Baader, Brandt, and Lutz 2005) the condition (\dagger) is only stated for subconcepts C of the ontology \mathcal{O} , thus (\dagger) requires a proof.

Example 2. The interpretations \mathcal{I} defined in the proof of Theorem 1 define canonical models $\mathcal{I}_{\mathcal{O}, A}$ with $\rho_A = a$ for the ontologies $\mathcal{O} \in \{\mathcal{O}_u, \mathcal{O}_n, \mathcal{O}_r, \mathcal{O}_i\}$. The interpretations \mathcal{I}' define canonical models $\mathcal{I}_{\mathcal{O}, \mathcal{A}^{\Sigma}}$ with $\mathcal{A}^{\Sigma}_{\mathcal{O}}$ the Σ -reduct of $\mathcal{I}_{\mathcal{O}, A}$ regarded as an ABox and $\rho_A = a'$.

The *directed unfolding* of a pointed Σ -ABox \mathcal{A}, x into a pointed Σ -ABox \mathcal{A}^u, x that is ditree-shaped modulo $\Sigma \cap \mathbf{N}_I$ is defined in the standard way. In the *rooted directed unfolding*, nodes that cannot be reached from x via role names are dropped.

Assume now that \mathcal{O} is in normal form and A a concept name. Let $\mathcal{A}^{\Sigma}_{\mathcal{O}, A}$ be the Σ -reduct of the canonical model $\mathcal{I}_{\mathcal{O}, A}$, regarded as an ABox. Denote by $\mathcal{A}^{\Sigma, u}_{\mathcal{O}, A}, \rho_A$ the directed unfolding of $\mathcal{A}^{\Sigma}_{\mathcal{O}, A}, \rho_A$, by $\mathcal{A}^{\downarrow \Sigma}_{\mathcal{O}, A}, \rho_A$ the sub-ABox of $\mathcal{A}^{\Sigma}_{\mathcal{O}, A}$ rooted in ρ_A , and by $\mathcal{A}^{\downarrow \Sigma, u}_{\mathcal{O}, A}, \rho_A$ its rooted directed unfolding. Theorem 2 is a direct consequence of the following characterization of interpolants.

Theorem 5. *There exists a polynomial p such that the following conditions are equivalent for all $\mathcal{EL}\mathcal{R}\mathcal{O}_u$ -ontologies $\mathcal{O}_1, \mathcal{O}_2$ in normal form, concept names A, B , and $\Sigma = \text{sig}(\mathcal{O}_1, A) \cap \text{sig}(\mathcal{O}_2, B)$:*

1. An $\mathcal{EL}\mathcal{O}_u$ -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 finite subset \mathcal{A} of $\mathcal{A}^{\Sigma, u}_{\mathcal{O}_1 \cup \mathcal{O}_2, A}$ with $|\text{ind}(\mathcal{A})| \leq 2^{p(\|\mathcal{O}_1 \cup \mathcal{O}_2\|)}$ such that the $\mathcal{EL}\mathcal{O}_u$ -concept corresponding to \mathcal{A}, ρ_A is an $\mathcal{EL}\mathcal{O}_u$ -interpolant for $A \sqsubseteq B$ under $\mathcal{O}_1, \mathcal{O}_2$.

The same equivalences hold if in Points 1 to 3, $\mathcal{EL}\mathcal{O}_u$ is replaced by $\mathcal{EL}\mathcal{O}$, $\mathcal{A}^{\Sigma}_{\mathcal{O}_1 \cup \mathcal{O}_2, A}$ by $\mathcal{A}^{\downarrow \Sigma}_{\mathcal{O}_1 \cup \mathcal{O}_2, A}$, and $\mathcal{A}^{\Sigma, u}_{\mathcal{O}_1 \cup \mathcal{O}_2, A}$ by $\mathcal{A}^{\downarrow \Sigma, u}_{\mathcal{O}_1 \cup \mathcal{O}_2, A}$.

In Point 3, \mathcal{A} can be computed in exponential time, if it exists.

Note that the polynomial time decidability of interpolant existence follows from Point 2 of Theorem 5 (and the tractability of $\mathcal{EL}\mathcal{R}\mathcal{O}_u$ (Baader, Brandt, and Lutz 2005)).

Example 3. Our proof of Theorem 2 can be regarded as an application of Theorem 5: by Example 2, the interpretations \mathcal{I} and \mathcal{I}' coincide with the canonical models $\mathcal{I}_{\mathcal{O}, A}$ and $\mathcal{I}_{\mathcal{O}, \mathcal{A}^{\Sigma}_{\mathcal{O}, A}}$ and so $\rho_A = a' \notin A^{\mathcal{I}_{\mathcal{O}, \mathcal{A}^{\Sigma}_{\mathcal{O}, A}}}$ is equivalent to $\mathcal{O}, \mathcal{A}^{\Sigma}_{\mathcal{O}, A} \not\models A(\rho_A)$ (Point 2 in Theorem 5).

The following example illustrates the difference between the existence of explicit definitions in $\mathcal{EL}\mathcal{O}$ and $\mathcal{EL}\mathcal{O}_u$ and thus the need for moving to the ABoxes $\mathcal{A}^{\downarrow \Sigma}_{\mathcal{O}, A}$, and $\mathcal{A}^{\downarrow \Sigma, u}_{\mathcal{O}, A}$ if one does not admit the universal role in explicit definitions.

Example 4. Let $\mathcal{O} = \{A \sqsubseteq \{b\}, A \sqsubseteq \exists r. B, B \sqsubseteq \exists s. A\}$ and let $\Sigma = \{b, B\}$. Then A is explicitly $\mathcal{EL}\mathcal{O}_u(\Sigma)$ -definable under \mathcal{O} since $\mathcal{O} \models A \equiv \{b\} \sqcap \exists u. B$ but A

is not explicitly $\mathcal{EL}\mathcal{O}(\Sigma)$ -definable. Note that in this case $\mathcal{A}_{\mathcal{O},A}^{\Sigma} = \{\{b\}(\rho_A), B(y)\}$ but $\mathcal{A}_{\mathcal{O},A}^{\downarrow\Sigma} = \{\{b\}(\rho_A)\}$.

We next sketch the proof idea for Theorem 5 for the case with universal role in interpolants. We show “1. \Rightarrow 2.”, observe that “3. \Rightarrow 1.” is trivial, and then sketch the proof of “2. \Rightarrow 3.” and the exponential time algorithm computing interpolants, details are provided in the appendix of the full version. For “1. \Rightarrow 2.” assume that C is an $\mathcal{EL}\mathcal{O}_u(\Sigma)$ -concept with (i) $\mathcal{O}_1 \cup \mathcal{O}_2 \models A \sqsubseteq C$ and (ii) $\mathcal{O}_1 \cup \mathcal{O}_2 \models C \sqsubseteq B$. By (\dagger) and (i), $\mathcal{A}_{\mathcal{O}_1 \cup \mathcal{O}_2, A}^{\Sigma} \models C(\rho_A)$. But then by (ii) $\mathcal{O}_1 \cup \mathcal{O}_2, \mathcal{A}_{\mathcal{O}_1 \cup \mathcal{O}_2, A}^{\Sigma} \models B(\rho_A)$, as required.

If one does not impose a bound on the size of \mathcal{A} in Point 3, then one can prove “2. \Rightarrow 3.” using compactness and a generalization of *unraveling tolerance* according to which $\mathcal{O}_1 \cup \mathcal{O}_2, \mathcal{A}_{\mathcal{O}_1 \cup \mathcal{O}_2, A}^{\Sigma}$ and $\mathcal{O}_1 \cup \mathcal{O}_2, \mathcal{A}_{\mathcal{O}_1 \cup \mathcal{O}_2, A}^{\Sigma, u}$ entail the same $C(\rho_A)$ (Lutz and Wolter 2017; Hernich et al. 2020). As we are interested in an exponential bound on the size of \mathcal{A} (and a deterministic exponential time algorithm computing it) we require a more syntactic approach. Our proof of “2. \Rightarrow 3.” is based on *derivation trees* which represent a derivation of a fact $C(a)$ from an ontology \mathcal{O} and ABox \mathcal{A} using a labeled tree. Our derivation trees generalize those introduced in (Bienvenu, Lutz, and Wolter 2013; Baader et al. 2016) to languages with nominals and role inclusions. Reflecting the use of individual names and concept names in the construction of the domain of the canonical model (Baader, Brandt, and Lutz 2005), we assume $a \in \Delta := \text{ind}(\mathcal{A}) \cup ((N_C \cup N_I) \cap \text{sig}(\mathcal{O}))$ and $C \in \Theta := \{\top\} \cup (N_C \cap \text{sig}(\mathcal{O})) \cup \{\{a\} \mid a \in N_I \cap \text{sig}(\mathcal{O})\}$. Then a derivation tree (T, V) for $(a, C) \in \Delta \times \Theta$ is a tree T with a labeling function $V : T \rightarrow \Delta \times \Theta$ such that $V(\varepsilon) = (a, C)$ and (V, T) satisfies rules stating under which conditions the label of n is derived in one step from the labels of the successors of n . To illustrate, the existence of successors n_1, n_2 of n with $V(n_1) = (a, C_1)$ and $V(n_2) = (a, C_2)$ justifies $V(n) = (a, C)$ if $\mathcal{O} \models C_1 \sqcap C_2 \sqsubseteq C$. The rules are given in the appendix of the full version, we only discuss the rule used to capture derivations using RIs: $V(n) = (a_1, C)$ is justified if there are role names r_2, \dots, r_{2k-2}, r such that (a_{2k}, C') is a label of a successor of n , $\mathcal{O} \models \exists r.C' \sqsubseteq C$, $\mathcal{O} \models r_2 \circ \dots \circ r_{2k-2} \sqsubseteq r$, and the situation depicted in Figure 4 holds, where the “dotted lines” stand for ‘either $a_i = a_{i+1}$ or some $(a_i, \{c\}), (a_{i+1}, \{c\})$ with $c \in N_I$ are labels of successors of n ’, and \hat{r}_i stands for ‘either $r(a_i, a_{i+1}) \in \mathcal{A}$ or some (a_i, C_i) is a label of a successor of n and $\mathcal{O} \models C_i \sqsubseteq \exists r_i.\{a_{i+1}\}$ if $a_{i+1} \in N_I$ and $\mathcal{O} \models C_i \sqsubseteq \exists r_i.a_{i+1}$ if $a_{i+1} \in N_C$. Moreover, for all $a_i \neq a_1, 1 \leq i \leq 2k$, there exists a successor of n with label (a_i, D) for some D . The soundness of this rule should be clear, completeness can be shown similarly to the analysis of canonical models.

The length of the sequence a_1, \dots, a_{2k} can be exponential (for instance, in Example 1 for the fact (ρ_A, A) in $\mathcal{O}_p, \mathcal{A}_{\mathcal{O}_p, A}^{\Sigma_1}$). One can show, however, that its length can be bounded without affecting completeness by $2^q(\|\mathcal{O}\| + \|\mathcal{A}\|)$ with q a polynomial. The following lemma summarizes the main properties of derivation trees.

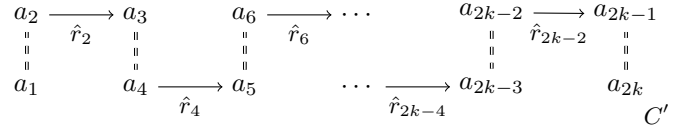


Figure 4: Rule for Role Inclusions.

Lemma 3. *Let \mathcal{O} be an $\mathcal{EL}\mathcal{RO}_u$ -ontology in normal form and \mathcal{A} a finite $\text{sig}(\mathcal{O})$ -ABox. Then*

1. $\mathcal{O}, \mathcal{A} \models A(x)$ if and only if there is a derivation tree for $A(x)$ in \mathcal{O}, \mathcal{A} . Moreover, if a derivation tree exists, then there exists one of depth and outdegree bounded by $(\|\mathcal{A}\| + \|\mathcal{O}\|) \times \|\mathcal{O}\|$ which can be constructed in exponential time in $\|\mathcal{O}\| + \|\mathcal{A}\|$.
2. If (T, V) is a derivation tree for $A(x)$ in \mathcal{O}, \mathcal{A} of at most exponential size, then one can construct in exponential time (in $\|\mathcal{A}\| + \|\mathcal{O}\|$) a derivation tree (T', V') for $A(x)$ in $\mathcal{O}, \mathcal{A}^u$ with \mathcal{A}^u the directed unfolding of \mathcal{A} modulo $\Sigma = \text{sig}(\mathcal{A}) \cap N_I$ and T' of the same depth as T and such that the outdegree of T' does not exceed $\max\{3, 3n\}$ with n the length of the longest chain $a_1 \dots a_n$ used in the rule for RIs in the derivation tree (T, V) .

Proof. We sketch the idea. For Point 1, the bound on the depth of derivation trees can be proved by observing that one can assume (using a standard pumping argument) that the labels of distinct nodes on a single path are distinct and the bound on the outdegree can be proved by observing that one can trivially assume that all successor nodes of a node have distinct labels. For the construction of derivation trees, let F_n denote the set of facts in $\Delta \times \Theta$ for which there is a derivation tree of depth at most n . Then one can construct in exponential time derivation trees for all facts in any F_n , $n \leq (\|\mathcal{A}\| + \|\mathcal{O}\|) \times \|\mathcal{O}\|$ by starting with derivation trees of depth 0 for members of F_0 , and then constructing derivation trees of depth $i + 1$ for members of F_{i+1} using the trees for members of F_0, \dots, F_i . For Point 2, the transformation of (T, V) into (T', V') is by induction over rule application, the only interesting step being the rule for RIs. Using the ontology \mathcal{O}_p of Example 1 one can see that the exponential blow-up of the outdegree is unavoidable. \square

We are now in the position to complete the sketch of the proof of “2. \Rightarrow 3.” Assume that Point 2 holds. Then $\mathcal{O}_1 \cup \mathcal{O}_2, \mathcal{A}_{\mathcal{O}_1 \cup \mathcal{O}_2, A}^{\Sigma} \models B(\rho_A)$. By Point 1 of Lemma 3 we can construct a derivation tree (T, V) for (ρ_A, B) in $\mathcal{O}_1 \cup \mathcal{O}_2, \mathcal{A}_{\mathcal{O}_1 \cup \mathcal{O}_2, A}^{\Sigma}$ of polynomial depth and outdegree in exponential time. By Point 2 of Lemma 3 we can transform (T, V) into a derivation tree (T', V') for (ρ_A, B) in $\mathcal{O}_1 \cup \mathcal{O}_2, \mathcal{A}_{\mathcal{O}_1 \cup \mathcal{O}_2, A}^{\Sigma, u}$ in exponential time. Now let \mathcal{A} be the restriction of $\mathcal{A}_{\mathcal{O}_1 \cup \mathcal{O}_2, A}^{\Sigma, u}$ to all $x \in \text{ind}(\mathcal{A}_{\mathcal{O}_1 \cup \mathcal{O}_2, A}^{\Sigma, u})$ which occur in a label of V' . Then (T', V') is also a derivation tree for (ρ_A, B) in $\mathcal{O}_1 \cup \mathcal{O}_2, \mathcal{A}$ and so $\mathcal{O}_1 \cup \mathcal{O}_2, \mathcal{A} \models B(\rho_A)$. It follows that the $\mathcal{EL}\mathcal{O}_u(\Sigma)$ -concept corresponding to \mathcal{A} is an interpolant for $A \sqsubseteq B$ under $\mathcal{O}_1 \cup \mathcal{O}_2$. Its size is at most exponential in $\|\mathcal{O}_1 \cup \mathcal{O}_2\|$ since (T', V') is at most exponential in $\|\mathcal{O}_1 \cup \mathcal{O}_2\| + \|\mathcal{A}_{\mathcal{O}_1 \cup \mathcal{O}_2, A}^{\Sigma}\|$, and so also in $\|\mathcal{O}_1 \cup \mathcal{O}_2\|$.

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

We analyze interpolants and explicit definitions for \mathcal{ELI} and its extensions with nominals and universal roles, and show the following result from the introduction.

Theorem 3. *For \mathcal{ELI} and any extension with any combination of nominals, the universal role, or \perp , the existence of interpolants and explicit definitions is EXPTIME-complete. If an interpolant/explicit definition exists, then there exists one of at most double exponential size that can be computed in double exponential time. This bound is optimal.*

The double exponential lower bound on the size of explicit definitions and interpolants is shown in the appendix of the full version. The proof is inspired by similar lower bounds for the size of FO-rewritings and uniform interpolants (Lutz and Wolter 2010; Nikitina and Rudolph 2014). To prove the remaining claims of Theorem 3, we lift Theorem 5 to \mathcal{ELI} . The main differences are that (1) we now associate undirected graphs with ABoxes and also unfold along inverse roles; (2) that canonical models become potentially infinite but tree-shaped; (3) that therefore deciding the new variant of Point 2 of Theorem 5 is not an instance of standard entailment checking in \mathcal{ELI} , instead we give a reduction to emptiness checking for tree automata; and (4) that to bound the size of \mathcal{A} in Point 3, we employ transfer sequences (and not derivation trees) to represent how facts are derived.

In more detail, associate with every ABox \mathcal{A} the undirected graph $G_{\mathcal{A}}^u = (\text{ind}(\mathcal{A}), \bigcup_{r \in \text{NR}} \{\{x, y\} \mid r(x, y) \in \mathcal{A}\})$. We say that \mathcal{A} is *tree-shaped* if $G_{\mathcal{A}}^u$ 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 observe that $\mathcal{ELIO}_u(\Sigma)$ -concepts correspond to pointed Σ -ABoxes \mathcal{A}, x such that \mathcal{A} is tree-shaped modulo $\text{N}_1 \cap \Sigma$. $\mathcal{ELIO}(\Sigma)$ -concepts correspond to *weakly rooted* pointed Σ -ABoxes \mathcal{A}, x such that \mathcal{A} is tree-shaped modulo $\text{N}_1 \cap \Sigma$, where \mathcal{A}, x is called weakly rooted if for every $y \in \text{ind}(\mathcal{A})$ there is a path from x to y in $G_{\mathcal{A}}^u$.

For every \mathcal{ELIO}_u -ontology \mathcal{O} and concept A there exists a (potentially infinite) pointed canonical model $\mathcal{I}_{\mathcal{O}, A}, \rho_A$ such that the ABox $\mathcal{A}_{\mathcal{O}, A}$ corresponding to $\mathcal{I}_{\mathcal{O}, A}$ is tree-shaped modulo $\text{N}_1 \cap \text{sig}(\mathcal{O})$. The property (\dagger) used in the context of canonical models for tractable extensions of \mathcal{EL} holds here as well. We also require the *undirected unfolding* of a pointed Σ -ABox \mathcal{A}, x into a pointed Σ -ABox \mathcal{A}^*, x which is tree-shaped modulo $\Sigma \cap \text{N}_1$. In the *rooted undirected unfolding*, nodes that cannot be reached from x via roles are dropped.

Assume now that \mathcal{O} is in normal form and A a concept name. Let $\mathcal{A}_{\mathcal{O}, A}^{\Sigma}$ be the Σ -reduct of the canonical model $\mathcal{I}_{\mathcal{O}, A}$, regarded as an ABox. Denote by $\mathcal{A}_{\mathcal{O}, A}^{\Sigma, *}, \rho_A$ the undirected unfolding of $\mathcal{A}_{\mathcal{O}, A}^{\Sigma}, \rho_A$, by $\mathcal{A}_{\mathcal{O}, A}^{\downarrow w \Sigma}, \rho_A$ the sub-ABox of $\mathcal{A}_{\mathcal{O}, A}^{\Sigma}$ weakly rooted in ρ_A , and by $\mathcal{A}_{\mathcal{O}, A}^{\downarrow w \Sigma, *}, \rho_A$ its rooted undirected unfolding. Then we lift Theorem 5 as follows.

Theorem 6. *There exists a polynomial p such that the following conditions are equivalent for all \mathcal{ELIO}_u -ontologies $\mathcal{O}_1, \mathcal{O}_2$ in normal form, concept names A, B , and $\Sigma = \text{sig}(\mathcal{O}_1, A) \cap \text{sig}(\mathcal{O}_2, B)$:*

1. An \mathcal{ELIO}_u -interpolant for $A \sqsubseteq B$ under $\mathcal{O}_1, \mathcal{O}_2$ exists;
2. $\mathcal{O}_1 \cup \mathcal{O}_2, \mathcal{A}_{\mathcal{O}_1 \cup \mathcal{O}_2, A}^{\Sigma} \models B(\rho_A)$;
3. there exists a finite subset \mathcal{A} of $\mathcal{A}_{\mathcal{O}_1 \cup \mathcal{O}_2, A}^{\Sigma, *}$ with $|\text{ind}(\mathcal{A})| \leq 2^{2^{p(\|\mathcal{O}_1 \cup \mathcal{O}_2\|)}}$ such that the \mathcal{ELIO}_u -concept corresponding to \mathcal{A}, ρ_A is an \mathcal{ELIO}_u -interpolant for $A \sqsubseteq B$ under $\mathcal{O}_1, \mathcal{O}_2$.

The same equivalences hold if in Points 1 to 3, \mathcal{ELIO}_u is replaced by \mathcal{ELIO} , $\mathcal{A}_{\mathcal{O}_1 \cup \mathcal{O}_2, A}^{\Sigma}$ by $\mathcal{A}_{\mathcal{O}_1 \cup \mathcal{O}_2, A}^{\downarrow w \Sigma}$, and $\mathcal{A}_{\mathcal{O}_1 \cup \mathcal{O}_2, A}^{\Sigma, *}$ by $\mathcal{A}_{\mathcal{O}_1 \cup \mathcal{O}_2, A}^{\downarrow w \Sigma, *}$.

In Point 3, \mathcal{A} can be computed in double exponential time, if it exists.

We first sketch how tree automata are used to show that Point 2 entails an exponential time upper bound for deciding the existence of an interpolant. To this end we represent finite prefix-closed subsets \mathcal{A} of $\mathcal{A}_{\mathcal{O}_1 \cup \mathcal{O}_2, A}^{\Sigma}$ as trees and design

- a non-deterministic tree automaton over finite trees (NTA), \mathfrak{A}_1 , that accepts exactly those trees that represent prefix-closed finite subsets of $\mathcal{A}_{\mathcal{O}_1 \cup \mathcal{O}_2, A}^{\Sigma}$;
- a two-way alternating tree automaton over finite trees (2ATA), \mathfrak{A}_2 , that accepts exactly those trees that represent a pointed ABox \mathcal{A}, ρ with $\mathcal{O}_1 \cup \mathcal{O}_2, \mathcal{A} \models B(\rho)$.

Similar tree automata techniques have been used e.g. in (Jung et al. 2020). \mathfrak{A}_1 is constructed using the definition of canonical models; its states are essentially types occurring in the canonical model and it can be constructed in exponential time. The 2ATA \mathfrak{A}_2 tries to construct a derivation tree for $B(\rho)$ in $\mathcal{O}_1 \cup \mathcal{O}_2, \mathcal{A}$, given as input a tree representing \mathcal{A}, ρ . It has polynomially many states, and can thus be turned into an equivalent NTA with exponentially many states (Vardi 1998). By taking the intersection with \mathfrak{A}_1 , one can then check in exponential time whether $L(\mathfrak{A}_1) \cap L(\mathfrak{A}_2) \neq \emptyset$, that is, whether $\mathcal{O}_1 \cup \mathcal{O}_2, \mathcal{A}_{\mathcal{O}_1 \cup \mathcal{O}_2, A}^{\Sigma} \models B(\rho_A)$.

We return to the proof of Theorem 6. The interesting implication is “2. \Rightarrow 3.” and the double exponential computation of interpolants. In this case we use transfer sequences to obtain a bound on the size of the subset \mathcal{A} of $\mathcal{A}_{\mathcal{O}_1 \cup \mathcal{O}_2, A}^{\Sigma, *}$ needed to derive $B(\rho_A)$ (we note that for \mathcal{ELI} without nominals one can also use the automata encoding above). Transfer sequences describe how facts are derived in a tree-shaped ABox and allow to determine when individuals a and b behave sufficiently similar so that the subtree rooted at a can be replaced by the subtree rooted at b (Bienvenu, Lutz, and Wolter 2013) without affecting a derivation. This technique can be used to show that one can always choose a prefix closed subset \mathcal{A} of $\mathcal{A}_{\mathcal{O}_1 \cup \mathcal{O}_2, A}^{\Sigma, *}$ of at most exponential depth. This also implies that \mathcal{A} can be obtained in double exponential time by constructing the canonical model up to depth $2^q(\|\mathcal{O}_1 \cup \mathcal{O}_2\|)$ with q a polynomial.

8 Expressive Horn Description Logics

We address two questions regarding expressive Horn-DLs. (1) Can our results for \mathcal{ELI} and extensions be lifted to more expressive Horn-DLs? (2) In the examples provided in the proof of Theorem 1 we sometimes (for example, for \mathcal{EL}_u and \mathcal{ELI}) construct explicit Horn-DL definitions to show implicit definability of concept names. Are Horn-DL concepts always sufficient to obtain an explicit definition if an implicit definition exists? We provide a positive answer to (1) if one only admits \mathcal{ELIO}_u -concepts (or fragments) as interpolants/explicit definitions and a negative answer to (2) in the sense that \mathcal{ELI} and various other Horn-DLs do not enjoy the CIP/PBDP even if one admits Horn-DL concepts as interpolants/explicit definitions.

We introduce expressive Horn DLs (Hustadt, Motik, and Sattler 2005), presented here in the form proposed in (Lutz and Wolter 2012). *Horn- \mathcal{ALCIO}_u -concepts* R and *Horn- \mathcal{ALCIO}_u -CIs* $L \sqsubseteq R$ are defined by the syntax rules

$$R, R' ::= \top \mid \perp \mid A \mid \neg A \mid \{a\} \mid \neg\{a\} \mid R \sqcap R' \mid L \rightarrow R \mid \exists r.R \mid \forall r.R$$

$$L, L' ::= \top \mid \perp \mid A \mid L \sqcap L' \mid L \sqcup L' \mid \exists r.L$$

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.

Theorem 7. *Let $(\mathcal{L}, \mathcal{L}')$ be the pair (Horn- \mathcal{ALCI} , \mathcal{ELI}) or the pair (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.*

Moreover, if an \mathcal{L} -interpolant/explicit definition exists, then there exists one of at most double exponential size that can be computed in double exponential time.

Theorem 7 follows from Theorem 3 and the fact that for any \mathcal{L} -ontology one can construct in polynomial time an \mathcal{L}' -ontology in normal form that is a conservative extension of \mathcal{L} (see (Bienvenu et al. 2016) for a similar result). We next show that despite the fact that Horn- \mathcal{ALCI} -concepts sometimes provide explicit definitions if none exist in \mathcal{ELI} (proof of Theorem 1), they are not sufficient to prove the CIP/PBDP.

Theorem 8. *There exists an ontology \mathcal{O} in Horn- \mathcal{ALC} (and in \mathcal{ELI}), a signature Σ , and a concept name A such that A is implicitly definable using Σ under \mathcal{O} but does not have an explicit Horn- $\mathcal{ALCI}_u(\Sigma)$ -definition.*

Proof. We modify the ontology used in the proof of Point 1 of Theorem 1. Let $\Sigma = \{B, D_1, E, r, r_1\}$ and let \mathcal{O} contain $B \sqcap \exists r.(C \sqcap E) \sqsubseteq A$ and the following CIs:

$$A \sqsubseteq B, \quad B \sqsubseteq \forall r.F, \quad B \sqsubseteq \exists r.C, \quad C \sqsubseteq F \sqcap \forall r_1.D_1,$$

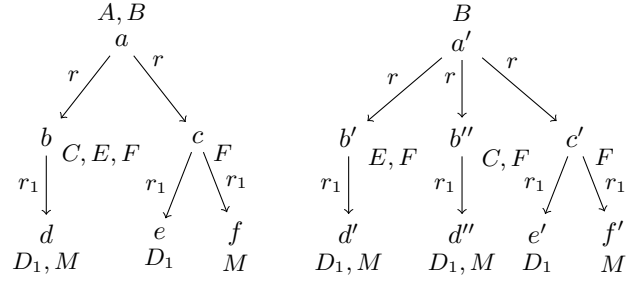


Figure 5: Interpretations \mathcal{I} (left) and \mathcal{I}' (right).

$$F \sqsubseteq \exists r_1.D_1 \sqcap \exists r_1.M, \\ A \sqsubseteq \forall r.((F \sqcap \exists r_1.(D_1 \sqcap M)) \rightarrow E).$$

Intuitively, the final two CIs should be read as

$$F \sqsubseteq \exists r_1.D_1 \\ A \sqsubseteq \forall r.((F \sqcap \forall r_1.D_1) \rightarrow E)$$

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

A is implicitly definable using Σ under \mathcal{O} since

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

To show that A is not explicitly Horn- $\mathcal{ALCI}_u(\Sigma)$ -definable under \mathcal{O} consider the interpretations \mathcal{I} and \mathcal{I}' in Figure 5. The claim follows from the facts that \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 . The latter can be proved by observing that there exists a Horn- $\mathcal{ALCI}_u(\Sigma)$ -simulation between \mathcal{I} and \mathcal{I}' (Jung et al. 2019) containing $(\{a\}, a)$, we refer the reader to the appendix of the full version. To obtain an example in \mathcal{ELI} , it suffices to take a conservative extension of \mathcal{O} in \mathcal{ELI} . \square

9 Discussion

For a few important extensions of $\mathcal{EL}/\mathcal{ELI}$ the complexity of interpolant and explicit definition existence remains to be investigated. Examples include extensions of \mathcal{ELI} with role inclusions, and extensions of \mathcal{EL} or \mathcal{ELI} with functional roles or more general number restrictions. It would also be of interest to investigate interpolant existence if Horn-concepts are admitted as interpolants (using, for example, the games introduced in (Jung et al. 2019)). Finally, the question arises whether there exists at all a decidable Horn language extending, say, Horn- \mathcal{ALCI} , with the CIP/PBDP. We note that Horn-FO enjoys the CIP (Exercise 6.2.6 in (Chang and Keisler 1998)) but is undecidable and that we show in the appendix of the full version that the Horn fragment of the guarded fragment does not enjoy the CIP/PBDP.

Acknowledgments

This research was supported by the EPSRC UK grant EP/S032207/1.

References

- Areces, C.; Blackburn, P.; and Marx, M. 2001. Hybrid logics: Characterization, interpolation and complexity. *J. Symb. Log.* 66(3):977–1010.
- Artale, A.; Jung, J. C.; Mazzullo, A.; Ozaki, A.; and Wolter, F. 2021a. Living without Beth and Craig: Explicit definitions and interpolants in description logics with nominals and role hierarchies. In *Proc. of AAAI*.
- Artale, A.; Mazzullo, A.; Ozaki, A.; and Wolter, F. 2021b. On free description logics with definite descriptions. In *Proc. of KR*.
- Baader, F.; Bienvenu, M.; Lutz, C.; and Wolter, F. 2016. Query and predicate emptiness in ontology-based data access. *J. Artif. Intell. Res.* 56:1–59.
- Baader, F.; Horrocks, I.; Lutz, C.; and Sattler, U. 2017. *An Introduction to Description Logic*. Cambridge University Press.
- Baader, F.; Brandt, S.; and Lutz, C. 2005. Pushing the \mathcal{EL} envelope. In *Proc. of IJCAI*, 364–369.
- Benedikt, M.; Leblay, J.; ten Cate, B.; and Tsamoura, E. 2016. *Generating Plans from Proofs: The Interpolation-based Approach to Query Reformulation*. Synthesis Lectures on Data Management. Morgan & Claypool Publishers.
- Benedikt, M.; Kostylev, E. V.; Mogavero, F.; and Tsamoura, E. 2017. Reformulating queries: Theory and practice. In *IJCAI*, 837–843.
- Bienvenu, M.; Hansen, P.; Lutz, C.; and Wolter, F. 2016. First order-rewritability and containment of conjunctive queries in horn description logics. In *IJCAI*, 965–971.
- Bienvenu, M.; Lutz, C.; and Wolter, F. 2013. First-order rewritability of atomic queries in Horn description logics. In *Proc. of IJCAI*.
- Borgida, A.; Toman, D.; and Weddell, G. E. 2016. On referring expressions in query answering over first order knowledge bases. In *Proc. of KR*, 319–328.
- Chang, C., and Keisler, H. J. 1998. *Model Theory*. Elsevier.
- Deutsch, A.; Popa, L.; and Tannen, V. 2006. Query reformulation with constraints. *SIGMOD Rec.* 35(1):65–73.
- Geleta, D.; Payne, T. R.; and Tamma, V. A. M. 2016. An investigation of definability in ontology alignment. In Blomqvist, E.; Ciancarini, P.; Poggi, F.; and Vitali, F., eds., *Proc. of EKAW*, 255–271.
- Hernich, A.; Lutz, C.; Papacchini, F.; and Wolter, F. 2020. Dichotomies in ontology-mediated querying with the guarded fragment. *ACM Trans. Comput. Log.* 21(3):20:1–20:47.
- Hustadt, U.; Motik, B.; and Sattler, U. 2005. Data complexity of reasoning in very expressive description logics. In *IJCAI*, 466–471.
- Jung, J. C., and Wolter, F. 2021. Living without Beth and Craig: Definitions and interpolants in the guarded and two-variable fragments. In *Proc. of LICS*.
- Jung, J. C.; Papacchini, F.; Wolter, F.; and Zakharyashev, M. 2019. Model comparison games for horn description logics. In *Proc. of LICS*, 1–14. IEEE.
- Jung, J. C.; Lutz, C.; Martel, M.; and Schneider, T. 2020. Conservative extensions in horn description logics with inverse roles. *J. Artif. Intell. Res.* 68:365–411.
- Konev, B.; Lutz, C.; Walther, D.; and Wolter, F. 2009. Formal properties of modularisation. In *Modular Ontologies*, volume 5445 of *Lecture Notes in Computer Science*. Springer. 25–66.
- Konev, B.; Lutz, C.; Ponomaryov, D. K.; and Wolter, F. 2010. Decomposing description logic ontologies. In *Proc. of KR*. AAAI Press.
- Koopmann, P., and Schmidt, R. A. 2015. Uniform interpolation and forgetting for ALC ontologies with aboxes. In *Proc. of AAAI*, 175–181. AAAI Press.
- Lutz, C., and Wolter, F. 2010. Deciding inseparability and conservative extensions in the description logic EL. *J. Symb. Comput.* 45(2):194–228.
- Lutz, C., and Wolter, F. 2011. Foundations for uniform interpolation and forgetting in expressive description logics. In *Proc. of IJCAI*, 989–995. IJCAI/AAAI.
- Lutz, C., and Wolter, F. 2012. Non-uniform data complexity of query answering in description logics. In Brewka, G.; Eiter, T.; and McIlraith, S. A., eds., *Proc. of KR*.
- Lutz, C., and Wolter, F. 2017. The data complexity of description logic ontologies. *Logical Methods in Computer Science* 13(4).
- Lutz, C.; Piro, R.; and Wolter, F. 2011. Description logic tboxes: Model-theoretic characterizations and rewritability. In Walsh, T., ed., *Proc. of IJCAI*, 983–988. IJCAI/AAAI.
- Lutz, C.; Seylan, I.; and Wolter, F. 2012. An automata-theoretic approach to uniform interpolation and approximation in the description logic EL. In *Proc. of KR*. AAAI Press.
- Lutz, C.; Seylan, I.; and Wolter, F. 2019. The data complexity of ontology-mediated queries with closed predicates. *Logical Methods in Computer Science* 15(3).
- Maksimova, L., and Gabbay, D. 2005. *Interpolation and Definability in Modal and Intuitionistic Logics*. Clarendon Press.
- Nikitina, N., and Rudolph, S. 2014. (Non-)succinctness of uniform interpolants of general terminologies in the description logic EL. *Artif. Intell.* 215:120–140.
- Place, T., and Zeitoun, M. 2016. Separating regular languages with first-order logic. *Log. Methods Comput. Sci.* 12(1).
- Seylan, I.; Franconi, E.; and de Bruijn, J. 2009. Effective query rewriting with ontologies over dboxes. In *Proc. of IJCAI*, 923–925.
- Sofronie-Stokkermans, V. 2008. Interpolation in local theory extensions. *Log. Methods Comput. Sci.* 4(4).
- ten Cate, B.; Conradie, W.; Marx, M.; and Venema, Y. 2006. Definitorially complete description logics. In *Proc. of KR*, 79–89. AAAI Press.
- ten Cate, B.; Franconi, E.; and Seylan, I. 2013. Beth definability in expressive description logics. *J. Artif. Intell. Res.* 48:347–414.

- ten Cate, B. 2005. Interpolation for extended modal languages. *J. Symb. Log.* 70(1):223–234.
- Toman, D., and Weddell, G. E. 2011. *Fundamentals of Physical Design and Query Compilation*. Synthesis Lectures on Data Management. Morgan & Claypool Publishers.
- Toman, D., and Weddell, G. E. 2021. FO rewritability for OMQ using both definability and interpolation. In Homola, M.; Ryzhikov, V.; and Schmidt, R. A., eds., *Proc. of DL*. CEUR-WS.org.
- Vardi, M. Y. 1998. Reasoning about the past with two-way automata. In *Proc. of ICALP'98*, 628–641.