

# Living Without Beth and Craig: Explicit Definitions and Interpolants in Description Logics with Nominals

Alessandro Artale<sup>1</sup>, Jean Christoph Jung<sup>2</sup>, Andrea Mazzullo<sup>1</sup>, Ana Ozaki<sup>3</sup>,  
and Frank Wolter<sup>4</sup>

<sup>1</sup> Free University of Bozen-Bolzano      <sup>2</sup> University of Bremen  
<sup>3</sup> University of Bergen      <sup>4</sup> University of Liverpool

**Abstract.** The Craig interpolation property (CIP) states that an interpolant for an implication exists iff it is valid. The Beth definability property (BDP) states that an explicit definition exists iff a formula stating implicit definability is valid. Thus, they transform potentially hard existence problems into deduction problems in the underlying logic. Description Logics with nominals do not have the CIP nor BDP, but in particular, deciding and computing explicit definitions of concepts or individuals has many potential applications in ontology engineering and ontology-based data management. In this article we show two main results: even without Craig and Beth, the existence of interpolants and explicit definitions is decidable in the description logics with nominals  $\mathcal{ALCCO}$  and  $\mathcal{ALCZO}$ . 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.

## 1 Introduction

The *Craig Interpolation Property* (CIP) for first-order logic (FO) states that an implication  $\varphi \Rightarrow \psi$  is valid in FO iff there exists a formula  $\chi$  in FO using only the common symbols of  $\varphi$  and  $\psi$  such that  $\varphi \Rightarrow \chi$  and  $\chi \Rightarrow \psi$  are both valid.  $\chi$  is then called an interpolant for  $\varphi \Rightarrow \psi$ . The CIP of FO and numerous other logics is generally regarded as one of the most important and useful results in formal logic, with numerous applications [43]. Description logics (DLs) are no exception; indeed, the CIP has been intensively investigated [10,40,27,11,33,21]. A particularly important consequence of the CIP is the *Beth definability property* (BDP), which states that a relation or constant is implicitly definable iff it is explicitly definable. In other words, a relation or constant is uniquely determined by a theory iff there exists a definition for it in that theory.

The BDP has been used in ontology engineering to extract equivalent acyclic terminologies from ontologies [10,11], it has been investigated in ontology-based data management to equivalently rewrite ontology-mediated queries [40], and it has been proposed to support the construction of alignments between ontologies [21]. The CIP is often used as a tool to compute explicit definitions [10,11].

It is also the basic logical property that ensures the robust behaviour of ontology modules [26]. In the form of parallel interpolation it has been investigated in [27] to decompose ontologies. In [33], it is used to study P/NP dichotomies in ontology-based query answering. The BDP is also related to the computation of referring expressions in linguistics [30] and in ontology-based data management [7]. In this case, the focus is on computing an explicit definition (or description) for an individual rather than for arbitrary concepts. More recently, it has been observed that the CIP is closely related to the existence of strongly separating concepts for positive and negative examples given as data items in a knowledge base [15,22,23].

The CIP 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 BDP do not always hold. Particularly important examples of failure are DLs with nominals (or, equivalently, hybrid modal logics that add nominals to propositional modal logic). The CIP and BDP fail massively in these cases as even for very simple implications such as  $(\{a\} \sqcap \exists r.\{a\}) \sqsubseteq (\{b\} \rightarrow \exists r.\{b\})$  no interpolant exists. Moreover, 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 [9].

The aim of this paper is to start an investigation of the complexity of deciding the existence of interpolants and explicit definitions for DLs in which this cannot be deduced using the CIP or BDP. We start by considering  $\mathcal{ALCO}$  and its extensions by inverse roles and/or the universal role and prove that the existence of interpolants and the existence of explicit definitions are both 2EXPTIME complete, thus confirming the suspicion that these are much harder problems than deduction if one has to live without Beth and Craig.

The upper bound proof is based on a straightforward characterization of the non-existence of interpolants by the existence of certain bisimulations between pointed models. We then pursue a mosaic based approach by introducing mosaics that are sets of types over the input ontologies/concepts which can be satisfied in bisimilar nodes. Natural constraints for sets of such mosaics characterize when they can be linked together to construct, simultaneously, models of the input ontologies and concepts and an appropriate bisimulation between them. The double exponential upper bound is then naturally explained by the observation that there are double exponentially many mosaics. Formally, the lower bound is proved by a reduction of the word problem for exponentially space-bounded alternating Turing machines.

## 2 Related Work

The CIP and the BDP have been investigated extensively. They have found applications in formal verification [37], theory combinations [14,16], and in database theory for query rewriting under views [36] and query reformulation and compilation [41,6]. Of particular relevance for this work is the investigation of interpolation and definability in modal logic in general [35] and in hybrid modal logic in particular [1,9]. Also related is work on interpolation in guarded logics [19,18,3,5,4].

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 [32,34,38,28] and in modal logic [44,29,20]. 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. It seems that interpolant and explicit definition existence have not yet been investigated for logics that do not enjoy the CIP or BDP. 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 briefly discuss the relationship to computing referring expressions. It has been convincingly argued [8] 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.” Examples include generated ref expressions in object oriented databases and blank node identifiers in RDF. 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  $\mathcal{O}$  (in symbols  $\mathcal{O} \models \{a\} \equiv C$ ) and present such a concept  $C$  to the human user. In [7,8,42], 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.

## 3 Preliminaries

Let  $N_C$ ,  $N_R$ , and  $N_I$  be mutually disjoint and countably infinite sets of *concept names*, *role names*, 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

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

where  $a$  ranges over individual names,  $A$  over concept names, and  $s$  over role names. We use  $C \sqcup D$  as abbreviation for  $\neg(\neg C \sqcap \neg D)$  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

[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]	dual of [Forth]

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

universal role,  $\mathcal{ALCO}^u$ , obtained by dropping inverse roles, and  $\mathcal{ALCO}$ , which is obtained by dropping the universal role and inverse roles.

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

A *signature*  $\Sigma$  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  $\text{sig}(X)$  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  $\text{sig}(C) \subseteq \Sigma$ . The *size* of a (finite) syntactic object  $X$ , denoted  $\|X\|$ , is the number of symbols needed to write it, with every occurrence of a concept and role name contributing one.

The semantics is defined in terms of *interpretations*  $\mathcal{I} = (\Delta^{\mathcal{I}}, \cdot^{\mathcal{I}})$  as usual, see [2]. An interpretation  $\mathcal{I}$  *satisfies* a CI  $C \sqsubseteq D$  if  $C^{\mathcal{I}} \subseteq D^{\mathcal{I}}$ . We say that  $\mathcal{I}$  is a *model* of an ontology  $\mathcal{O}$  if it satisfies all inclusions in it. We say that a CI  $C \sqsubseteq D$  follows from an ontology  $\mathcal{O}$ , in symbols  $\mathcal{O} \models C \sqsubseteq D$ , if every model of  $\mathcal{O}$  satisfies  $C \sqsubseteq D$ . 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  $C^{\mathcal{I}} \neq \emptyset$ .

We associate every interpretation  $\mathcal{I}$  with an undirected graph  $G_{\mathcal{I}} = (V, E)$  where  $V = \Delta^{\mathcal{I}}$  and  $E = \{\{d, e\} \mid (d, e) \in s^{\mathcal{I}} \text{ for some } s \in \mathbf{N}_{\mathbf{R}}\}$ . We say that  $\mathcal{I}$  is a *tree* if  $G_{\mathcal{I}}$  is a tree (without self loops) and there are no multi-edges, that is,  $(d, e) \in s_1^{\mathcal{I}}$  implies  $(d, e) \notin s_2^{\mathcal{I}}$  for all distinct roles  $s_1, s_2$ .

We next recall model-theoretic characterizations of when elements 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  $\Sigma$ , 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  $\mathcal{ALCO}$ . Let  $\Sigma$  be a signature. A relation  $S \subseteq \Delta^{\mathcal{I}} \times \Delta^{\mathcal{J}}$  is an  $\mathcal{ALCO}(\Sigma)$ -*bisimulation* if conditions [AtomC], [AtomI], [Forth] and [Back] from Figure 1 hold, where  $A$  ranges over all concept names in  $\Sigma$ ,  $a$  over all individual names in  $\Sigma$ , and  $r$  over all role names in  $\Sigma$ . We write  $\mathcal{I}, d \sim_{\mathcal{ALCO}, \Sigma} \mathcal{J}, e$  and call  $\mathcal{I}, d$  and  $\mathcal{J}, e$   $\mathcal{ALCO}(\Sigma)$ -*bisimilar* if there exists an  $\mathcal{ALCO}(\Sigma)$ -bisimulation  $S$  such that  $(d, e) \in S$ . For  $\mathcal{ALCIO}$ , we

define  $\sim_{\mathcal{ALC}\mathcal{IO},\Sigma}$  analogously, but now demand that in Figure 1  $r$  additionally ranges over inverse roles. For  $\mathcal{ALC}\mathcal{O}^u$  and  $\mathcal{ALC}\mathcal{IO}^u$  we extend the respective conditions by demanding that the domain  $\text{dom}(S)$  and range  $\text{ran}(S)$  of  $S$  contain  $\Delta^{\mathcal{I}}$  and  $\Delta^{\mathcal{J}}$ , respectively.

The next lemma summarizes the model-theoretic characterizations for all relevant DLs [31,17]. For  $\omega$ -saturated structures, we refer the reader to [13].

**Lemma 1.** *Let  $\mathcal{I}, d$  and  $\mathcal{J}, e$  be pointed interpretations and  $\omega$ -saturated. Let  $\mathcal{L} \in \{\mathcal{ALC}\mathcal{O}, \mathcal{ALC}\mathcal{IO}, \mathcal{ALC}\mathcal{O}^u, \mathcal{ALC}\mathcal{IO}^u\}$  and  $\Sigma$  a signature. Then*

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

*For the “if”-direction, the  $\omega$ -saturatedness condition can be dropped.*

## 4 Notions Studied and Main Result

We introduce the Craig interpolation property and Beth definability property as defined in [11] and observe that none of the DLs considered in this article enjoys any of the two properties. We then introduce *interpolant existence* and *explicit definition existence*, the decision problems studied in this paper. We also remind the reader of the relationship between the Craig interpolation property and the Beth definability property and introduce a few instances of explicit definability existence that are of particular interest. Finally we formulate the main result of this article.

Let  $\mathcal{O}_1, \mathcal{O}_2$  be ontologies and let  $C_1, C_2$  be  $\mathcal{L}$ -concepts. We set  $\text{sig}(\mathcal{O}, C) = \text{sig}(\mathcal{O}) \cup \text{sig}(C)$ , 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

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

If  $\mathcal{O}_1 = \mathcal{O}_2 = \emptyset$ , then we drop  $\mathcal{O}_1$  and  $\mathcal{O}_2$  and speak of an  $\mathcal{L}$ -interpolant for  $C_1 \sqsubseteq C_2$ .

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

The example given in the introduction was first observed in [9] and shows that none of the DLs introduced above has the CIP. This is true even for empty ontologies. We are interested in the following decision problem.

**Definition 2.** *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  $\mathcal{O}_1 \cup \mathcal{O}_2$ .*

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.

*Remark 1.* Observe that if  $\mathcal{L}$  admits the universal role, then  $\mathcal{L}$ -interpolant existence can be reduce in polynomial time to  $\mathcal{L}$ -interpolant existence under empty ontologies. To see this, consider any input  $\mathcal{O}_1, \mathcal{O}_2, C_1, C_2$ . Then let

$$D_i = \forall u. \left( \bigwedge_{F_1 \sqsubseteq F_2 \in \mathcal{O}_i} (F_1 \rightarrow F_2) \right) \rightarrow C_i,$$

for  $i = 1, 2$ . Then any  $\mathcal{L}$ -concept  $D$  is an  $\mathcal{L}$ -interpolant for  $C_1 \sqsubseteq C_2$  under  $\mathcal{O}_1 \cup \mathcal{O}_2$  iff  $D$  is an  $\mathcal{L}$ -interpolant for  $D_1 \sqsubseteq D_2$ . It also follows that  $\mathcal{L}$  enjoys the CIP iff it enjoys the CIP for CIs under empty ontologies.

We next define the relevant definability notions. Let  $\mathcal{O}$  be an ontology and  $C$  a concept. Let  $\Sigma \subseteq \text{sig}(\mathcal{O}, C)$  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 concept  $C$  is called *implicitly definable from  $\Sigma$  under  $\mathcal{O}$*  if the  $\Sigma$ -reduct of any model  $\mathcal{I}$  of  $\mathcal{O}$  determines the set  $C^{\mathcal{I}}$  or, equivalently, if  $\mathcal{O} \cup \mathcal{O}_\Sigma \models C \equiv C_\Sigma$ , where  $\mathcal{O}_\Sigma$  and  $C_\Sigma$  are obtained from  $\mathcal{O}$  and, respectively,  $C$  by replacing every non- $\Sigma$  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  $\Sigma$  under  $\mathcal{O}$ , for any language  $\mathcal{L}$ . A logic enjoys the Beth definability property if the converse implication holds as well:

**Definition 3.** A DL has the Beth definable property (BDP) if for any  $\mathcal{L}$ -ontology  $\mathcal{O}$ ,  $\mathcal{L}$ -concept  $C$ , and signature  $\Sigma \subseteq \text{sig}(\mathcal{O}, C)$  the following holds: if  $C$  is implicitly definable from  $\Sigma$  under  $\mathcal{O}$ , then  $C$  is explicitly  $\mathcal{L}(\Sigma)$ -definable under  $\mathcal{O}$ .

*Remark 2.* Observe that the CIP implies the BDP. To see this, assume that an  $\mathcal{L}$ -ontology  $\mathcal{O}$ ,  $\mathcal{L}$ -concept  $C$ , and a signature  $\Sigma$  are given, and that  $C$  is implicitly definable from  $\Sigma$  under  $\mathcal{O}$ . Then  $\mathcal{O} \cup \mathcal{O}_\Sigma \models C \equiv C_\Sigma$ , with  $\mathcal{O}_\Sigma$  and  $C_\Sigma$  as defined 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  $\mathcal{O}$ .

For DLs extending  $\mathcal{ALC}$  with inverse roles, transitive roles, functional roles, role hierarchies, and combinations thereof, a full classification as to whether they enjoy the CIP and BDP or not has been established [11]. It turns out that among those DLs precisely the DLs without role hierarchies have both the CIP and BDP and those with role hierarchies do not have either. It follows that  $\mathcal{L}$ -interpolant existence and explicit  $\mathcal{L}$ -definition existence (defined below) are in EXPTIME for DLs ranging from  $\mathcal{ALC}$  to  $\mathcal{SZF}$ , the extension of  $\mathcal{ALC}$  with inverse roles, transitive roles, and functional roles.

The following example shows that none of the DLs with nominals considered in this paper has the BDP and that this holds even if the signature  $\Sigma$  is restricted to  $\Sigma = \text{sig}(\mathcal{O}) \setminus \{A\}$ , where  $A$  is a concept name or nominal.

*Example 1.* Let  $\mathcal{O}$  be defined as

$$\mathcal{O} = \{ \{a\} \sqsubseteq \exists r. \{a\}, A \sqcap \neg \{a\} \sqsubseteq \forall r. (\neg \{a\} \rightarrow \neg A), \neg A \sqcap \neg \{a\} \sqsubseteq \forall r. (\neg \{a\} \rightarrow A) \}$$

Thus,  $\mathcal{O}$  implies that  $a$  is reflexive and that no node distinct from  $a$  is reflexive. Let  $\Sigma = \{r, A\}$ . Then  $\{a\}$  is implicitly definable from  $\Sigma$  under  $\mathcal{O}$  since  $\mathcal{O} \models \forall x((x = a) \leftrightarrow r(x, x))$ .  $\{a\}$  is not explicitly  $\mathcal{L}(\Sigma)$ -definable under  $\mathcal{O}$  for any DL  $\mathcal{L}$  between  $\mathcal{ALCO}$  and  $\mathcal{ALCIO}^u$ : consider the model  $\mathcal{I}$  with  $\Delta^{\mathcal{I}} = \{c, d\}$ ,  $a^{\mathcal{I}} = c$ ,  $r^{\mathcal{I}} = \{(c, c), (c, d), (d, c)\}$ ,  $A^{\mathcal{I}} = \{c, d\}$ . Then  $\mathcal{I}$  is a model of  $\mathcal{O}$  and the relation  $S = \Delta^{\mathcal{I}} \times \Delta^{\mathcal{I}}$  is an  $\mathcal{ALCIO}^u(\Sigma)$ -bisimulation on  $\mathcal{I}$ . Thus, Lemma 1 implies

$$\mathcal{I}, c \equiv_{\mathcal{ALCIO}^u, \Sigma} \mathcal{I}, d$$

and there is no explicit  $\mathcal{ALCIO}^u(\Sigma)$ -definition for  $\{a\}$  under  $\mathcal{O}$  as any such definition would apply to  $d$  as well.

We are interested in the following decision problem.

**Definition 4.** Let  $\mathcal{L}$  be a DL. Explicit  $\mathcal{L}$ -definition existence is the problem to decide for any  $\mathcal{L}$ -ontology  $\mathcal{O}$ ,  $\mathcal{L}$ -concept  $C$ , and signature  $\Sigma \subseteq \text{sig}(\mathcal{O}, C)$  whether there exists an explicit  $\mathcal{L}(\Sigma)$ -definition of  $C$  under  $\mathcal{O}$ .

Observe that explicit definition existence reduces to checking implicit definability for logics with the BDP but that this is not the case for logics without the BDP. Also observe that the following reduction is a direct consequence of the argument presented in Remark 2.

**Lemma 2.** Let  $\mathcal{L}$  be a DL. There is a polynomial time reduction of explicit  $\mathcal{L}$ -definition existence to  $\mathcal{L}$ -interpolant existence.

If  $C$  ranges over concept names  $A$  and  $\Sigma = \text{sig}(\mathcal{O}) \setminus \{A\}$ , then we speak of *narrow explicit  $\mathcal{L}$ -definition existence*. If  $C$  ranges over nominals  $\{a\}$ , then we speak of  *$\mathcal{L}$ -referring expression existence* and if, in addition,  $\Sigma = \text{sig}(\mathcal{O}) \setminus \{a\}$ , then we speak of *narrow  $\mathcal{L}$ -referring expression existence*. The following example illustrates explicit definability of nominals.

*Example 2.* Consider the  $\mathcal{ALCIO}$  ontology  $\mathcal{O}$ , about detectives and spies, that consists of the following CIs ( $\mathcal{O}$  is a variant of an ontology introduced in [39]):

$$\begin{aligned} \mathcal{O} = \{ & \exists \text{suspects}.\top \sqsubseteq \text{Detective}, \text{Detective} \sqsubseteq \forall \text{deceives}.\perp, \text{Detective} \sqsubseteq \neg \text{Spy}, \\ & \text{Detective} \equiv \{d_1\} \sqcup \{d_2\} \sqcup \{d_3\}, \{s_1\} \sqsubseteq \neg \text{Spy}, \{s_4\} \sqsubseteq \text{Spy}, \\ & \{s_1\} \sqsubseteq \exists \text{deceives}.\{s_2\}, \{s_2\} \sqsubseteq \exists \text{deceives}.\{s_3\}, \{s_3\} \sqsubseteq \exists \text{deceives}.\{s_4\}, \\ & \{s_4\} \sqsubseteq \forall \text{deceives}^{\neg}.\text{Spy}, \{d_1\} \sqsubseteq \forall \text{suspects}.\{s_1\}, \{d_3\} \sqsubseteq \forall \text{suspects}.\{s_4\}, \\ & \{d_2\} \sqsubseteq \exists \text{suspects}.\{s_2\} \sqcap \exists \text{suspects}.\{s_3\} \\ & \}. \end{aligned}$$

Reasoning by cases, it can be seen that

$$\mathcal{O} \models \{d_2\} \equiv \exists \text{suspects}.\{ \text{Spy} \sqcap \exists \text{deceives}^{\neg}.\neg \text{Spy} \},$$

thus, for  $\Sigma = \{\text{Spy}, \text{suspects}, \text{deceives}\}$ , there is an explicit  $\mathcal{ALCIO}(\Sigma)$  definition of  $\{d_2\}$  under  $\mathcal{O}$ . Another definition of  $\{d_2\}$  under  $\mathcal{O}$  is given by  $\exists \text{suspects}.\{s_2\} \sqcap \exists \text{suspects}.\{s_3\}$ . On the other hand, for  $\Sigma' = \{\text{suspects}\}$ , there does not exist any explicit  $\mathcal{ALCIO}(\Sigma')$  definition of  $\{d_2\}$  under  $\mathcal{O}$ .

In the rest of the paper we establish tight complexity results for the introduced decision problems. In particular, we show the following theorem.

**Theorem 1.** *Let  $\mathcal{L} \in \{\mathcal{ALCCO}, \mathcal{ALCCIO}, \mathcal{ALCCO}^u, \mathcal{ALCCIO}^u\}$ . Then the problems  $\mathcal{L}$ -interpolant existence, (narrow) explicit  $\mathcal{L}$ -definition existence, and (narrow)  $\mathcal{L}$ -referring expression existence are all 2EXPTIME-complete.*

## 5 Model-theoretic Characterizations

We provide model-theoretic characterizations of the (non-)existence of interpolants and explicit definitions using bisimulations.

**Definition 5 (Joint consistency modulo  $\mathcal{L}(\Sigma)$ -bisimulations).** *Let  $\mathcal{L}$  be a DL. Let  $\mathcal{O}_1, \mathcal{O}_2$  be  $\mathcal{L}$ -ontologies,  $C_1, C_2$  be  $\mathcal{L}$ -concepts, and  $\Sigma \subseteq \text{sig}(\mathcal{O}_1, \mathcal{O}_2, C_1, C_2)$  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 following result characterizes the existence of interpolants using joint consistency modulo  $\mathcal{L}(\Sigma)$ -bisimulations.

**Theorem 2 (Interpolants).** *Let  $\mathcal{L} \in \{\mathcal{ALCCO}, \mathcal{ALCCIO}, \mathcal{ALCCO}^u, \mathcal{ALCCIO}^u\}$ . Let  $\mathcal{O}_1, \mathcal{O}_2$  be  $\mathcal{L}$ -ontologies and let  $C_1, C_2$  be  $\mathcal{L}$ -concepts, and  $\Sigma = \text{sig}(\mathcal{O}_1, C_1) \cap \text{sig}(\mathcal{O}_2, C_2)$ . Then the following conditions are equivalent:*

1. *there does not exist any  $\mathcal{L}$ -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.*

**Proof.** “1  $\Rightarrow$  2”. Assume there is no  $\mathcal{L}$ -interpolant for  $C_1 \sqsubseteq C_2$  under  $\mathcal{O}_1 \cup \mathcal{O}_2$ . Let  $\Sigma = \text{sig}(\mathcal{O}_1, C_1) \cap \text{sig}(\mathcal{O}_2, C_2)$  and define

$$\Gamma = \{D \mid \mathcal{O}_1 \cup \mathcal{O}_2 \models C_1 \sqsubseteq D, D \in \mathcal{L}(\Sigma)\}.$$

Then  $\mathcal{O}_1 \cup \mathcal{O}_2 \not\models D \sqsubseteq C_2$ , for any  $D \in \Gamma$ . As  $\Gamma$  is closed under conjunction and by compactness, there exists a model  $\mathcal{J}$  of  $\mathcal{O}_1 \cup \mathcal{O}_2$  and a node  $d$  in it such that  $d \in D^{\mathcal{J}}$  for all  $D \in \Gamma$  but  $d \notin C_2^{\mathcal{J}}$ . Consider the full  $\Sigma$ -type  $t_{\mathcal{J}}^{\mathcal{L}(\Sigma)}(d)$  of  $d$  in  $\mathcal{J}$ , defined as the set of all  $\mathcal{L}(\Sigma)$ -concepts  $D$  such that  $d \in D^{\mathcal{J}}$ . Then by compactness there exists a model  $\mathcal{I}$  of  $\mathcal{O}_1 \cup \mathcal{O}_2$  and a node  $e$  in it such that  $e \in C_1^{\mathcal{I}}$  and  $e \in D^{\mathcal{I}}$  for all  $D \in t_{\mathcal{J}}^{\mathcal{L}(\Sigma)}(d)$ . Thus  $\mathcal{I}, e \equiv_{\mathcal{L}, \Sigma} \mathcal{J}, d$ . We may assume both  $\mathcal{I}$  and  $\mathcal{J}$  are  $\omega$ -saturated. By Lemma 1,  $\mathcal{I}, e \sim_{\mathcal{L}, \Sigma} \mathcal{J}, d$ .

“2  $\Rightarrow$  1”. Assume an  $\mathcal{L}$ -interpolant  $D$  for  $C_1 \sqsubseteq C_2$  under  $\mathcal{O}_1 \cup \mathcal{O}_2$  exists. Assume that Condition 2 holds, that is, there are models  $\mathcal{I}_1$  and  $\mathcal{I}_2$  of  $\mathcal{O}_1 \cup \mathcal{O}_2$  and  $d_i \in \Delta^{\mathcal{I}_i}$  for  $i = 1, 2$  such that  $d_1 \in C_1^{\mathcal{I}_1}$  and  $d_2 \notin C_2^{\mathcal{I}_2}$  and  $\mathcal{I}_1, d_1 \sim_{\mathcal{L}, \Sigma} \mathcal{I}_2, d_2$ , where  $\Sigma = \text{sig}(\mathcal{O}_1, C_1) \cap \text{sig}(\mathcal{O}_2, C_2)$ . Then, by Lemma 1,  $\mathcal{I}_1, d_1 \equiv_{\mathcal{L}, \Sigma} \mathcal{I}_2, d_2$ . But then from  $d_1 \in C_1^{\mathcal{I}_1}$  we obtain  $d_1 \in D^{\mathcal{I}_1}$  and so  $d_2 \in D^{\mathcal{I}_2}$  which implies  $d_2 \in C_2^{\mathcal{I}_2}$ , a contradiction.  $\square$



We now characterize the existence of explicit definitions using joint consistency modulo  $\mathcal{L}(\Sigma)$ -bisimulations. Recall that by Lemma 2 there is a polynomial time reduction of explicit definition existence to interpolant existence. In fact, we will use the model-theoretic characterization of interpolant existence to prove the 2EXPTIME upper bound and we use the characterization of explicit definition existence to prove the matching complexity lower bound.

**Theorem 3 (Definitions).** *Let  $\mathcal{L} \in \{\mathcal{ALCCO}, \mathcal{ALCCIO}, \mathcal{ALCCO}^u, \mathcal{ALCCIO}^u\}$ . Let  $\mathcal{O}$  be an  $\mathcal{L}$ -ontology,  $C$  an  $\mathcal{L}$ -concept, and  $\Sigma \subseteq \text{sig}(\mathcal{O}, C)$  a signature. Then the following conditions are equivalent:*

1. *there does not exist any 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.*

**Proof.** Similar to the proof of Theorem 2. □

The following reduction is a consequence of Theorems 2 and 3, respectively.

**Theorem 4.** *Let  $\mathcal{L} \in \{\mathcal{ALCCO}, \mathcal{ALCCIO}, \mathcal{ALCCO}^u, \mathcal{ALCCIO}^u\}$ . Then  $\mathcal{L}$ -interpolant existence and explicit  $\mathcal{L}$ -definition existence are polynomial time reducible to the complement of joint consistency modulo  $\mathcal{L}$ -bisimulations.*

## 6 The 2EXPTIME upper bound

We provide a double exponential time algorithm that decides joint consistency modulo  $\mathcal{L}$ -bisimulations, for the DLs  $\mathcal{L}$  considered in this article.

**Theorem 5.** *Let  $\mathcal{L} \in \{\mathcal{ALCCO}, \mathcal{ALCCIO}, \mathcal{ALCCO}^u, \mathcal{ALCCIO}^u\}$ . Then joint consistency modulo  $\mathcal{L}$ -bisimulations is in 2EXPTIME.*

We pursue a mosaic-style decision procedure based on types. Assume  $\mathcal{L} \in \{\mathcal{ALCCO}, \mathcal{ALCCIO}, \mathcal{ALCCO}^u, \mathcal{ALCCIO}^u\}$ . 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 sub-concepts of concepts in  $\mathcal{O}_1, \mathcal{O}_2, C_1, C_2$ . A  $\Xi$ -type  $t$  is a maximal subset of  $\Xi$  such that there exists a model  $\mathcal{I}$  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  $\Xi$ -type realized at  $d$  in  $\mathcal{I}$ . Let  $T(\Xi)$  denote the set of all  $\Xi$ -types. Let  $r$  be a role. A pair  $(t_1, t_2)$  of  $\Xi$ -types  $t_1, t_2$  is  $r$ -coherent, in symbols  $t_1 \rightsquigarrow_r t_2$ , if there exists a model  $\mathcal{I}$  and  $(d_1, d_2) \in r^{\mathcal{I}}$  such that  $t_i = \text{tp}_{\Xi}(\mathcal{I}, d_i)$  for  $i = 1, 2$ .

We construct models using pairs  $(T_1, T_2) \subseteq T(\Xi) \times T(\Xi)$  such that all  $t \in T_i$  are satisfiable in  $\mathcal{L}(\Sigma)$ -bisimilar nodes of models of  $\mathcal{O}_i$ ,  $i = 1, 2$ . The following example illustrates the reason for working with sets of such pairs of sets of types.

*Example 3.* Consider the ontology  $\mathcal{O}$ , interpretation  $\mathcal{I}$ , and signature  $\Sigma$  defined in Example 1.  $\mathcal{I}$  witnesses that  $\mathcal{O}, \{a\}$  and  $\mathcal{O}, \neg\{a\}$  are jointly consistent modulo  $\mathcal{ALC}\mathcal{IO}^u(\Sigma)$ -bisimulations. The bisimulation was given by  $S = \Delta^{\mathcal{I}} \times \Delta^{\mathcal{I}}$  and required just two nodes to be bisimilar. This is not always the case. To illustrate, add to  $\mathcal{O}$  the CI  $\neg\{a\} \sqsubseteq \exists r. \neg\{a\}$  and denote by  $\mathcal{O}'$  the resulting ontology. In the interpretation showing that  $\mathcal{O}', \{a\}$  and  $\mathcal{O}', \neg\{a\}$  are jointly consistent modulo  $\mathcal{ALC}\mathcal{IO}^u(\Sigma)$ -bisimulations, one now requires at least three distinct nodes satisfying distinct types that are  $\mathcal{ALC}\mathcal{IO}^u(\Sigma)$ -bisimilar. In fact, it is straightforward to introduce a counter up to  $2^n - 1$  using  $n$  concept names not in  $\Sigma$  such that in the resulting ontology  $\mathcal{O}''$ , we still have that  $\mathcal{O}'', \{a\}$  and  $\mathcal{O}'', \neg\{a\}$  are jointly consistent modulo  $\mathcal{ALC}\mathcal{IO}^u(\Sigma)$ -bisimulations but in any interpretation witnessing this there exist  $2^n$  distinct nodes realizing distinct types that are  $\mathcal{ALC}\mathcal{IO}^u(\Sigma)$ -bisimilar.

We now formulate conditions on a set  $\mathcal{S} \subseteq 2^{T(\Xi) \times T(\Xi)}$  that ensures 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  $\mathcal{L}(\Sigma)$ -bisimilar. We lift the definition of  $r$ -coherence from pairs of types to pairs of elements of  $2^{T(\Xi) \times T(\Xi)}$ . 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  $t \rightsquigarrow_r t'$ . 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  $(T_1, T_2) \rightsquigarrow_r (T'_1, T'_2)$  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  $t \rightsquigarrow_r t'$ .

We first formulate conditions that ensure that nominals are interpreted as singletons and that individuals in  $\Sigma$  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 satisfiable in a model of  $\mathcal{O}_i$ ,  $i = 1, 2$ .

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

1.  $\Sigma$ -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 there exists  $t' \in T'_i$  with  $C \in t'$  and  $t \rightsquigarrow_r t'$ .
3.  $\Sigma$ -existential saturation: for  $i = 1, 2$ , if  $(T_1, T_2) \in \mathcal{S}$  and  $\exists r.C \in t \in T_i$ , where  $r$  is a role name in  $\Sigma$ , then there exist  $(T'_1, T'_2) \in \mathcal{S}$  such that  $(T_1, T_2) \rightsquigarrow_r (T'_1, T'_2)$  and there exists  $t' \in T'_i$  with  $C \in t'$  such that  $t \rightsquigarrow_r t'$ .

The conditions above are sufficient for  $\mathcal{ALCO}$ . If inverse roles or the universal role are present then we strengthen Condition 3 to Condition 3 $\mathcal{I}$  and add the Condition 4 $u$ , respectively:

- 3 $\mathcal{I}$ . Condition 3 with ‘ $(T_1, T_2) \rightsquigarrow_r (T'_1, T'_2)$ ’ replaced by ‘ $(T_1, T_2) \rightsquigarrow_r (T'_1, T'_2)$ ’.  
 4 $u$ . if  $(T_1, T_2) \in \mathcal{S}$ , then  $T_i \neq \emptyset$ , for  $i = 1, 2$ .

Thus,  $\mathcal{S}$  is  $\mathcal{ALCIO}(\Sigma)$ -good if the conditions above hold with Condition 3 replaced by Condition 3 $\mathcal{I}$  and  $\mathcal{S}$  is  $\mathcal{ALCO}^u(\Sigma)$ -good and, respectively,  $\mathcal{ALCIO}^u(\Sigma)$ -good if also Condition 4 $u$  holds.

**Lemma 3.** *Let  $\mathcal{L} \in \{\mathcal{ALCO}, \mathcal{ALCIO}, \mathcal{ALCO}^u, \mathcal{ALCIO}^u\}$ . Assume  $\mathcal{O}_1, \mathcal{O}_2$  are  $\mathcal{L}$ -ontologies,  $C_1, C_2$  are  $\mathcal{L}$ -concepts, and let  $\Sigma \subseteq \text{sig}(\Xi)$  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  $\Xi$ -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.** “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  $\Xi$ -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  $T_1, T_2$  are maximal such that there are

- $d_t \in \Delta^{\mathcal{I}_1}$  realizing  $t$  in  $\mathcal{I}_1$ , for all  $t \in T_1$ , and
- $e_t \in \Delta^{\mathcal{I}_2}$  realizing  $t$  in  $\mathcal{I}_2$ , for all  $t \in T_2$

and all  $\{d_t \mid t \in T_1\} \cup \{e_t \mid t \in T_2\}$  are  $\mathcal{L}(\Sigma)$ -bisimilar. It is straightforward to show that  $\mathcal{S}$  is  $\mathcal{L}(\Sigma)$ -good and satisfies Point 2.

“2  $\Rightarrow$  1”. Assume  $\mathcal{S}$  is  $\mathcal{L}(\Sigma)$ -good and we have  $\Xi$ -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 we construct interpretations  $\mathcal{I}_1$  and  $\mathcal{I}_2$  by setting

$$\begin{aligned} \Delta^{\mathcal{I}_i} &:= \{(t, (T_1, T_2)) \mid t \in T_i \text{ and } (T_1, T_2) \in \mathcal{S}\} \\ r^{\mathcal{I}_i} &:= \{((t, p), (t', p')) \in \Delta^{\mathcal{I}_i} \times \Delta^{\mathcal{I}_i} \mid p \rightsquigarrow_r p' \text{ and } t \rightsquigarrow_r t'\} & r \in \Sigma \\ r^{\mathcal{I}_i} &:= \{((t, p), (t', p')) \in \Delta^{\mathcal{I}_i} \times \Delta^{\mathcal{I}_i} \mid t \rightsquigarrow_r t'\} & r \notin \Sigma \\ A^{\mathcal{I}_i} &:= \{(t, p) \in \Delta^{\mathcal{I}_i} \mid A \in t\} \\ a^{\mathcal{I}_i} &:= (t, (T_1, T_2)) \in \Delta^{\mathcal{I}_i} & a \in t \in T_i \end{aligned}$$

If  $\mathcal{L}$  admits inverse roles then the definition is the same except that we replace in the definition of  $r^{\mathcal{I}_i}$  for  $r \in \Sigma$  the condition ‘ $p \rightsquigarrow_r p'$ ’ by ‘ $p \rightsquigarrow_r p'$ ’.

One can show by induction that for  $i = 1, 2$  and all  $D \in \Xi$  and  $(t, p) \in \Delta^{\mathcal{I}_i}$ :  $D \in t$  iff  $(t, p) \in D^{\mathcal{I}_i}$ . Thus  $\mathcal{I}_i$  is a model  $\mathcal{O}_i$  for  $i = 1, 2$ . Moreover, let

$$S = \{((t_1, p_1), (t_2, p_2)) \in \Delta^{\mathcal{I}_1} \times \Delta^{\mathcal{I}_2} \mid p_1 = p_2\}.$$

Then  $S$  is a  $\mathcal{L}(\Sigma)$ -bisimulation between  $\mathcal{I}_1$  and  $\mathcal{I}_2$  witnessing that  $\mathcal{O}_1, C_1$  and  $\mathcal{O}_2, C_2$  are jointly consistent modulo  $\mathcal{L}(\Sigma)$ -bisimulations.  $\square$

**Lemma 4.** *Let  $\mathcal{L} \in \{\mathcal{ALCCO}, \mathcal{ALCCIO}, \mathcal{ALCCO}^u, \mathcal{ALCCIO}^u\}$ . 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  $\Sigma \subseteq \text{sig}(\Xi)$  Condition 2 of Lemma 3 holds.*

**Proof.** We start with  $\mathcal{ALCCO}$ . Assume  $\mathcal{O}_1, \mathcal{O}_2, C_1, C_2$ , and  $\Sigma$  are given. We can enumerate in double exponential time the maximal sets  $\mathcal{U} \subseteq 2^{T(\Xi) \times T(\Xi)}$  that are good for nominals and for  $\mathcal{O}_1, \mathcal{O}_2$ : simply remove from  $2^{T(\Xi) \times T(\Xi)}$  all  $(T_1, T_2)$  such that  $T_i$  contains a  $t$  that is not satisfiable in any model of  $\mathcal{O}_i$  and then take the sets  $\mathcal{U}$  in which for each nominal  $a \in \text{sig}(\Xi)$  the types  $t_a^i$  and the pairs  $(T_1, T_2)$  in which  $t_a^i$  occurs have been selected, for  $i = 1, 2$ . Also make sure that either  $(\{t_a^1\}, \{t_a^2\}) \in \mathcal{U}$  or  $(\{t_a^1\}, \emptyset), (\emptyset, \{t_a^2\}) \in \mathcal{U}$ . Then we eliminate from any such  $\mathcal{U}$  recursively all pairs that are not  $\Sigma$ -concept name coherent, existentially saturated, or  $\Sigma$ -existentially saturated. Let  $\mathcal{S}_0 \subseteq \mathcal{U}$  be the largest fixpoint of this procedure. Then one can easily show that there exists a set  $\mathcal{S}$  satisfying Condition 2 of Lemma 3 iff there exists a maximal set  $\mathcal{U}$  that is good for nominals and  $\mathcal{O}_1, \mathcal{O}_2$  such that the largest fixpoint  $\mathcal{S}_0$  satisfies Condition 2 of Lemma 3. The elimination procedure is in double exponential time.

The modifications needed for the remaining DLs are straightforward: for DLs with inverse roles modify the recursive elimination procedure by considering Condition 3 $\mathcal{I}$  for  $\Sigma$ -existential saturation and for DLs with the universal role remove any  $(T_1, T_2)$  with  $T_i = \emptyset$  for  $i = 1$  or  $i = 2$  from any  $\mathcal{U}$ .  $\square$

Theorem 5 is a direct consequence of Lemmas 3 and 4.

## 7 The 2EXPTIME lower bound

We show that for any of the DLs  $\mathcal{L}$  considered in this paper narrow  $\mathcal{L}$ -referring expression existence is 2EXPTIME-hard. It follows that  $\mathcal{L}$ -explicit definition existence and  $\mathcal{L}$ -interpolant existence are also 2EXPTIME-hard. By Theorems 2 and 3, it suffices to prove the following result.

**Lemma 5.** *Let  $\mathcal{L} \in \{\mathcal{ALCCO}, \mathcal{ALCCIO}, \mathcal{ALCCO}^u, \mathcal{ALCCIO}^u\}$ . It is 2EXPTIME-hard to decide for an  $\mathcal{L}$ -ontology  $\mathcal{O}$ , individual name  $b$ , and signature  $\Sigma \subseteq \text{sig}(\mathcal{O}) \setminus \{b\}$  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  $\Sigma = \text{sig}(\mathcal{O}) \setminus \{b\}$ .*

We reduce the word problem for exponentially space bounded alternating Turing machines (ATMs). We actually use a slightly unusual ATM model which is easily seen to be equivalent to the standard model.

An *alternating Turing machine (ATM)* is a tuple  $M = (Q, \Theta, \Gamma, q_0, \Delta)$  where  $Q = Q_\exists \uplus Q_\forall$  is the set of states that consists of *existential states* in  $Q_\exists$  and *universal states* in  $Q_\forall$ . Further,  $\Theta$  is the input alphabet and  $\Gamma$  is the tape alphabet that contains a *blank symbol*  $\square \notin \Theta$ ,  $q_0 \in Q_\exists$  is the *starting state*, and the *transition relation*  $\Delta$  is of the form  $\Delta \subseteq Q \times \Gamma \times Q \times \Gamma \times \{L, R\}$ . The set  $\Delta(q, a) := \{(q', a', M) \mid (q, a, q', a', M) \in \Delta\}$  must contain exactly two or zero elements for every  $q \in Q$  and  $a \in \Gamma$ . Moreover, the state  $q'$  must be from  $Q_\forall$  if

$q \in Q_{\exists}$  and from  $Q_{\exists}$  otherwise, that is, existential and universal states alternate. Note that there is no accepting state. The ATM accepts if it runs forever and rejects otherwise. Starting from the standard ATM model, this can be achieved by assuming that exponentially space bounded ATMs terminate on any input and then modifying them to enter an infinite loop from the accepting state.

A *configuration* of an ATM is a word  $wqw'$  with  $w, w' \in \Gamma^*$  and  $q \in Q$ . We say that  $wqw'$  is *existential* if  $q$  is, and likewise for *universal*. *Successor configurations* are defined in the usual way. Note that every configuration has exactly two successor configurations.

A *computation tree* of an ATM  $M$  on input  $w$  is an infinite tree whose nodes are labeled with configurations of  $M$  such that

- the root is labeled with the initial configuration  $q_0w$ ;
- if a node is labeled with an existential configuration  $wqw'$ , then it has a single successor which is labeled with a successor configuration of  $wqw'$ ;
- if a node is labeled with a universal configuration  $wqw'$ , then it has two successors which are labeled with the two successor configurations of  $wqw'$ .

An ATM  $M$  *accepts* an input  $w$  if there is a computation tree of  $M$  on  $w$ .

We reduce the word problem for  $2^n$ -space bounded ATMs which is well-known to be 2EXPTIME-hard [12]. We first provide the reduction for  $\mathcal{ALCO}$  using an ontology  $\mathcal{O}$  and signature  $\Sigma$  such that  $\mathcal{O}$  contains concept names that are not in  $\Sigma$ . We set

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

The idea of the reduction is as follows. 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  $\rho$ . Along  $\rho$ , a counter counts modulo  $2^n$  using concept names not in  $\Sigma$ . Additionally, in each point of  $\rho$  starts an infinite tree along role  $s$  that is supposed to mimick the computation tree of  $M$ . Along this tree, two counters are maintained:

- one counter starting at 0 and counting modulo  $2^n$  to divide the tree in subpaths of length  $2^n$ ; each such path of length  $2^n$  represents a configuration;
- another counter starting at *the value of the counter on  $\rho$*  and also counting modulo  $2^n$ .

To link successive configurations we use that if there exist models  $\mathcal{I}$  and  $\mathcal{J}$  of  $\mathcal{O}$  such that  $\mathcal{I}, b^{\mathcal{I}} \sim_{\mathcal{ALCO}, \Sigma} \mathcal{J}, d$  for some  $d \neq b^{\mathcal{J}}$  it follows that in  $\mathcal{J}$  all nodes on some  $r$ -path  $\rho$  through  $d$  are  $\mathcal{ALCO}(\Sigma)$ -bisimilar. Thus, each node on  $\rho$  is the starting point of  $s$ -trees with identical  $\Sigma$ -decorations. As on the  $m$ th  $s$ -tree the second counter starts at all nodes at distances  $k \times 2^n - m$ , for all  $k \geq 1$ , we are in the position to coordinate all positions at all successive configurations.

The ontology  $\mathcal{O}$  is constructed as follows. We enforce that any node  $d$  that does not equal  $b$  and has an  $r$ -successor satisfies a concept name  $I_s$  that triggers an  $A_i$ -counter along the role name  $r$  and starts  $s$ -trees. We thus have for concept names  $A_i, \bar{A}_i$ ,  $i < n$ :

$$\begin{aligned}
\neg\{b\} \cap \exists r. \top &\subseteq I_s \\
I_s &\subseteq \prod_{i < n} (A_i \sqcup \bar{A}_i) \\
I_s &\subseteq \exists r. \top \cap \forall r. I_s \\
A_i \cap \prod_{j < i} A_j &\subseteq \forall r. \bar{A}_i \\
\bar{A}_i \cap \prod_{j < i} A_j &\subseteq \forall r. A_i \\
A_i \cap \bigsqcup_{j < i} \bar{A}_j &\subseteq \forall r. A_i \\
\bar{A}_i \cap \bigsqcup_{j < i} \bar{A}_j &\subseteq \forall r. \bar{A}_i
\end{aligned}$$

Using the concept names  $I_s$ , we start the  $s$ -trees with two counters, realized using concept names  $U_i, \bar{U}_i$  and  $V_i, \bar{V}_i$ ,  $i < n$ , and initialized to 0 and the value of the  $A$ -counter, respectively:

$$\begin{aligned}
I_s &\subseteq (U = 0) \\
I_s \cap A_j &\subseteq V_j & j < n \\
I_s \cap \bar{A}_j &\subseteq \bar{V}_j & j < n \\
\top &\subseteq \exists s. \top
\end{aligned}$$

Here,  $(U = 0)$  is an abbreviation for the concept  $\prod_{i=1}^n \bar{U}_i$ , we use similar abbreviations below. The counters  $U_i$  and  $V_i$  are incremented along  $s$  analogously to how  $A_i$  is incremented along  $r$ , so we omit details. Configurations of  $M$  are represented between two consecutive points having  $U$ -counter value 0. We next enforce the structure of the computation tree, assuming that  $q_0 \in Q_\forall$ :

$$\begin{aligned}
I_s &\subseteq B_\forall \\
(U < 2^n - 1) \cap B_\forall &\subseteq \forall s. B_\forall \\
(U < 2^n - 1) \cap B_\exists^i &\subseteq \forall s. B_\exists^i & i \in \{1, 2\} \\
(U = 2^n - 1) \cap B_\forall &\subseteq \forall s. (B_\exists^1 \sqcup B_\exists^2) \\
(U = 2^n - 1) \cap (B_\exists^1 \sqcup B_\exists^2) &\subseteq \forall s. B_\forall \\
(U = 2^n - 1) \cap B_\forall &\subseteq \exists s. Z \cap \exists s. \neg Z
\end{aligned}$$

These sentences enforce that all points which represent a configuration satisfy exactly one of  $B_\forall, B_\exists^1, B_\exists^2$  indicating the kind of configuration and, if existential, also a choice of the transition function. The symbol  $Z \in \Sigma$  enforces the branching.

We next set the initial configuration, for input  $w = a_0, \dots, a_{n-1}$ .

$$\begin{aligned}
I_s &\sqsubseteq A_{q_0, a_0} \\
I_s &\sqsubseteq \forall s^k. A_{a_k} & 0 < k < n \\
I_s &\sqsubseteq \forall s^{n+1}. \text{Blank} \\
\text{Blank} &\sqsubseteq A_{\square} \\
\text{Blank} \sqcap (U < 2^n - 1) &\sqsubseteq \forall s. \text{Blank}
\end{aligned}$$

To coordinate consecutive configurations, we associate with  $M$  functions  $f_i$ ,  $i \in \{1, 2\}$  that map the content of three consecutive cells of a configuration to the content of the middle cell in the  $i$ -th successor configuration (assuming an arbitrary order on the set  $\Delta(q, a)$ ). In what follows, we ignore the cornercases that occur at the border of configurations; they are treated in a similar way. Clearly, for each possible such triple  $(\sigma_1, \sigma_2, \sigma_3) \in \Gamma \cup (Q \times \Gamma)$ , there is an  $\mathcal{ALC}$  concept  $C_{\sigma_1, \sigma_2, \sigma_3}$  which is true at an element  $a$  of the computation tree iff  $a$  is labeled with  $A_{\sigma_1}$ ,  $a$ 's  $s$ -successor  $b$  is labeled with  $A_{\sigma_2}$ , and  $b$ 's  $s$ -successor  $c$  is labeled with  $A_{\sigma_3}$ . Now, in each configuration, we synchronize elements with  $V$ -counter 0 by including for every  $\sigma = (\sigma_1, \sigma_2, \sigma_3)$  and  $i \in \{1, 2\}$  the following sentences:

$$\begin{aligned}
(V = 2^n - 1) \sqcap (U < 2^n - 2) \sqcap C_{\sigma_1, \sigma_2, \sigma_3} &\sqsubseteq \forall s. A_{f_1(\sigma)}^1 \sqcap \forall s. A_{f_2(\sigma)}^2 \\
(V = 2^n - 1) \sqcap (U < 2^n - 2) \sqcap C_{\sigma_1, \sigma_2, \sigma_3} \sqcap B_{\exists}^i &\sqsubseteq \forall s. A_{f_i(\sigma)}^i
\end{aligned}$$

The concept names  $A_{\sigma}^i$  are used as markers (not in  $\Sigma$ ) and are propagated along  $s$  for  $2^n$  steps, exploiting the  $V$ -counter. The superscript  $i \in \{1, 2\}$  determines the successor configuration that the symbol is referring to. After crossing the end of a configuration, the symbol  $\sigma$  is propagated using concept names  $A'_{\sigma}$  (the superscript is not needed anymore because the branching happens at the end of the configuration, based on  $Z$ ).

$$\begin{aligned}
(U < 2^n - 1) \sqcap A_{\sigma}^i &\sqsubseteq \forall s. A_{\sigma}^i \\
(U = 2^n - 1) \sqcap B_{\forall} \sqcap A_{\sigma}^1 &\sqsubseteq \forall s. (\neg Z \sqcup A'_{\sigma}) \\
(U = 2^n - 1) \sqcap B_{\forall} \sqcap A_{\sigma}^2 &\sqsubseteq \forall s. (Z \sqcup A'_{\sigma}) \\
(U = 2^n - 1) \sqcap B_{\exists}^i \sqcap A_{\sigma}^i &\sqsubseteq \forall s. A'_{\sigma} & i \in \{1, 2\} \\
(V < 2^n - 1) \sqcap A'_{\sigma} &\sqsubseteq \forall s. A'_{\sigma} \\
(V = 2^n - 1) \sqcap A'_{\sigma} &\sqsubseteq \forall s. A_{\sigma}
\end{aligned}$$

For those  $(q, a)$  with  $\Delta(q, a) = \emptyset$ , we add the concept inclusion

$$A_{q, a} \sqsubseteq \perp.$$

The following lemma establishes correctness of the reduction.

**Lemma 6.** *The following conditions are equivalent:*

1.  $M$  accepts  $w$ .
2. there exist models  $\mathcal{I}$  and  $\mathcal{J}$  of  $\mathcal{O}$  such that  $\mathcal{I}, b^{\mathcal{I}} \sim_{\mathcal{ALCO}, \Sigma} \mathcal{J}, d$ , for some  $d \neq b^{\mathcal{J}}$ .

**Proof.** “ $1 \Rightarrow 2$ ”. If  $M$  accepts  $w$ , there is a computation tree of  $M$  on  $w$ . We construct a single interpretation  $\mathcal{I}$  with  $\mathcal{I}, b^{\mathcal{I}} \sim_{\mathcal{ALCO}, \Sigma} \mathcal{I}, d$  for some  $d \neq b^{\mathcal{I}}$  as follows. Let  $\hat{\mathcal{J}}$  be the infinite tree-shaped interpretation that represents the computation tree of  $M$  on  $w$  as described above, that is, configurations are represented by sequences of  $2^n$  elements linked by role  $s$  and labeled by  $B_{\forall}, B_{\exists}^1, B_{\exists}^2$  depending on whether the configuration is universal or existential, and in the latter case the superscript indicates which choice has been made for the existential state. Finally, the first element of the first successor configuration of a universal configuration is labeled with  $Z$ . Observe that  $\hat{\mathcal{J}}$  interprets only the symbols in  $\Sigma$  as non-empty. Now, we obtain interpretation  $\mathcal{I}_k$ ,  $k < 2^n$  from  $\hat{\mathcal{J}}$  by interpreting non- $\Sigma$ -symbols as follows:

- the root of  $\mathcal{I}_k$  satisfies  $I_s$ ;
- the  $U$ -counter starts at 0 at the root and counts modulo  $2^n$  along each  $s$ -path;
- the  $V$ -counter starts at  $k$  at the root and counts modulo  $2^n$  along each  $s$ -path;
- the auxiliary concept names of the shape  $A_{\sigma}^i$  and  $A'_{\sigma}$  are interpreted in a minimal way so as to satisfy the concept inclusions listed above. Note that the respective concept inclusions are Horn, hence there is no choice.

Now obtain  $\mathcal{I}$  from  $\hat{\mathcal{J}}$  and the  $\mathcal{I}_k$  by creating an infinite outgoing  $r$ -path  $\rho$  from some node  $d$  (with the corresponding  $A$ -counter) and adding all  $\mathcal{I}_k$  to every node on the  $r$ -path, identifying the roots of the  $\mathcal{I}_k$  with the node on the path. Additionally, add  $\hat{\mathcal{J}}$  to  $b^{\mathcal{I}} = b$  by identifying  $b$  with the root of  $\hat{\mathcal{J}}$ . It should be clear that  $\mathcal{I}$  is as required. In particular, the reflexive and symmetric closure of

- all pairs  $(b, e), (e, e')$ , with  $e, e'$  on  $\rho$ , and
- all pairs  $(e, e'), (e', e'')$ , with  $e$  in  $\hat{\mathcal{J}}$  and  $e', e''$  copies of  $e$  in the trees  $\mathcal{I}_k$ .

is an  $\mathcal{ALCO}(\Sigma)$ -bisimulation  $S$  on  $\mathcal{I}$  with  $(b, d) \in S$ .

“ $2 \Rightarrow 1$ ”. Assume that  $\mathcal{I}, b^{\mathcal{I}} \sim_{\mathcal{ALCO}, \Sigma} \mathcal{J}, d$  for some  $d \neq b^{\mathcal{J}}$ . As argued above, due to the  $r$ -self loop at  $b^{\mathcal{I}}$ , from  $d$  there has to be an outgoing infinite  $r$ -path on which all  $s$ -trees are  $\mathcal{ALCO}(\Sigma)$ -bisimilar. Since  $\mathcal{I}$  is a model of  $\mathcal{O}$ , all these  $s$ -trees are additionally labeled with some auxiliary concept names not in  $\Sigma$ , depending on the distance from their roots on  $\rho$ . Using the concept inclusions in  $\mathcal{O}$  and the arguments given in their description, it can be shown that all  $s$ -trees contain a computation tree of  $M$  on input  $w$  (which is solely represented with concept names in  $\Sigma$ ).  $\square$

The same ontology  $\mathcal{O}$  can be used for the remaining three DLs. For  $\mathcal{ALCO}^u$ , exactly the same proof works. For the DLs with inverse roles the infinite  $r$ -path  $\rho$  has to be extended to an infinite  $r^-$ -path.

Using the ontology  $\mathcal{O}$  defined above we define a new ontology  $\mathcal{O}'$  to obtain the 2EXPTIME lower bound for signatures  $\Sigma' = \text{sig}(\mathcal{O}') \setminus \{b\}$ . Fix a role name



$r_E$  for any concept name  $E \in \text{sig}(\mathcal{O}) \setminus \Sigma$ . Now replace in  $\mathcal{O}$  any occurrence of  $E \in \text{sig}(\mathcal{O}) \setminus \Sigma$  by  $\exists r_E.\{b\}$  and denote the resulting ontology by  $\mathcal{O}'$ .

**Lemma 7.** *The following conditions are equivalent:*

1.  $M$  accepts  $w$ .
2. there exist models  $\mathcal{I}$  and  $\mathcal{J}$  of  $\mathcal{O}'$  such that  $\mathcal{I}, b^{\mathcal{I}} \sim_{\mathcal{ALCO}, \Sigma'} \mathcal{J}, d$ , for some  $d \neq b^{\mathcal{J}}$ .

**Proof.** “1  $\Rightarrow$  2”. We modify the interpretation  $\mathcal{I}$  defined in the proof of Lemma 6 in such a way that we obtain a model of  $\mathcal{O}'$  and such that the  $\mathcal{ALCO}(\Sigma)$ -bisimulation  $S$  on  $\mathcal{I}$  defined in that proof is, in fact, an  $\mathcal{ALCO}(\Sigma')$ -bisimulation on the new interpretation. Formally, obtain  $\mathcal{I}'$  from  $\mathcal{I}$  by interpreting every  $r_E$ ,  $E \in \text{sig}(\mathcal{O}) \setminus \Sigma$  as follows:

- (i) there is an  $r_E$ -edge from  $e$  to  $b^{\mathcal{I}}$  for all  $e \in E^{\mathcal{I}}$ ;
- (ii) there is an  $r_E$ -edge from  $e$  to all nodes on the path  $\rho$  for all  $(e, e') \in S$  and  $e' \in E^{\mathcal{I}}$ ;
- (iii) there are no more  $r_E$ -edges.

Note that, by (i),  $\mathcal{I}'$  is a model of  $\mathcal{O}'$ . By (ii), the relation  $S$  defined in the proof of Lemma 6 is an  $\mathcal{ALCO}(\Sigma')$ -bisimulation. In particular, by (i), elements  $e' \in E^{\mathcal{I}}$  have now an  $r_E$ -edge to  $b^{\mathcal{I}}$ , so any element  $e$  bisimilar to  $e'$ , that is,  $(e, e') \in S$ , needs an  $r_E$ -successor to some element bisimilar to  $b^{\mathcal{I}}$ . Since all elements on the path  $\rho$  are bisimilar to  $b^{\mathcal{I}}$ , these  $r_E$ -successors exist due to (ii).

“2  $\Rightarrow$  1”. This direction remains the same as in the proof of Lemma 6.  $\square$

The extension to our DLs with inverse roles and the universal role is again straightforward.

## 8 Conclusions

DLs with nominals do not have the CIP nor the BDP. Thus, deciding the existence of interpolants and explicit definitions cannot be reduced to validity checking. In fact, we have shown that both problems are 2EXPTIME complete for DLs ranging from  $\mathcal{ALCO}$  to  $\mathcal{ALCIO}^u$ . Numerous interesting questions remain to be explored. Firstly, what is the size of interpolants and explicit definitions and how to compute them? The techniques introduced in this paper should be a good starting point. Secondly, while for  $\mathcal{ALCIO}$  and logics with the universal role, the 2EXPTIME lower bound holds already under empty ontologies, this remains open for  $\mathcal{ALCO}$ . Finally, there are many important DLs and other fragments of FO which fail to have the CIP and BDP and for which interpolant existence and explicit definition existence are of interest. For instance, we have recently proved that interpolant existence and explicit definition existence are 3EXPTIME-complete for the guarded fragment in general, and 2EXPTIME-complete if arities are bounded by a constant [25]. Is an increase by one exponential the rule when one moves from validity to interpolant/explicit definition

existence for logics without the CIP/BDP? We conjecture that the same holds for DLs with role hierarchies (which also fail to have the CIP and BDP) [11,26]. What about the two-variable fragment of FO? What about Horn DLs where recent semantic characterizations [24] might be helpful?

## References

1. Areces, C., Blackburn, P., Marx, M.: Hybrid logics: Characterization, interpolation and complexity. *J. Symb. Log.* 66(3), 977–1010 (2001)
2. Baader, F., Horrocks, I., Lutz, C., Sattler, U.: *An Introduction to Description Logic*. Cambridge University Press (2017)
3. Bárány, V., Benedikt, M., ten Cate, B.: Some model theory of guarded negation. *J. Symb. Log.* 83(4), 1307–1344 (2018)
4. Benedikt, M., ten Cate, B., Vanden Boom, M.: Interpolation with decidable fix-point logics. In: *Proc. of LICS*. pp. 378–389. IEEE Computer Society (2015)
5. Benedikt, M., ten Cate, B., Vanden Boom, M.: Effective interpolation and preservation in guarded logics. *ACM Trans. Comput. Log.* 17(2), 8:1–8:46 (2016)
6. Benedikt, M., Leblay, J., ten Cate, B., Tsamoura, E.: *Generating Plans from Proofs: The Interpolation-based Approach to Query Reformulation*. Synthesis Lectures on Data Management, Morgan & Claypool Publishers (2016)
7. Borgida, A., Toman, D., Weddell, G.E.: On referring expressions in query answering over first order knowledge bases. In: *Proc. of KR*. pp. 319–328 (2016)
8. Borgida, A., Toman, D., Weddell, G.E.: Concerning referring expressions in query answers. In: *Proc. of IJCAI*. pp. 4791–4795 (2017)
9. ten Cate, B.: Interpolation for extended modal languages. *J. Symb. Log.* 70(1), 223–234 (2005)
10. ten Cate, B., Conradie, W., Marx, M., Venema, Y.: Definitorially complete description logics. In: *Proc. of KR*. pp. 79–89 (2006)
11. ten Cate, B., Franconi, E., Seylan, I.: Beth definability in expressive description logics. *J. Artif. Intell. Res.* 48, 347–414 (2013)
12. Chandra, A.K., Kozen, D.C., Stockmeyer, L.J.: Alternation. *J. ACM* 28, 114–133 (1981)
13. Chang, C., Keisler, H.J.: *Model Theory*. Elsevier (1998)
14. Cimatti, A., Griggio, A., Sebastiani, R.: Interpolant generation for UTVPI. In: *Proc. of CADE*. pp. 167–182. Springer (2009)
15. Funk, M., Jung, J.C., Lutz, C., Pulcini, H., Wolter, F.: Learning description logic concepts: When can positive and negative examples be separated? In: *Proc. of IJCAI*. pp. 1682–1688 (2019)
16. Goel, A., Krstic, S., Tinelli, C.: Ground interpolation for combined theories. In: *Proc. of CADE*. pp. 183–198. Springer (2009)
17. Goranko, V., Otto, M.: Model theory of modal logic. In: *Handbook of Modal Logic*, pp. 249–329. Elsevier (2007)
18. Hoogland, E., Marx, M.: Interpolation and definability in guarded fragments. *Studia Logica* 70(3), 373–409 (2002)
19. Hoogland, E., Marx, M., Otto, M.: Beth definability for the guarded fragment. In: *Proc. of LPAR*. pp. 273–285. Lecture Notes in Computer Science, Springer (1999)
20. Iemhoff, R.: Uniform interpolation and the existence of sequent calculi. *Annals of Pure and Applied Logic* 170(11), 102711 (2019)

21. Jiménez-Ruiz, E., Payne, T.R., Solimando, A., Tamma, V.A.M.: Limiting logical violations in ontology alignment through negotiation. In: Proc. of KR. pp. 217–226. AAAI Press (2016)
22. Jung, J.C., Lutz, C., Pulcini, H., Wolter, F.: Logical separability of incomplete data under ontologies. In: Proc. of KR. IJCAI (2020)
23. Jung, J.C., Lutz, C., Pulcini, H., Wolter, F.: Separating positive and negative data examples by concepts and formulas: The case of restricted signature. In: submitted (2020)
24. Jung, J.C., Papacchini, F., Wolter, F., Zakharyashev, M.: Model comparison games for horn description logics. In: Proc. of LICS. pp. 1–14. IEEE (2019)
25. Jung, J.C., Wolter, F.: Living without beth and craig: Explicit definitions and interpolants in the guarded fragment (2020), available at <http://arxiv.org/abs/2007.01597>
26. Konev, B., Lutz, C., Walther, D., Wolter, F.: Formal properties of modularisation. In: Modular Ontologies, Lecture Notes in Computer Science, vol. 5445, pp. 25–66. Springer (2009)
27. Konev, B., Lutz, C., Ponomaryov, D.K., Wolter, F.: Decomposing description logic ontologies. In: Proc. of KR. AAAI Press (2010)
28. Koopmann, P., Schmidt, R.A.: Uniform interpolation and forgetting for ALC ontologies with aboxes. In: Proc. of AAAI. pp. 175–181. AAAI Press (2015)
29. Kowalski, T., Metcalfe, G.: Uniform interpolation and coherence. *Annals of pure and applied logic* 170(7), 825–841 (2019)
30. Krahmer, E., van Deemter, K.: Computational generation of referring expressions: A survey. *Computational Linguistics* 38(1), 173–218 (2012)
31. Lutz, C., Piro, R., Wolter, F.: Description logic TBoxes: Model-theoretic characterizations and rewritability. In: Proc. of IJCAI (2011)
32. Lutz, C., Seylan, I., Wolter, F.: An automata-theoretic approach to uniform interpolation and approximation in the description logic EL. In: Proc. of KR. AAAI Press (2012)
33. Lutz, C., Seylan, I., Wolter, F.: The data complexity of ontology-mediated queries with closed predicates. *Logical Methods in Computer Science* 15(3) (2019)
34. Lutz, C., Wolter, F.: Foundations for uniform interpolation and forgetting in expressive description logics. In: Proc. of IJCAI. pp. 989–995. IJCAI/AAAI (2011)
35. Maksimova, L., Gabbay, D.: *Interpolation and Definability in Modal and Intuitionistic Logics*. Clarendon Press (2005)
36. Marx, M.: Queries determined by views: Pack your views. In: Proc. of PODS. p. 23–30. ACM (2007)
37. McMillan, K.L.: Interpolation and sat-based model checking. In: Proc. of CAV. pp. 1–13. Springer (2003)
38. Nikitina, N., Rudolph, S.: (Non-)succinctness of uniform interpolants of general terminologies in the description logic EL. *Artif. Intell.* 215, 120–140 (2014)
39. Schaerf, A.: On the complexity of the instance checking problem in concept languages with existential quantification. *J. Intell. Inf. Syst.* 2(3), 265–278 (1993)
40. Seylan, I., Franconi, E., de Bruijn, J.: Effective query rewriting with ontologies over dboxes. In: Proc. of IJCAI. pp. 923–925 (2009)
41. Toman, D., Weddell, G.E.: *Fundamentals of Physical Design and Query Compilation*. Synthesis Lectures on Data Management, Morgan & Claypool Publishers (2011)
42. Toman, D., Weddell, G.E.: Identity resolution in ontology based data access to structured data sources. In: Proc. of PRICAI. pp. 473–485 (2019)

- 43. Van Benthem, J.: The many faces of interpolation. *Synthese* pp. 451–460 (2008)
- 44. Visser, A., et al.: Uniform interpolation and layered bisimulation. In: Gödel'96: Logical foundations of mathematics, computer science and physics—Kurt Gödel's legacy, pp. 139–164. Association for Symbolic Logic (1996)