Living Without Beth and Craig: Definitions and Interpolants in Description Logics with Nominals and Role Inclusions

Alessandro Artale,[1] Jean Christoph Jung, [2] Andrea Mazzullo,[1] Ana Ozaki,[3] Frank Wolter [4]

[1] Free University of Bozen-Bolzano
[2] University of Hildesheim
[3] University of Bergen
[4] University of Liverpool
{artale,mazzullo}@inf.unibz.it, jungj@uni-hildesheim.de, ana.ozaki@uib.no, wolter@liverpool.ac.uk

Abstract

The Craig interpolation property (CIP) states that an interpolant for an implication exists iff it is valid. The projective Beth definability property (PBDP) states that an explicit definition exists iff a formula stating implicit definability is valid. Thus, the CIP and PBDP transform potentially hard existence problems into deduction problems in the underlying logic. Description Logics with nominals and/or role inclusions do not enjoy the CIP nor PBDP, but interpolants and explicit definitions have many potential applications in ontology engineering and ontology-based data management. In this article we show the following: even without Craig and Beth, the existence of interpolants and explicit definitions is decidable in description logics with nominals and/or role inclusions such as $\mathcal{ALCO}$, $\mathcal{ALCH}$ and $\mathcal{ALCHIO}$. However, living without Craig and Beth makes this problem harder than deduction: we prove that the existence problems become 2ExpTime-complete, thus one exponential harder than validity. The existence of explicit definitions is 2ExpTime-hard even if one asks for a definition of a nominal using any symbol distinct from that nominal, but it becomes ExpTime-complete if one asks for a definition of a concept name using any symbol distinct from that concept name.

Introduction

The *Craig Interpolation Property* (CIP) for a logic $\mathcal{L}$ states that an implication $\varphi \Rightarrow \psi$ is valid in $\mathcal{L}$ iff there exists a formula χ in $\mathcal{L}$ using only the common symbols of φ and ψ such that $\varphi \Rightarrow \chi$ and $\chi \Rightarrow \psi$ are both valid in $\mathcal{L}$. χ is then called an $\mathcal{L}$-interpolant for $\varphi \Rightarrow \psi$. The CIP is generally regarded as one of the most important and useful properties in formal logic (Van Benthem 2008), with numerous applications ranging from formal verification (McMillan 2003), to theory combinations (Cimatti, Griggio, and Sebastiani 2009; Goel, Krstic, and Tinelli 2009; Calvanese et al. 2020) and query reformulation and rewriting in databases (Marx 2007; Toman and Weddell 2011; Benedikt et al. 2016). Description logics (DLs) are no exception (ten Cate et al. 2006; Seylan, Franconi, and de Bruijn 2009; Konev et al. 2010; ten Cate, Franconi, and Seylan 2013; Lutz, Seylan, and Wolter 2019; Jiménez-Ruiz et al. 2016). A particularly important consequence of the CIP in DLs is the *projective Beth definability property* (PBDP), which states that a concept is implicitly definable using a signature Σ of symbols iff it is explicitly definable using Σ. If the concept is a concept name and Σ the set of all symbols distinct from that concept name, then we speak of the (non-projective) Beth definability property (BDP).

The BDP and PBDP have been used in ontology engineering to extract equivalent acyclic terminologies from ontologies (ten Cate et al. 2006; ten Cate, Franconi, and Seylan 2013), they have been investigated in ontology-based data management to equivalently rewrite ontology-mediated queries (Seylan, Franconi, and de Bruijn 2009; Toman and Weddell 2020), and they have been proposed to support the construction of alignments between ontologies (Jiménez-Ruiz et al. 2016). The CIP is often used as a tool to compute explicit definitions (ten Cate et al. 2006; ten Cate, Franconi, and Seylan 2013; Toman and Weddell 2020). It is also the basic logical property that ensures the robust behaviour of ontology modules (Konev et al. 2009). In the form of parallel interpolation it has been investigated in (Konev et al. 2010) to decompose ontologies. In (Lutz, Seylan, and Wolter 2019), it is used to study P/NP dichotomies in ontology-based query answering. The PBDP is also related to the computation of referring expressions in linguistics (Krahmer and van Deemter 2012) and in ontology-based data management (Borgida, Toman, and Weddell 2016). It has been convincingly argued (Borgida, Toman, and Weddell 2017) that very often in applications the individual names used in ontologies are insufficient "to allow humans to figure out what real-world objects they refer to." A natural way to address this problem is to check for such an individual name a whether there exists a concept C not using a that provides an explicit definition of a under the ontology $\mathcal{O}$ and present such a concept C to the human user. Also very recently, it has been observed that strongly separating concepts for positive and negative examples given as data items in a knowledge base can be represented as interpolants, for appropriately defined ontologies and implications (Funk et al. 2019; Jung et al. 2020a,b). Thus, under the approach to DL concept learning proposed in (Fanizzi, d'Amato, and Esposito 2008; Lisi 2012), searching for a solution to the concept learning problem can be reduced to computing an interpolant.

The CIP, PBDP, and BDP are so powerful because intuitively very hard existence questions are reduced to straightforward deduction questions: an interpolant *exists* iff an implication is valid and an explicit definition *exists* iff a straightforward formula stating implicit definability is valid. The existence problems are thus not harder than validity. For example, in the DL $\mathcal{ALC}$, the existence of an interpolant or an explicit definition can be decided in EXPTIME simply because deduction in $\mathcal{ALC}$ is in EXPTIME (and without ontology even in PSPACE).

Unfortunately, the CIP and the PBDP fail to hold for many standard DLs. Particularly important examples of failure are the extension $\mathcal{ALCO}$ of $\mathcal{ALC}$ with nominals, the extension $\mathcal{ALCH}$ of $\mathcal{ALC}$ with role inclusions, and extensions of these with inverse roles and the universal role. To illustrate, even for very simple implications such as $(\{a\} \sqcap \exists r.\{a\}) \sqsubseteq (\{b\} \rightarrow \exists r.\{b\})$ no $\mathcal{ALCO}$-interpolant exists. Moreover, at least for nominals, there is no satisfactory way to extend the expressive power of (expressive) DLs with nominals to ensure the existence of interpolants as validity is undecidable in any extension of $\mathcal{ALCO}$ with the CIP (ten Cate 2005a).

The aim of this paper is to investigate the complexity of deciding the existence of interpolants and explicit definitions for DLs in which this cannot be deduced using the CIP or PBDP. We consider $\mathcal{ALCO}$ and $\mathcal{ALCH}$ and their extensions by inverse roles and/or the universal role. We note that both role inclusions and nominals are part of the OWL 2 DL standard and are used in many real-world ontologies (Whetzel et al. 2011). We prove that the existence of interpolants and the existence of explicit definitions are both 2EXPTIME-complete in all cases, thus confirming the suspicion that these are harder problems than deduction if one has to live without Beth and Craig. For DLs with nominals, the 2EXPTIME lower bound even holds if one asks for an explicit definition of a nominal over the signature containing all symbols distinct from that nominal, a scenario that is of particular interest for the study of referring expressions.

For the BDP, the situation is different as $\mathcal{ALCH}$ and its extensions without nominals enjoy the BDP (ten Cate 2005b; ten Cate et al. 2006). Moreover, while $\mathcal{ALCO}$ and $\mathcal{ALCHO}$ do not enjoy the BDP (ten Cate 2005b; ten Cate et al. 2006), we show here that their extensions with the universal role (ten Cate 2005b; ten Cate et al. 2006) and/or inverse roles do. In fact, despite the fact that $\mathcal{ALCO}$ and $\mathcal{ALCHO}$ do not enjoy the BDP, we show that for all DLs considered in this paper the problem to decide the existence of an explicit definition of a concept name over the signature containing all symbols distinct from that concept name is EXPTIME-complete, thus not harder than deduction.

Detailed proofs are provided in the full version (Artale et al. 2020).

Related Work

The CIP, PBDP, and BDP have been investigated extensively. In addition to the work discussed in the introduction, we mention the investigation of interpolation and definability in modal logic in general (Maksimova and Gabbay 2005) and in hybrid modal logic in particular (Areces, Blackburn, and Marx 2001; ten Cate 2005a). Also related is work on interpolation in guarded logics (Hoogland, Marx, and Otto 1999; Hoogland and Marx 2002; Bárány, Benedikt, and ten Cate 2018; Benedikt, ten Cate, and Vanden Boom 2016, 2015).

Relevant work on Craig interpolation and Beth definability in description logic has been discussed in the introduction. Craig interpolation should not be confused with work on uniform interpolation, both in description logic (Lutz, Seylan, and Wolter 2012; Lutz and Wolter 2011; Nikitina and Rudolph 2014; Koopmann and Schmidt 2015) and in modal logic (Visser et al. 1996; Kowalski and Metcalfe 2019; Iemhoff 2019). 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 hardly been investigated for logics that do not enjoy the CIP or PBDP. Exceptions include linear temporal logic, LTL, for which the decidability of interpolant existence has been shown in (Place and Zeitoun 2016; Henkell 1988; Henkell et al. 2010) and the guarded fragment for which decidability and 3EXPTIME completenss for interpolant existence are shown in (Jung and Wolter 2020). This is in contrast to work on uniform interpolants in description logics which has in fact focused on the existence and computation of uniform interpolants that do not always exist.

Finally, we note that in (Borgida, Toman, and Weddell 2016, 2017; Toman and Weddell 2019), the authors propose the use of referring expressions in a query answering context with weaker DLs. The focus is on using functional roles to generate referring expressions for individuals for which there might not be a denoting individual name at all in the language.

Preliminaries

We use standard DL notation, see (Baader et al. 2017) for details. Let $\mathsf{N_C}$, $\mathsf{N_R}$, and $\mathsf{N_I}$ be mutually disjoint and countably infinite sets of *concept*, *role*, and *individual names*. A *role* is a role name s or an *inverse role* s^-, with s a role name and $(s^-)^- = s$. We use u to denote the *universal role*. A *nominal* takes the form $\{a\}$, with a an individual name. An $\mathcal{ALCIO}^u$*-concept* is defined according to the syntax rule

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

where a ranges over individual names, A over concept names, and r over roles. We use $C \sqcup D$ as abbreviation for $\neg(\neg C \sqcap \neg D)$, $C \rightarrow D$ for $\neg C \sqcup D$, and $\forall r.C$ for $\neg\exists r.(\neg C)$. We use several fragments of $\mathcal{ALCIO}^u$, including $\mathcal{ALCIO}$, obtained by dropping the universal role, $\mathcal{ALCO}^u$, obtained by dropping inverse roles, $\mathcal{ALCO}$, obtained from $\mathcal{ALCO}^u$ by dropping the universal role, and $\mathcal{ALC}$, obtained from $\mathcal{ALCO}$ by dropping nominals. If $\mathcal{L}$ is any of the DLs above, then an *$\mathcal{L}$-concept inclusion ($\mathcal{L}$-CI)* takes the form $C \sqsubseteq D$ with C and D $\mathcal{L}$-concepts. An *$\mathcal{L}$-ontology* is a finite set of $\mathcal{L}$-CIs. We also consider DLs with *role inclusions (RIs)*, expressions of the form $r \sqsubseteq s$, where r and s are roles. As usual, the addition of RIs is indicated by adding the letter $\mathcal{H}$ to the name of the DL, where inverse roles occur in RIs only if the DL admits inverse roles. Thus, for example,

[AtomC]	for all $(d, e) \in S$: $d \in A^{\mathcal{I}}$ iff $e \in A^{\mathcal{J}}$
[AtomI]	for all $(d, e) \in S$: $d = a^{\mathcal{I}}$ iff $e = a^{\mathcal{J}}$
[Forth]	if $(d, e) \in S$ and $(d, d') \in r^{\mathcal{I}}$, then there is a e' with $(e, e') \in r^{\mathcal{J}}$ and $(d', e') \in S$.
[Back]	if $(d, e) \in S$ and $(e, e') \in r^{\mathcal{J}}$, then there is a d' with $(d, d') \in r^{\mathcal{I}}$ and $(d', e') \in S$.

Figure 1: Conditions on $S \subseteq \Delta^{\mathcal{I}} \times \Delta^{\mathcal{J}}$.

$\mathcal{ALCH}$*-ontologies* are finite sets of $\mathcal{ALC}$-CIs and RIs not using inverse roles and $\mathcal{ALCHIO}^u$*-ontologies* are finite sets of $\mathcal{ALCIO}^u$-CIs and RIs. In what follows we use $\mathsf{DL}_{\mathsf{nr}}$ to denote the set of DLs $\mathcal{ALCO}$, $\mathcal{ALCIO}$, $\mathcal{ALCH}$, $\mathcal{ALCHO}$, $\mathcal{ALCHIO}$, and their extensions with the universal role. To simplify notation we do not drop the letter $\mathcal{H}$ when speaking about the concepts and CIs of a DL with RIs. Thus, for example, we sometimes use the expressions $\mathcal{ALCHO}$-concept and $\mathcal{ALCHO}$-CI to denote $\mathcal{ALCO}$-concepts and CIs, respectively.

The semantics is defined in terms of *interpretations* $\mathcal{I} = (\Delta^{\mathcal{I}}, \cdot^{\mathcal{I}})$ as usual, see (Baader et al. 2017). An interpretation $\mathcal{I}$ *satisfies* an $\mathcal{L}$-CI $C \sqsubseteq D$ if $C^{\mathcal{I}} \subseteq D^{\mathcal{I}}$ and an RI $r \sqsubseteq s$ if $r^{\mathcal{I}} \subseteq s^{\mathcal{I}}$. We say that $\mathcal{I}$ is a *model* of an ontology $\mathcal{O}$ if it satisfies all inclusions in it. We say that an inclusion α follows from an ontology $\mathcal{O}$, in symbols $\mathcal{O} \models \alpha$, if every model of $\mathcal{O}$ satisfies α. We write $\mathcal{O} \models C \equiv D$ if $\mathcal{O} \models C \sqsubseteq D$ and $\mathcal{O} \models D \sqsubseteq C$. A concept C is *satisfiable* w.r.t. an ontology $\mathcal{O}$ if there is a model $\mathcal{I}$ of $\mathcal{O}$ with $C^{\mathcal{I}} \neq \emptyset$.

A *signature* Σ is a set of concept, role, and individual names, uniformly referred to as *symbols*. Following standard practice we do not regard the universal role as a symbol but as a logical connective. Thus, the universal role is not contained in any signature. We use $\mathrm{sig}(X)$ to denote the set of symbols used in any syntactic object X such as a concept or an ontology. An $\mathcal{L}(\Sigma)$*-concept* is an $\mathcal{L}$-concept C with $\mathrm{sig}(C) \subseteq \Sigma$.

We next recall model-theoretic characterizations of when nodes in interpretations are indistinguishable by concepts formulated in one of the DLs $\mathcal{L}$ introduced above. A *pointed interpretation* is a pair $\mathcal{I}, d$ with $\mathcal{I}$ an interpretation and $d \in \Delta^{\mathcal{I}}$. For pointed interpretations $\mathcal{I}, d$ and $\mathcal{J}, e$ and a signature Σ, we write $\mathcal{I}, d \equiv_{\mathcal{L},\Sigma} \mathcal{J}, e$ and say that $\mathcal{I}, d$ and $\mathcal{J}, e$ are $\mathcal{L}(\Sigma)$*-equivalent* if $d \in C^{\mathcal{I}}$ iff $e \in C^{\mathcal{J}}$, for all $\mathcal{L}(\Sigma)$-concepts C.

As for the model-theoretic characterizations, we start with $\mathcal{ALC}$. Let Σ be a signature. A relation $S \subseteq \Delta^{\mathcal{I}} \times \Delta^{\mathcal{J}}$ is an $\mathcal{ALC}(\Sigma)$*-bisimulation* if conditions [AtomC], [Forth] and [Back] from Figure 1 hold, where A and r range over all concept and role names in Σ, respectively. We write $\mathcal{I}, d \sim_{\mathcal{ALC},\Sigma} \mathcal{J}, e$ and call $\mathcal{I}, d$ and $\mathcal{J}, e$ $\mathcal{ALC}(\Sigma)$*-bisimilar* if there exists an $\mathcal{ALC}(\Sigma)$-bisimulation S such that $(d, e) \in S$. For $\mathcal{ALCO}$, we define $\sim_{\mathcal{ALCO},\Sigma}$ analogously, but now demand that, in Figure 1, also condition [AtomI] holds for all individual names $a \in \Sigma$. For languages $\mathcal{L}$ with inverse roles, we demand that, in Figure 1, r additionally ranges over inverse roles. For languages $\mathcal{L}$ with the universal role we extend the respective conditions by demanding that the domain $\mathrm{dom}(S)$ and range $\mathrm{ran}(S)$ of S contain $\Delta^{\mathcal{I}}$ and $\Delta^{\mathcal{J}}$, respectively. If a DL $\mathcal{L}$ has RIs, then we use $\mathcal{I}, d \sim_{\mathcal{L},\Sigma} \mathcal{J}, e$ to state that $\mathcal{I}, d \sim_{\mathcal{L}',\Sigma} \mathcal{J}, e$ for the fragment $\mathcal{L}'$ of $\mathcal{L}$ without RIs.

The next lemma summarizes the model-theoretic characterizations for all relevant DLs (Lutz, Piro, and Wolter 2011; Goranko and Otto 2007). For the definition of ω-saturated structures, we refer the reader to (Chang and Keisler 1998).

Lemma 1 *Let $\mathcal{I}, d$ and $\mathcal{J}, e$ be pointed interpretations and ω-saturated. Let $\mathcal{L} \in \mathsf{DL}_{\mathsf{nr}}$ and Σ a signature. Then*

$$\mathcal{I}, d \equiv_{\mathcal{L},\Sigma} \mathcal{J}, e \quad \textit{iff} \quad \mathcal{I}, d \sim_{\mathcal{L},\Sigma} \mathcal{J}, e.$$

For the "if"-direction, the ω-saturatednesses condition can be dropped.

Craig Interpolation and Beth Definability

We introduce the Craig interpolation property (CIP) and the (projective) Beth definability property ((P)BDP) as defined in (ten Cate, Franconi, and Seylan 2013). We recall their relationship and observe that no DL in $\mathsf{DL}_{\mathsf{nr}}$ enjoys the CIP or PBDP while all except $\mathcal{ALCO}$ and $\mathcal{ALCHO}$ enjoy the BDP.

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

We next define the relevant definability notions. Let $\mathcal{O}$ be an ontology and C a concept. Let $\Sigma \subseteq \mathrm{sig}(\mathcal{O}, C)$ be a signature. An $\mathcal{L}(\Sigma)$-concept D is an *explicit $\mathcal{L}(\Sigma)$-definition of C under $\mathcal{O}$* if $\mathcal{O} \models C \equiv D$. We call C *explicitly definable in $\mathcal{L}(\Sigma)$ under $\mathcal{O}$* if there is an explicit $\mathcal{L}(\Sigma)$-definition of C 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 C is called *implicitly definable from Σ under $\mathcal{O}$* if the Σ-reduct of any model $\mathcal{I}$ of $\mathcal{O}$ determines the set $C^{\mathcal{I}}$; in other words, if $\mathcal{I}$ and $\mathcal{J}$ are both models of $\mathcal{O}$ such that $\mathcal{I}_{|\Sigma} = \mathcal{J}_{|\Sigma}$, then $C^{\mathcal{I}} = C^{\mathcal{J}}$. It is easy to see that implicit definability can be reformulated as a standard reasoning problem as follows: a concept C is implicitly definable from Σ under $\mathcal{O}$ iff $\mathcal{O} \cup \mathcal{O}_{\Sigma} \models C \equiv C_{\Sigma}$, where $\mathcal{O}_{\Sigma}$ and C_{Σ} are obtained from $\mathcal{O}$ and, respectively, C by replacing every non-Σ symbol uniformly by a fresh symbol. If a concept 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 2 *A DL $\mathcal{L}$ has the projective Beth definability property (PBDP) if for any $\mathcal{L}$-ontology $\mathcal{O}$, $\mathcal{L}$-concept C, and signature $\Sigma \subseteq \mathrm{sig}(\mathcal{O}, C)$ the following holds: if C is implicitly definable from Σ under $\mathcal{O}$, then C is explicitly $\mathcal{L}(\Sigma)$-definable under $\mathcal{O}$.*

It is known that the CIP and PBDP are tightly linked (ten Cate, Franconi, and Seylan 2013).

Lemma 2 *If $\mathcal{L}$ enjoys the CIP, then $\mathcal{L}$ enjoys the PBDP.*

To see this, assume that an $\mathcal{L}$-concept C is implicitly definable from Σ under an $\mathcal{L}$-ontology $\mathcal{O}$, for some signature Σ. Then $\mathcal{O} \cup \mathcal{O}_\Sigma \models C \equiv C_\Sigma$, with $\mathcal{O}_\Sigma$ and C_Σ as above. Take an $\mathcal{L}$-interpolant D for $C \sqsubseteq C_\Sigma$ under $\mathcal{O} \cup \mathcal{O}_\Sigma$. Then D is an explicit $\mathcal{L}(\Sigma)$-definition of C under $\mathcal{O}$.

An important special case of explicit definability is the explicit definability of a concept name A from $\mathrm{sig}(\mathcal{O}) \setminus \{A\}$ under an ontology $\mathcal{O}$. For this case, we consider the following *non-projective version* of the Beth definability property: A DL $\mathcal{L}$ enjoys the Beth definability property (BDP) if for any $\mathcal{L}$-ontology $\mathcal{O}$ and any concept name A the following holds: if A is implicitly definable from $\mathrm{sig}(\mathcal{O}) \setminus \{A\}$ under $\mathcal{O}$, then A is explicitly $\mathcal{L}(\mathrm{sig}(\mathcal{O}) \setminus \{A\})$-definable under $\mathcal{O}$. Clearly, the PBDP implies the BDP, but not vice versa.

Many DLs including $\mathcal{ALC}$, $\mathcal{ALCI}$, and $\mathcal{ALCI}^u$, enjoy CIP and (P)BDP. However, DLs supporting nominals or role inclusions do not enjoy PBDP and thus, by Lemma 2, also not the CIP (ten Cate, Franconi, and Seylan 2013). The following theorem summarizes the situation.

Theorem 3 *(1) No $\mathcal{L} \in \mathsf{DL}_{\mathsf{nr}}$ enjoys the CIP or the PBDP.*
(2) All $\mathcal{L} \in \mathsf{DL}_{\mathsf{nr}} \setminus \{\mathcal{ALCO}, \mathcal{ALCHO}\}$ enjoy the BDP. $\mathcal{ALCO}$ and $\mathcal{ALCHO}$ do not enjoy the BDP.

To see Part (1) of the theorem, we provide two example ontologies. Consider first the $\mathcal{ALCO}$-ontology

$$\begin{aligned}\mathcal{O}_1 \quad = \quad & \{\ \{a\} \sqsubseteq \exists r.\{a\},\ A \sqcap \neg\{a\} \sqsubseteq \forall r.(\neg\{a\} \rightarrow \neg A),\\ & \neg A \sqcap \neg\{a\} \sqsubseteq \forall r.(\neg\{a\} \rightarrow A)\ \}.\end{aligned}$$

Thus, $\mathcal{O}_1$ implies that a is reflexive and that no node distinct from a is reflexive. Let $\Sigma = \{r, A\}$. Then $\{a\}$ is implicitly definable from Σ under $\mathcal{O}_1$ since $\mathcal{O}_1 \models \forall x((x = a) \leftrightarrow r(x,x))$, but one can show that $\{a\}$ is not explicitly $\mathcal{L}(\Sigma)$-definable under $\mathcal{O}_1$ for any $\mathcal{L} \in \mathsf{DL}_{\mathsf{nr}}$ with nominals.

Consider now the $\mathcal{ALCH}$-ontology

$$\begin{aligned}\mathcal{O}_2 \quad = \quad & \{\ r \sqsubseteq r_1,\ r \sqsubseteq r_2,\ \neg\exists r.\top \sqcap \exists r_1.A \sqsubseteq \forall r_2.\neg A,\\ & \neg\exists r.\top \sqcap \exists r_1.\neg A \sqsubseteq \forall r_2.A\ \}\end{aligned}$$

from (ten Cate, Franconi, and Seylan 2013). Then $\exists r.\top$ is implicitly definable from $\Sigma = \{r_1, r_2\}$ under $\mathcal{O}_2$ as $\mathcal{O}_2 \models \forall x\big(\exists y\, r(x,y) \leftrightarrow \exists y\,(r_1(x,y) \wedge r_2(x,y))\big)$, but it is not explicitly definable from Σ under $\mathcal{O}_2$ in any DL from $\mathsf{DL}_{\mathsf{nr}}$.

Part (2) of Theorem 3 for $\mathcal{L} \in \mathsf{DL}_{\mathsf{nr}}$ without nominals or with the universal role follows from Theorems 2.5.4 and 6.2.4 in (ten Cate 2005b), respectively, see also (ten Cate et al. 2006). That $\mathcal{ALCO}$ and $\mathcal{ALCHO}$ do not enjoy the BDP is shown in (ten Cate et al. 2006). It remains to prove that $\mathcal{ALCIO}$ and $\mathcal{ALCHIO}$ enjoy the BDP. This is done in the appendix using a generalized version of cartesian products called bisimulation products.

Notions Studied and Main Result

The failure of CIP and (P)BDP reported in Theorem 3 motivates the investigation of the respective decision problems of *interpolant existence* and *projective and non-projective definition existence*, which are defined as follows.

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 $\mathcal{O}_1 \cup \mathcal{O}_2$.*

Definition 4 *Let $\mathcal{L}$ be a DL.* Projective $\mathcal{L}$-definition existence *is the problem to decide for any $\mathcal{L}$-ontology $\mathcal{O}$, $\mathcal{L}$-concept C, and signature $\Sigma \subseteq sig(\mathcal{O}, C)$ whether there exists an explicit $\mathcal{L}(\Sigma)$-definition of C under $\mathcal{O}$.* (Non-projective) $\mathcal{L}$-definition existence *is the sub-problem where C ranges only over concept names A and $\Sigma = sig(\mathcal{O}) \setminus \{A\}$.*

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. Similarly, projective definition existence reduces to checking implicit definability for logics with the PBDP but not for logics without the PBDP. Also observe that the following reduction can be proved similarly to the proof of Lemma 2.

Lemma 4 *Let $\mathcal{L}$ be a DL. There is a polynomial time reduction of projective $\mathcal{L}$-definition existence to $\mathcal{L}$-interpolant existence.*

In terms of applications of the introduced decision problems, we note that non-projective definition existence is particularly relevant for the extraction of acyclic terminologies from ontologies (ten Cate et al. 2006), while the flexibility of projective definition existence is useful in most other applications discussed in the introduction. When it comes to computing referring expressions as discussed in the introduction, we are interested in the case when C ranges over nominals $\{a\}$. We then speak of *projective $\mathcal{L}$-referring expression existence* and of *(non-projective) $\mathcal{L}$-referring expression existence*, if $\Sigma = \mathrm{sig}(\mathcal{O}) \setminus \{a\}$.

The main concern of the present paper is to study the computational complexity of the introduced decision problems. As a preliminary step, we provide model-theoretic characterizations for the existence of interpolants and explicit definitions in terms of bisimulations as captured in the following central notion.

Definition 5 (Joint consistency) *Let $\mathcal{L} \in \mathsf{DL}_{\mathsf{nr}}$. Let $\mathcal{O}_1, \mathcal{O}_2$ be $\mathcal{L}$-ontologies, C_1, C_2 be $\mathcal{L}$-concepts, and $\Sigma \subseteq sig(\mathcal{O}_1, \mathcal{O}_2, C_1, C_2)$ be a signature. Then $\mathcal{O}_1, C_1$ and $\mathcal{O}_2, C_2$ are called* jointly consistent modulo $\mathcal{L}(\Sigma)$-bisimulations *if there exist pointed models $\mathcal{I}_1, d_1$ and $\mathcal{I}_2, d_2$ such that $\mathcal{I}_i$ is a model of $\mathcal{O}_i$, $d_i \in C_i^{\mathcal{I}_i}$, for $i = 1, 2$, and $\mathcal{I}_1, d_1 \sim_{\mathcal{L},\Sigma} \mathcal{I}_2, d_2$.*

The associated decision problem, *joint consistency modulo $\mathcal{L}$-bisimulations*, is defined in the expected way. The following result characterizes the existence of interpolants using joint consistency modulo $\mathcal{L}(\Sigma)$-bisimulations. The proof uses Lemma 1.

Theorem 5 *Let $\mathcal{L} \in \mathsf{DL}_{\mathsf{nr}}$. Let $\mathcal{O}_1, \mathcal{O}_2$ be $\mathcal{L}$-ontologies and let C_1, C_2 be $\mathcal{L}$-concepts, and $\Sigma = sig(\mathcal{O}_1, C_1) \cap sig(\mathcal{O}_2, C_2)$. Then the following conditions are equivalent:*

1. *there is no $\mathcal{L}$-interpolant for $C_1 \sqsubseteq C_2$ under $\mathcal{O}_1 \cup \mathcal{O}_2$;*
2. *$\mathcal{O}_1 \cup \mathcal{O}_2, C_1$ and $\mathcal{O}_1 \cup \mathcal{O}_2, \neg C_2$ are jointly consistent modulo $\mathcal{L}(\Sigma)$-bisimulations.*

The following characterization of the existence of explicit definitions is a direct consequence of Theorem 5.

Theorem 6 *Let $\mathcal{L} \in \mathsf{DL}_{\mathsf{nr}}$. Let $\mathcal{O}$ be an $\mathcal{L}$-ontology, C an $\mathcal{L}$-concept, and $\Sigma \subseteq sig(\mathcal{O}, C)$ a signature. Then the following conditions are equivalent:*

1. *there is no explicit $\mathcal{L}(\Sigma)$-definition of C under $\mathcal{O}$;*
2. *$\mathcal{O}, C$ and $\mathcal{O}, \neg C$ are jointly consistent modulo $\mathcal{L}(\Sigma)$-bisimulations.*

In the remainder of the paper we establish the following tight complexity results for the introduced decision problems.

Theorem 7 *Let $\mathcal{L} \in \mathsf{DL}_{\mathsf{nr}}$. Then, (i) $\mathcal{L}$-interpolant existence and projective $\mathcal{L}$-definition existence are* 2EXPTIME*-complete, (ii) if $\mathcal{L}$ admits nominals, then both projective and non-projective $\mathcal{L}$-referring expression existence are* 2EXPTIME*-complete, and (iii) non-projective $\mathcal{L}$-definition existence is* EXPTIME*-complete.*

Observe that the characterizations given in Theorems 5 and 6 provide reductions of interpolant and definition existence to the complement of joint consistency modulo $\mathcal{L}$-bisimulations. Hence for the upper bounds in Points *(i)* and *(ii)*, it suffices to decide the latter in double exponential time which is what we do in the next section. After that, we provide lower bounds for definition existence and referring expression existence which imply the corresponding lower bounds for interpolant existence via Lemma 4. Finally, we show the upper bounds of Point *(iii)*; the lower bounds are inherited from validity.

The 2ExpTime Upper Bound

We provide a double exponential time mosaic-style algorithm that decides joint consistency modulo $\mathcal{L}$-bisimulations, for all $\mathcal{L} \in \mathsf{DL}_{\mathsf{nr}}$.

Theorem 8 *Let $\mathcal{L} \in \mathsf{DL}_{\mathsf{nr}}$. Then joint consistency modulo $\mathcal{L}$-bisimulations is in* 2EXPTIME.

Assume $\mathcal{L} \in \mathsf{DL}_{\mathsf{nr}}$. We may assume that $\mathcal{L}$ extends $\mathcal{ALCHO}$. Consider $\mathcal{L}$-ontologies $\mathcal{O}_1$ and $\mathcal{O}_2$ and $\mathcal{L}$-concepts C_1 and C_2. Let $\Sigma \subseteq \text{sig}(\mathcal{O}_1, \mathcal{O}_2, C_1, C_2)$ be a signature. Let $\Xi = \text{sub}(\mathcal{O}_1, \mathcal{O}_2, C_1, C_2)$ denote the closure under single negation of the set of subconcepts of concepts in $\mathcal{O}_1, \mathcal{O}_2, C_1, C_2$. A Ξ-*type* t is a subset of Ξ such that there exists an interpretation $\mathcal{I}$ and $d \in \Delta^{\mathcal{I}}$ with $t = \text{tp}_\Xi(\mathcal{I}, d)$, where

$$\text{tp}_\Xi(\mathcal{I}, d) = \{C \in \Xi \mid d \in C^{\mathcal{I}}\}$$

is the Ξ-*type realized at d in $\mathcal{I}$*. Let $T(\Xi)$ denote the set of all Ξ-types. Let r be a role. A pair (t_1, t_2) of Ξ-types t_1, t_2 is *r-coherent for $\mathcal{O}_i$*, in symbols $t_1 \rightsquigarrow_{r,\mathcal{O}_i} t_2$, if the following condition holds for all roles s with $\mathcal{O}_i \models r \sqsubseteq s$: (1) if $\neg\exists s.C \in t_1$, then $C \notin t_2$ and (2) if $\neg\exists s^-.C \in t_2$, then $C \notin t_1$. We aim to work with pairs $(T_1, T_2) \in 2^{T(\Xi)} \times 2^{T(\Xi)}$ such that all $t \in T_1 \cup T_2$ are realized in mutually $\mathcal{L}(\Sigma)$-bisimilar nodes of models of $\mathcal{O}_i$, for $i = 1, 2$.

Thus, we now formulate conditions on a set $\mathcal{S} \subseteq 2^{T(\Xi)} \times 2^{T(\Xi)}$ which ensure that one can construct from $\mathcal{S}$ models $\mathcal{I}_i$ of $\mathcal{O}_i$ such that for any pair $(T_1, T_2) \in \mathcal{S}$ and all $t \in T_i$, $i = 1, 2$, there are nodes $d_t \in \Delta^{\mathcal{I}_i}$ realizing t, such that all $d_t, t \in T_1 \cup T_2$ are mutually $\mathcal{L}(\Sigma)$-bisimilar. We lift the definition of r-coherence from pairs of types to pairs of elements of $2^{T(\Xi)} \times 2^{T(\Xi)}$. Let r be a role. We call a pair (T_1, T_2), (T_1', T_2') *r-coherent*, in symbols $(T_1, T_2) \rightsquigarrow_r (T_1', T_2')$, if for $i = 1, 2$ and any $t \in T_i$ there exists a $t' \in T_i'$ such that $t \rightsquigarrow_{r,\mathcal{O}_i} t'$. Moreover, to deal with DLs with inverse roles, we say that (T_1, T_2), (T_1', T_2') are *fully r-coherent*, in symbols $(T_1, T_2) \leftrightsquigarrow_r (T_1', T_2')$ if the converse holds as well: for $i = 1, 2$ and any $t' \in T_i'$ there exists a $t \in T_i$ such that $t \rightsquigarrow_{r,\mathcal{O}_i} t'$.

We first formulate conditions that ensure that nominals are interpreted as singletons and that individuals in Σ are preserved by the bisimulation. Say that $\mathcal{S}$ is *good for nominals* if for every individual name $a \in \text{sig}(\Xi)$ and $i = 1, 2$ there exists exactly one t_a^i with $\{a\} \in t_a^i \in \bigcup_{(T_1,T_2)\in\mathcal{S}} T_i$ and exactly one pair $(T_1, T_2) \in \mathcal{S}$ with $t_a^i \in T_i$. Moreover, if $a \in \Sigma$, then that pair either takes the form $(\{t_a^1\}, \{t_a^2\})$ or the form $(\{t_a^1\}, \emptyset)$ and $(\emptyset, \{t_a^2\})$, respectively.

Secondly, we ensure that the types used in $\mathcal{S}$ are consistent with $\mathcal{O}_1$ and $\mathcal{O}_2$, respectively. Say that $\mathcal{S}$ is *good for $\mathcal{O}_1, \mathcal{O}_2$* if $(\emptyset, \emptyset) \notin \mathcal{S}$ and for every $(T_1, T_2) \in \mathcal{S}$ all types $t \in T_i$ are realizable in a model of $\mathcal{O}_i$, $i = 1, 2$.

Finally, we need to ensure that concept names in Σ are preserved by the bisimulation and that the back and forth condition of bisimulations hold. $\mathcal{S}$ is called $\mathcal{ALCHO}(\Sigma)$-*good* if it is good for nominals and $\mathcal{O}_1, \mathcal{O}_2$, and the following conditions hold:

1. Σ-*concept name coherence*: for any concept name $A \in \Sigma$ and $(T_1, T_2) \in \mathcal{S}$, $A \in t$ iff $A \in t'$ for all $t, t' \in T_1 \cup T_2$;
2. *Existential saturation*: for $i = 1, 2$, if $(T_1, T_2) \in \mathcal{S}$ and $\exists r.C \in t \in T_i$, then there exists $(T_1', T_2') \in \mathcal{S}$ such that (1) there exists $t' \in T_i'$ with $C \in t'$ and $t \rightsquigarrow_{r,\mathcal{O}_i} t'$ and (2) if $\mathcal{O}_i \models r \sqsubseteq s$ with $s \in \Sigma$, then $(T_1, T_2) \rightsquigarrow_s (T_1', T_2')$.

If inverse roles or the universal role are present then we strengthen Condition 2 to Condition $2\mathcal{I}$ and add Condition $3u$, respectively:

$2\mathcal{I}$. Condition 2 with '$s \in \Sigma$' replaced by 's a role over Σ' and '$(T_1, T_2) \rightsquigarrow_s (T_1', T_2')$' replaced by '$(T_1, T_2) \leftrightsquigarrow_s (T_1', T_2')$'.

$3u$. if $(T_1, T_2) \in \mathcal{S}$, then $T_i \neq \emptyset$, for $i = 1, 2$.

Thus, $\mathcal{S}$ is $\mathcal{ALCHIO}(\Sigma)$-*good* if the conditions above hold with Condition 2 replaced by Condition $2\mathcal{I}$ and $\mathcal{S}$ is $\mathcal{ALCHO}^u(\Sigma)$-*good* and, respectively, $\mathcal{ALCHIO}^u(\Sigma)$-*good* if also Condition $3u$ holds.

Lemma 9 *Let $\mathcal{L} \in \mathsf{DL}_{\mathsf{nr}}$. Assume $\mathcal{O}_1, \mathcal{O}_2$ are $\mathcal{L}$-ontologies, C_1, C_2 are $\mathcal{L}$-concepts, and let $\Sigma \subseteq sig(\Xi)$ be a signature. The following conditions are equivalent:*

1. *$\mathcal{O}_1, C_1$ and $\mathcal{O}_2, C_2$ are jointly consistent modulo $\mathcal{L}(\Sigma)$-bisimulations.*
2. *there exists an $\mathcal{L}(\Sigma)$-good set $\mathcal{S}$ and Ξ-types t_1, t_2 with $C_1 \in t_1$ and $C_2 \in t_2$ such that $t_1 \in T_1$ and $t_2 \in T_2$ for some $(T_1, T_2) \in \mathcal{S}$.*

Proof. (sketch) "1 $\Rightarrow$ 2". Let $\mathcal{I}_1, d_1 \sim_{\mathcal{L},\Sigma} \mathcal{I}_2, d_2$ for models $\mathcal{I}_1$ of $\mathcal{O}_1$ and $\mathcal{I}_2$ of $\mathcal{O}_2$ such that d_1, d_2 realize Ξ-types t_1, t_2 and $C_1 \in t_1, C_2 \in t_2$. Define $\mathcal{S}$ by setting

$(T_1, T_2) \in \mathcal{S}$ if there is $d \in \Delta^{\mathcal{I}_i}$ for some $i \in \{1,2\}$ such that

$$T_j = \{\text{tp}_\Xi(\mathcal{I}_j, e) \mid e \in \Delta^{\mathcal{I}_j}, \mathcal{I}_i, d \sim_{\mathcal{L},\Sigma} \mathcal{I}_j, e\},$$

for $j = 1, 2$. One can show that $\mathcal{S}$ is $\mathcal{L}(\Sigma)$-good and satisfies Condition 2.

"2 ⇒ 1". Assume $\mathcal{S}$ is $\mathcal{L}(\Sigma)$-good and we have Ξ-types s_1, s_2 with $C_1 \in s_1$ and $C_2 \in s_2$ such that $s_1 \in S_1$ and $s_2 \in S_2$ for some $(S_1, S_2) \in \mathcal{S}$. If $\mathcal{L}$ does not admit inverse roles, then interpretations $\mathcal{I}_1$ and $\mathcal{I}_2$ witnessing Condition 1 are defined by setting

$$\begin{aligned}
\Delta^{\mathcal{I}_i} &:= \{(t, (T_1, T_2)) \mid t \in T_i \text{ and } (T_1, T_2) \in \mathcal{S}\}\\
r^{\mathcal{I}_i} &:= \{((t, p), (t', p')) \in \Delta^{\mathcal{I}_i} \times \Delta^{\mathcal{I}_i} \mid t \leadsto_{r,\mathcal{O}_i} t',\\
&\qquad \forall s \in \Sigma\, ((\mathcal{O}_i \models r \sqsubseteq s) \Rightarrow p \leadsto_s p')\}\\
A^{\mathcal{I}_i} &:= \{(t, p) \in \Delta^{\mathcal{I}_i} \mid A \in t\}\\
a^{\mathcal{I}_i} &:= (t, (T_1, T_2)) \in \Delta^{\mathcal{I}_i},\ \{a\} \in t \in T_i
\end{aligned}$$

If $\mathcal{L}$ admits inverse roles then replace '$s \in \Sigma$' by 's a role over Σ' and '$p \leadsto_s p'$' by '$p \leftrightsquigarrow_s p'$' in the definition of $r^{\mathcal{I}_i}$. ❑

The following lemma can now be established using a standard recursive bad mosaic elimination procedure.

Lemma 10 *Let $\mathcal{L} \in \mathsf{DL}_{\mathsf{nr}}$. Then it is decidable in double exponential time whether for $\mathcal{L}$-ontologies $\mathcal{O}_1, \mathcal{O}_2$, $\mathcal{L}$-concepts C_1, C_2, and a signature $\Sigma \subseteq sig(\Xi)$ there exists an $\mathcal{S}$ and t_1, t_2 satisfying Condition 2 of Lemma 9.*

Theorem 8 is a direct consequence of Lemmas 9 and 10.

2ExpTime Lower Bounds

We show that for any $\mathcal{L}$ in $\mathsf{DL}_{\mathsf{nr}}$, projective $\mathcal{L}$-definition existence is 2EXPTIME-hard and that, if $\mathcal{L}$ supports nominals, even (non-projective) $\mathcal{L}$-referring expression existence is 2EXPTIME-hard.

DLs with Nominals We start with DLs with nominals. By Theorems 5 and 6, it suffices to prove the following result.

Lemma 11 *Let $\mathcal{L} \in \mathsf{DL}_{\mathsf{nr}}$ admit nominals. It is* 2EXPTIME-*hard to decide for an $\mathcal{L}$-ontology $\mathcal{O}$, individual name b, and signature $\Sigma \subseteq sig(\mathcal{O}) \setminus \{b\}$ whether $\mathcal{O}, \{b\}$ and $\mathcal{O}, \neg\{b\}$ are jointly consistent modulo $\mathcal{L}(\Sigma)$-bisimulations. This is true even if b is the only individual in $\mathcal{O}$ and $\Sigma = sig(\mathcal{O}) \setminus \{b\}$.*

We reduce the word problem for 2^n-space bounded alternating Turing machines, which is known to be 2EXPTIME-hard (Chandra, Kozen, and Stockmeyer 1981). An *alternating Turing machine (ATM)* is a tuple $M = (Q, \Theta, \Gamma, q_0, \Delta)$ where Q is the set of states consisting of *existential* and *universal states*, Θ and Γ are input and tape alphabet, $q_0 \in Q$ is the initial state, and the transition relation Δ makes sure that existential and universal states alternate. We assume binary branching. The acceptance condition of our ATMs is defined in a slightly unusual way, without using accepting states: The ATM accepts if it runs forever on all branches and rejects otherwise. This is without loss of generality, since starting from the standard ATM model, this can be achieved by assuming that the ATM terminates on every input and then modifying it to enter an infinite loop from the accepting state. For a precise definition of ATMs and their acceptance condition, we refer the reader to the appendix.

The idea of the reduction is as follows. Given input word w of length n, we construct an ontology $\mathcal{O}$ such that M accepts w iff $\mathcal{O}, \{b\}$ and $\mathcal{O}, \neg\{b\}$ are jointly consistent modulo $\mathcal{L}(\Sigma)$-bisimulations, where

$$\Sigma = \{r, s, Z, B_\forall, B^1_\exists, B^2_\exists\} \cup \{A_\sigma \mid \sigma \in \Gamma \cup (Q \times \Gamma)\}.$$

We provide the reduction here for $\mathcal{L} = \mathcal{ALCO}$; the modifications required for $\Sigma = \text{sig}(\mathcal{O}) \setminus \{b\}$, inverse roles, and the universal role are given in the appendix. The ontology $\mathcal{O}$ enforces that $r(b, b)$ holds using the CI $\{b\} \sqsubseteq \exists r.\{b\}$. Moreover, any node distinct from b with an r-successor lies on an infinite r-path ρ, enforced by the CIs:

$$\neg\{b\} \sqcap \exists r.\top \sqsubseteq I_s \qquad I_s \sqsubseteq \exists r.\top \sqcap \forall r.I_s$$

Thus, if there exist models $\mathcal{I}$ and $\mathcal{J}$ of $\mathcal{O}$ such that $\mathcal{I}, b^\mathcal{I} \sim_{\mathcal{ALC},\Sigma} \mathcal{J}, d$ for some $d \neq b^\mathcal{J}$ it follows that in $\mathcal{J}$ all nodes on some r-path ρ through d are $\mathcal{ALCO}(\Sigma)$-bisimilar. The situation is illustrated in Figure 2 where dashed egdes represent the enforced bisimulation. In each point of ρ starts an infinite tree along role s that is supposed to mimick the computation of M: a configuration of M is represented by 2^n consecutive elements of this infinite tree and is encoded by the concept names $A_\sigma \in \Sigma$. Moreover, each configuration is labeled by $B_\forall$ (if it is universal) and $B^i_\exists$ (if it is existential; $i \in \{1, 2\}$ refers to the existential choice that is taken). All these trees, called T_* and T_i in Figure 2, have identical Σ-decorations due to the enforced bisimulation.

To coordinate successor configurations, we proceed as follows. Along ρ, a counter counts modulo 2^n using concept names not in Σ. Along the trees T_i, two counters are maintained:

- one counter starting at 0 and counting modulo 2^n to divide the tree in configurations of length 2^n;
- another counter starting at *the value of the counter on ρ* and also counting modulo 2^n.

As on the ith s-tree T_i the second counter starts at all nodes at distances $k \times 2^n - i$, for all $k \geq 1$, we are in the position to coordinate all positions at all successive configurations, using concept names not in Σ.

DLs with Role Inclusions By Theorems 5 and 6, it suffices to prove the following.

Lemma 12 *Let $\mathcal{L} \in \mathsf{DL}_{\mathsf{nr}}$ admit role inclusions. It is* 2EXPTIME-*hard to decide for an $\mathcal{L}$-ontology $\mathcal{O}$, concept C, and signature $\Sigma \subseteq sig(\mathcal{O})$ whether $\mathcal{O}, C$ and $\mathcal{O}, \neg C$ are jointly consistent modulo $\mathcal{L}(\Sigma)$-bisimulations.*

As in the proof of Lemma 11, we reduce the word problem for exponentially space bounded ATMs, using the same ATM model as above. In fact the only difference to the proof of Lemma 11 is the way in which we enforce that exponentially many elements are $\mathcal{L}(\Sigma)$-bisimilar. We show how to achieve this for $\mathcal{L} = \mathcal{ALCH}$ using signature $\Sigma' = (\Sigma \setminus \{r\}) \cup \{r_1, r_2\}$. The symbols in $\Sigma' \cap \Sigma$ play exactly the

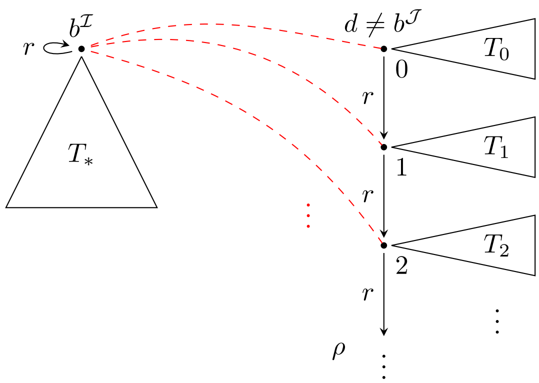

Figure 2: Enforced bisimulation in lower bound

same role as above. The modifications for the inclusion of inverse roles and the universal role are discussed in the appendix. The idea is to construct an ontology $\mathcal{O}'$ such that M accepts w iff $\mathcal{O}', C$ and $\mathcal{O}', \neg C$ are jointly consistent modulo $\mathcal{L}(\Sigma')$-bisimulations, for $C = \exists r^n.\top$.

We replace the nominal b by an r-chain of length n as follows (recall $r \notin \Sigma'$). The ontology $\mathcal{O}'$ contains the RIs $r \sqsubseteq r_1, r \sqsubseteq r_2$ and the CI $\neg\exists r^n.\top \sqcap \exists r_1^n.\top \sqsubseteq R$. The concept name R induces a binary tree T_R of depth n in which each inner node has an r_1- and an r_2-successor, and whose leaves carry counter values from 0 to $2^n - 1$, encoded via non-Σ concept names. To achieve the bisimilar elements, we use that if there exist models $\mathcal{I}$ and $\mathcal{J}$ of $\mathcal{O}'$ and $d \in \Delta^{\mathcal{I}}$, $e \in \Delta^{\mathcal{J}}$ such that

- $d \in (\exists r^n.\top)^{\mathcal{I}}$, $e \in (\neg\exists r^n.\top)^{\mathcal{J}}$;
- $\mathcal{I}, d \sim_{\mathcal{ALC},\Sigma'} \mathcal{J}, e$;

then it follows that in $\mathcal{J}$ there exists a binary tree T_R with root e and of depth n such that all leaves of T_R are $\mathcal{ALC}(\Sigma')$-bisimilar: Due to $\mathcal{I}, d \sim_{\mathcal{ALC},\Sigma'} \mathcal{J}, e$ and $d \in (\exists r_1^n.\top)^{\mathcal{I}}$, we have $e \in (\exists r_1^n.\top)^{\mathcal{J}}$ and thus $e \in R^{\mathcal{I}}$ which starts T_R. Moreover, as $r \sqsubseteq r_i$ for $i = 1, 2$, we have for the r-path starting at d in $\mathcal{I}$ for any r_1/r_2 sequence of length n a corresponding path of length n starting at e in $\mathcal{J}$. Thus, all leaves of T_R are $\mathcal{ALC}(\Sigma')$-bisimilar.

The rest of the proof is as above: $\mathcal{O}'$ enforces that every leaf of T_R is the start of an infinite tree along role s along which the same two counters are maintained; the second counter starts at *the value of the counter on the leaf*. The computation tree of M on input w is encoded as above and the coordination between consecutive configuration is achieved by the availability of the second counter and using non-Σ-symbols.

Non-Projective Definition Existence

We show that for $\mathcal{L}$ in $\mathsf{DL}_{\mathsf{nr}}$, non-projective $\mathcal{L}$-definition existence is in ExpTime. Note that by Theorem 3 (2), $\mathcal{ALCO}$ and $\mathcal{ALCHO}$ are the only DLs in $\mathsf{DL}_{\mathsf{nr}}$ that do not enjoy the BDP. Thus it suffices to consider these two languages.

To motivate our approach, observe that the addition of inverse roles or the universal role to $\mathcal{ALCO}$ or $\mathcal{ALCHO}$ restores the BDP. The following example from (ten Cate et al. 2006) illustrates what is happening: let $\mathcal{O}$ be the ontology containing $A \sqsubseteq \{a\}$, $\{b\} \sqcap B \sqsubseteq \exists r.(\{a\} \sqcap A)$, and $\{b\} \sqcap \neg B \sqsubseteq \exists r.(\{a\} \sqcap \neg A)$. Then A is explicitly definable under $\mathcal{O}$ by both $\{a\} \sqcap \exists r^-.(B \sqcap \{b\})$ and by $\{a\} \sqcap \exists u.(B \sqcap \{b\})$, but A is not explicitly $\mathcal{ALCO}(\{r, B, b, a\})$-definable under $\mathcal{O}$. This example motivates the following characterization, proved using Theorem 6 and bisimulation products. Let $\mathcal{I}, d$ be a pointed model and Σ a signature. Denote by $\Delta^{\mathcal{I}}_{\downarrow d,\Sigma}$ the smallest subset of $\Delta^{\mathcal{I}}$ such that $d \in \Delta^{\mathcal{I}}_{\downarrow d,\Sigma}$ and for all $(e, e') \in r^{\mathcal{I}}$ with r a role name in Σ, if $e \in \Delta^{\mathcal{I}}_{\downarrow d,\Sigma}$, then $e' \in \Delta^{\mathcal{I}}_{\downarrow d,\Sigma}$. The restriction of $\mathcal{I}$ to $\Delta^{\mathcal{I}}_{\downarrow d,\Sigma}$ is denoted $\mathcal{I}_{\downarrow d,\Sigma}$ and called the *interpretation generated by d w.r.t. Σ in $\mathcal{I}$*.

Lemma 13 *Let $\mathcal{O}$ be an $\mathcal{ALCHO}$-ontology, A a concept name, and $\Sigma = sig(\mathcal{O}) \setminus \{A\}$. Then A is not explicitly definable in $\mathcal{ALCHO}(\Sigma)$ under $\mathcal{O}$ iff there are pointed models $\mathcal{I}_1, d$ and $\mathcal{I}_2, d$ such that*

- *$\mathcal{I}_i$ is a model of $\mathcal{O}$, for $i = 1, 2$;*
- *the Σ-reducts of $\mathcal{I}_{1\downarrow d,\Sigma}$ and $\mathcal{I}_{2\downarrow d,\Sigma}$ are identical;*
- *$d \in A^{\mathcal{I}_1}$ and $d \notin A^{\mathcal{I}_2}$.*

Interpretations $\mathcal{I}_1$ and $\mathcal{I}_2$ witnessing that A is not $\mathcal{ALCO}(\{r, B, b, a\})$-definable under the ontology $\mathcal{O}$ defined above are obtained by taking $\Delta^{\mathcal{I}_i} = \{a, b\}$ with a, b interpreting themselves and $r^{\mathcal{I}_i} = \{(b, a)\}$, for $i = 1, 2$, and setting $A^{\mathcal{I}_1} = \{a\}$, $B^{\mathcal{I}_1} = \{b\}$, $A^{\mathcal{I}_2} = B^{\mathcal{I}_2} = \emptyset$. By reformulating the condition given in Lemma 13 as a concept satisfiability problem w.r.t. an $\mathcal{ALCHO}$-ontology, we obtain that non-projective $\mathcal{ALCHO}$-definition existence is in ExpTime, as required. The ExpTime upper bound for $\mathcal{ALCO}$ follows from the ExpTime upper bound for $\mathcal{ALCHO}$ as both languages have the same concept expressions.

Discussion

We have shown that deciding the existence of interpolants and explicit definitions is 2ExpTime-complete for DLs ranging from $\mathcal{ALCO}$ and $\mathcal{ALCH}$ to $\mathcal{ALCHIO}^u$. Our algorithms are not directly applicable in practice to decide the existence of interpolants or explicit definitions nor to compute them if they exist. We are optimistic, however, that the insights from the upper bound proof can be used to design complete tableau-like procedures that extend those in (ten Cate, Franconi, and Seylan 2013). From a theoretical viewpoint, also many interesting questions remain to be explored. For example, what is the size of interpolants and explicit definitions? The techniques introduced in this paper should be a good starting point. Also, while for $\mathcal{ALCHIO}$ and logics with the universal role, the 2ExpTime lower bound holds already under empty ontologies, this remains open for $\mathcal{ALCO}$ and $\mathcal{ALCH}$. Finally, there are many more DLs which do not enjoy the CIP and PBDP and for which the complexity of interpolant and explicit definition existence are open. Examples include expressive languages such as $\mathcal{SHOIQ}$ and Horn DLs with nominals where recent semantic characterizations (Jung et al. 2019) might be helpful.

Acknowledgements

Wolter was supported by EPSRC 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. 2020. Living Without Beth and Craig: Explicit Definitions and Interpolants in Description Logics with Nominals. Available at https://arxiv.org/abs/2007.02736.

Baader, F.; Horrocks, I.; Lutz, C.; and Sattler, U. 2017. *An Introduction to Description Logic*. Cambridge University Press.

Bárány, V.; Benedikt, M.; and ten Cate, B. 2018. Some Model Theory of Guarded Negation. *J. Symb. Log.* 83(4): 1307–1344.

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.; ten Cate, B.; and Vanden Boom, M. 2015. Interpolation with Decidable Fixpoint Logics. In *Proc. of LICS*, 378–389. IEEE Computer Society.

Benedikt, M.; ten Cate, B.; and Vanden Boom, M. 2016. Effective Interpolation and Preservation in Guarded Logics. *ACM Trans. Comput. Log.* 17(2): 8:1–8:46.

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.

Borgida, A.; Toman, D.; and Weddell, G. E. 2017. Concerning Referring Expressions in Query Answers. In *Proc. of IJCAI*, 4791–4795.

Calvanese, D.; Ghilardi, S.; Gianola, A.; Montali, M.; and Rivkin, A. 2020. Combined Covers and Beth Definability. In *Proc. of IJCAR (I)*, 181–200.

Chandra, A. K.; Kozen, D. C.; and Stockmeyer, L. J. 1981. Alternation. *J. ACM* 28: 114–133.

Chang, C.; and Keisler, H. J. 1998. *Model Theory*. Elsevier.

Cimatti, A.; Griggio, A.; and Sebastiani, R. 2009. Interpolant Generation for UTVPI. In *Proc. of CADE*, 167–182. Springer.

Fanizzi, N.; d'Amato, C.; and Esposito, F. 2008. DL-FOIL Concept Learning in Description Logics. In *Proc. of ILP*, 107–121.

Funk, M.; Jung, J. C.; Lutz, C.; Pulcini, H.; and Wolter, F. 2019. Learning Description Logic Concepts: When can Positive and Negative Examples be Separated? In *Proc. of IJCAI*, 1682–1688.

Goel, A.; Krstic, S.; and Tinelli, C. 2009. Ground Interpolation for Combined Theories. In *Proc. of CADE*, 183–198. Springer.

Goranko, V.; and Otto, M. 2007. Model theory of modal logic. In *Handbook of Modal Logic*, 249–329. Elsevier.

Henkell, K. 1988. Pointlike sets: the finest aperiodic cover of a finite semigroup. *J. Pure Appl. Algebra* 55(1-2): 85–126.

Henkell, K.; Rhodes, J.; ; and Steinberg, B. 2010. Aperiodic pointlikes and beyond. *Internat. J. Algebra Comput.* 20(2): 287–305.

Hoogland, E.; and Marx, M. 2002. Interpolation and Definability in Guarded Fragments. *Studia Logica* 70(3): 373–409.

Hoogland, E.; Marx, M.; and Otto, M. 1999. Beth Definability for the Guarded Fragment. In *Proc. of LPAR*, Lecture Notes in Computer Science, 273–285. Springer.

Iemhoff, R. 2019. Uniform interpolation and the existence of sequent calculi. *Annals of Pure and Applied Logic* 170(11): 102711.

Jiménez-Ruiz, E.; Payne, T. R.; Solimando, A.; and Tamma, V. A. M. 2016. Limiting Logical Violations in Ontology Alignnment Through Negotiation. In *Proc. of KR*, 217–226. AAAI Press.

Jung, J. C.; Lutz, C.; Pulcini, H.; and Wolter, F. 2020a. Logical Separability of Incomplete Data under Ontologies. In *Proc. of KR*.

Jung, J. C.; Lutz, C.; Pulcini, H.; and Wolter, F. 2020b. Separating Positive and Negative Data Examples by Concepts and Formulas: The Case of Restricted Signature. Available at https://arxiv.org/abs/2007.02669.

Jung, J. C.; Papacchini, F.; Wolter, F.; and Zakharyaschev, M. 2019. Model Comparison Games for Horn Description Logics. In *Proc. of LICS*, 1–14. IEEE.

Jung, J. C.; and Wolter, F. 2020. Living without Beth and Craig: Explicit Definitions and Interpolants in the Guarded Fragment. Available at http://arxiv.org/abs/2007.01597.

Konev, B.; Lutz, C.; Ponomaryov, D. K.; and Wolter, F. 2010. Decomposing Description Logic Ontologies. In *Proc. of KR*. AAAI Press.

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*, 25–66. Springer.

Koopmann, P.; and Schmidt, R. A. 2015. Uniform Interpolation and Forgetting for ALC Ontologies with ABoxes. In *Proc. of AAAI*, 175–181. AAAI Press.

Kowalski, T.; and Metcalfe, G. 2019. Uniform interpolation and coherence. *Annals of pure and applied logic* 170(7): 825–841.

Krahmer, E.; and van Deemter, K. 2012. Computational Generation of Referring Expressions: A Survey. *Computational Linguistics* 38(1): 173–218.

Lisi, F. A. 2012. A Formal Characterization of Concept Learning in Description Logics. In *Proc. of DL*.

Lutz, C.; Piro, R.; and Wolter, F. 2011. Description Logic TBoxes: Model-Theoretic Characterizations and Rewritability. In *Proc. of IJCAI*.

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).

Lutz, C.; and Wolter, F. 2011. Foundations for Uniform Interpolation and Forgetting in Expressive Description Logics. In *Proc. of IJCAI*, 989–995. IJCAI/AAAI.

Maksimova, L.; and Gabbay, D. 2005. *Interpolation and Definability in Modal and Intuitionistic Logics*. Clarendon Press.

Marx, M. 2007. Queries Determined by Views: Pack Your Views. In *Proc. of PODS*, 23–30. ACM. ISBN 9781595936851.

McMillan, K. L. 2003. Interpolation and SAT-Based Model Checking. In *Proc. of CAV*, 1–13. Springer.

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.

ten Cate, B. 2005a. Interpolation for extended modal languages. *J. Symb. Log.* 70(1): 223–234.

ten Cate, B. 2005b. *Model theory for extended modal languages*. Ph.D. thesis, University of Amsterdam. ILLC Dissertation Series DS-2005-01.

ten Cate, B.; Conradie, W.; Marx, M.; and Venema, Y. 2006. Definitorially Complete Description Logics. In *Proc. of KR*, 79–89.

ten Cate, B.; Franconi, E.; and Seylan, I. 2013. Beth Definability in Expressive Description Logics. *J. Artif. Intell. Res.* 48: 347–414.

Toman, D.; and Weddell, G. 2020. First Order Rewritability for Ontology Mediated Querying in Horn-DLFD. In *Proc. of Description Logic Workshop*. CEUR-WS.org.

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. 2019. Identity Resolution in Ontology Based Data Access to Structured Data Sources. In *Proc. of PRICAI*, 473–485.

Van Benthem, J. 2008. The many faces of interpolation. *Synthese* 451–460.

Visser, A.; et al. 1996. Uniform interpolation and layered bisimulation. In *Gödel'96: Logical foundations of mathematics, computer science and physics—Kurt Gödel's legacy*, 139–164. Association for Symbolic Logic.

Whetzel, P. L.; Noy, N. F.; Shah, N. H.; Alexander, P. R.; Nyulas, C.; Tudorache, T.; and Musen, M. A. 2011. BioPortal. *Nucleic acids research* 39(suppl 2): W541–W545. See also http://krr-nas.cs.ox.ac.uk/ontologies/readme.htm.