Horn Rewritability vs PTime Query Evaluation for Description Logic TBoxes

André Hernich[1], Carsten Lutz[2], Fabio Papacchini[1], and Frank Wolter[1]

[1] University of Liverpool, UK, {firstname.surname}@liverpool.ac.uk
[2] University of Bremen, Germany, clu@uni-bremen.de

Abstract. We study the following question: if $\mathcal{T}$ is a TBox that is formulated in an expressive DL $\mathcal{L}$ and all CQs can be evaluated in PTime w.r.t. $\mathcal{T}$, can $\mathcal{T}$ be replaced by a TBox $\mathcal{T}'$ that is formulated in the Horn-fragment of $\mathcal{L}$ and such that for all CQs and ABoxes, the answers w.r.t. $\mathcal{T}$ and $\mathcal{T}'$ coincide? Our main results are that this is indeed the case when $\mathcal{L}$ is the set of $\mathcal{ALCHI}$ or $\mathcal{ALCIF}$ TBoxes of quantifier depth 1 (which covers the majority of such TBoxes), but not for $\mathcal{ALCHIF}$ and $\mathcal{ALCQ}$ TBoxes of depth 1.

1 Introduction

In ontology-mediated querying, TBoxes are used to enrich incomplete data with domain knowledge, enabling more complete answers to queries [30, 5, 22]. Since query evaluation is coNP-hard in the presence of TBoxes formulated in expressive description logics (DLs) such as $\mathcal{ALC}$ and $\mathcal{SHIQ}$ [31, 23, 14], the identification of computationally more well-behaved setups has been an important goal of research [1, 9, 8, 28]. In particular, this has led to the introduction of Horn-DLs, syntactically defined fragments of expressive DLs that fall within the Horn-fragment of first-order logic [16, 29]. Widely known examples include the fragment Horn-$\mathcal{SHIQ}$ of $\mathcal{SHIQ}$ and the corresponding fragment Horn-$\mathcal{ALC}$ of $\mathcal{ALC}$ [16, 4]. In contrast to the expressive DLs in which they are included, these Horn-DLs admit the construction of universal models giving exactly the same answers to queries as the class of all models of a knowledge base. The existence of universal models can then be used to show that query evaluation is in PTime in data complexity, and to design practical query answering algorithms [12, 29, 21]. Thus, any TBox formulated in an expressive DL that falls within the corresponding Horn-DL is computationally well-behaved.

In this paper, we ask the converse question, concentrating on conjunctive queries (CQs): *is it the case that every TBox $\mathcal{T}$ formulated in an expressive DL and for which CQ-evaluation is in* PTime *is rewritable into a TBox $\mathcal{T}'$ formulated in the corresponding Horn-DL?* If one requires $\mathcal{T}$ to be logically equivalent to $\mathcal{T}'$, then the answer is "no" even for the basic expressive DL $\mathcal{ALC}$. But logical equivalence is an unnecessarily strong requirement in the context of ontology-mediated querying where it typically suffices to demand *CQ-inseparability*, that is, $\mathcal{T}$ and $\mathcal{T}'$ should give exactly the same answers to any CQ on any ABox; see for example [26, 6, 7] for more on CQ-inseparability. For an expressive DL $\mathcal{L}$, we are thus interested in the following property: we say that *rewritability into CQ-inseparable Horn-TBoxes captures* PTIME *query evaluation* if for every $\mathcal{L}$ TBox $\mathcal{T}$ such that CQ-evaluation w.r.t $\mathcal{T}$ is in PTIME there exists a Horn-$\mathcal{L}$ TBox $\mathcal{T}'$ such that $\mathcal{T}$ and $\mathcal{T}'$ are CQ-inseparable. Note that when $\mathcal{L}$ satisfies this property, then one can replace any $\mathcal{L}$ TBox $\mathcal{T}$ that enjoys PTIME CQ-evaluation by its rewriting $\mathcal{T}'$ and take advantage of the algorithms available for CQ-evaluation w.r.t. Horn-$\mathcal{L}$ TBoxes without affecting the answers to CQs.

The main result of this paper states that rewritability into CQ-inseparable Horn-TBoxes captures PTIME query evaluation when $\mathcal{L}$ is the class of $\mathcal{ALCHIF}$ TBoxes of depth one in which no role is included in a functional role, where the depth of a TBox is the nesting depth of quantifiers in its concepts. Despite the restriction to depth one, this result is rather general. We have analyzed 411 ontologies from the BioPortal repository. After removing all constructors that do not fall within $\mathcal{ALCHIF}$, 385 ontologies had depth 1 (sometimes modulo an easy equivalent rewriting). Moreover, the pre-processing used in most DL reasoners transforms an input TBox into a TBox of depth one by structural transformation. In our proof we show how to construct from a TBox $\mathcal{T}$ formulated in $\mathcal{L}$ a canonical Horn-TBox $\mathcal{T}_{\text{horn}}$ such that $\mathcal{T}_{\text{horn}}$ is a CQ-inseparable rewriting of $\mathcal{T}$ if and only if CQ-evaluation w.r.t. $\mathcal{T}$ is in PTIME. Informally, $\mathcal{T}_{\text{horn}}$ is constructed by first normalising $\mathcal{T}$ and then composing concept inclusions from the concepts in the normalised TBox. The correctness proof makes use of the fact that for $\mathcal{ALCHIF}$ TBoxes of depth one, CQ-evaluation is in PTIME iff the TBox admits universal models [15, 27].

We then show that this result is optimal in several ways. For example, the restriction regarding the interaction of role inclusions and functional roles cannot be dropped since for unrestricted $\mathcal{ALCHIF}$ TBoxes of depth one, rewritability into CQ-inseparable Horn-TBoxes does not capture PTIME query evaluation. We also show such a negative result for $\mathcal{ALCQ}$ TBoxes of depth one. Both results rely on the unique name assumption, which we generally make in this paper. Regarding TBoxes of depth larger than one, we remark that it is known that for $\mathcal{ALC}$ TBoxes of depth three rewritability into CQ-inseparable Horn-TBoxes does not capture PTIME query evaluation [27]; the status of $\mathcal{ALC}$ TBoxes of depth two is open.

Related Work. Rewritability from one language into another has been studied extensively in description logic. A large body of work investigates (the existence of) rewritings of ontology-mediated queries (OMQs) into an FO query or Datalog query that gives the same answers over all ABoxes [3, 4, 13]. The main difference to the work presented in this paper is that both the TBox and the CQ are given as input whereas in this paper we quantify over all CQs. In [18, 17, 10], the authors consider Horn-DL and $\mathcal{EL}$ rewritability of OMQs with atomic queries. Also closely related is work on the rewritability of TBoxes formulated in an expressive DL into a TBox formulated in a weaker DL that is either equivalent to or a conservative extension of the original TBox [25, 20].

2 Preliminaries

We use the notation from [2]. Let $\mathsf{N_C}$ and $\mathsf{N_R}$ be countably infinite sets of concept and role names, respectively. A *role* is a role name or the *inverse* r^- of a role name r. For standard DLs $\mathcal{L}$ between $\mathcal{ALC}$ and $\mathcal{ALCHIQ}$ we define $\mathcal{L}$ concept inclusions (CIs) and $\mathcal{L}$ TBoxes in the usual way. The *functionality assertions* in $\mathcal{ALCF}$ and its extensions take the form $\mathsf{func}(r)$, where r is a role name (a role if the DL admits inverse roles). *Role inclusions (RIs)* in $\mathcal{ALCH}$ and its extensions take the form $r \sqsubseteq s$, where r and s are role names (roles if the DL admits inverse roles). The only non-standard DL we consider is $\mathcal{ALCF}_\ell$ which is located between $\mathcal{ALCF}$ and $\mathcal{ALCQ}$ and in which one can use in addition to the constructors of $\mathcal{ALC}$ local functionality restrictions of the form $(\leq 1 r)$. In certain normal forms we also use $\mathcal{ELI}$ concepts which are constructed using $\top$, $\sqcap$, and $\exists r.C$, r a role.

An *ABox* $\mathcal{A}$ is a non-empty finite set of assertions of the form $A(a)$ and $r(a, b)$ with $A \in \mathsf{N_C}$, $r \in \mathsf{N_R}$, and a, b individual names.

Interpretations $\mathcal{I}$ and the extension $C^\mathcal{I}$ of a concept C are defined as usual. An interpretation $\mathcal{I}$ *satisfies* a CI $C \sqsubseteq D$ if $C^\mathcal{I} \subseteq D^\mathcal{I}$, an RI $r \sqsubseteq s$ if $r^\mathcal{I} \subseteq s^\mathcal{I}$, an assertion $A(a)$ if $a \in A^\mathcal{I}$, an assertion $r(a, b)$ if $(a, b) \in r^\mathcal{I}$, and a functionality assertion $\mathsf{func}(r)$ if $r^\mathcal{I}$ is a partial function. Note that we make the *standard name assumption*, that is, individual names are interpreted as themselves. This implies the unique name assumption.

An interpretation $\mathcal{I}$ is a *model* of a TBox $\mathcal{T}$ if it satisfies all inclusions and assertions in $\mathcal{T}$ and $\mathcal{I}$ is a *model* of an ABox $\mathcal{A}$ if it satisfies all assertions in $\mathcal{A}$. We call an ABox $\mathcal{A}$ *satisfiable w.r.t. a TBox* $\mathcal{T}$ if $\mathcal{A}$ and $\mathcal{T}$ have a common model.

The *depth* of an $\mathcal{ALCI}$ concept is the maximal number of nestings of the operators $\exists r.C$ and $\forall r.C$ in it; thus $\exists r.A$ has depth 1 and $\exists r.\forall r.A$ has depth 2. For DLs with number restrictions, nestings of number restrictions also contribute to the depth. The *depth* of a TBox is the maximal depth of the concepts that occur in it.

A Horn-$\mathcal{ALCI}$ CI takes the form $L \sqsubseteq R$, where L and R are built according to the following syntax rules:

$$R, R' ::= \top \mid \bot \mid A \mid \neg A \mid R \sqcap R' \mid \neg L \sqcup R \mid \exists r.R \mid \forall r.R$$
$$L, L' ::= \top \mid \bot \mid A \mid L \sqcap L' \mid L \sqcup L' \mid \exists r.L$$

A Horn-$\mathcal{ALCHIF}$ TBox is a finite set of Horn-$\mathcal{ALCI}$ CIs, RIs, and functionality assertions. Note that there are several alternative ways to define Horn-DLs [16, 24, 12, 19], our definition is from [27]. The results in this paper remain valid under the alternative definitions.

For a TBox $\mathcal{T}$, ABox $\mathcal{A}$ and conjunctive query (CQ) $q(\boldsymbol{x})$ we say that a tuple $\boldsymbol{a}$ of individuals in $\mathcal{A}$ of the same length as $\boldsymbol{x}$ is a *certain answer to* $q(\boldsymbol{x})$ *over* $\mathcal{A}$ *w.r.t.* $\mathcal{T}$, in symbols $\mathcal{T}, \mathcal{A} \models q(\boldsymbol{a})$, if $\mathcal{I} \models q(\boldsymbol{a})$ holds for all models $\mathcal{I}$ of $\mathcal{T}$ and $\mathcal{A}$. The *query evaluation problem for* $\mathcal{T}$ *and CQ* q is the problem to decide for an ABox $\mathcal{A}$ and a tuple $\boldsymbol{a}$ of individuals from $\mathcal{A}$, whether $\mathcal{T}, \mathcal{A} \models q(\boldsymbol{a})$. The

CQ-evaluation problem for $\mathcal{T}$ is in PTime if the query evaluation problem for $\mathcal{T}$ and q is in PTime for every CQ q.

The depth of TBoxes will play an important role in this paper. For deciding satisfiability and subsumption, TBoxes are often normalized to depth 1 in a pre-processing step. Since we are universally quantifying over all queries when defining the complexity of a TBox, such normalizations do not work in the sense that they can change the complexity of the TBox, see [27, 15].

To show that a TBox cannot be rewritten into a CQ-inseparable Horn-TBox, we shall sometimes use products of interpretations; we only need the product of two interpretations, see [11, 25] for the general case. Let $\mathcal{I}_1$ and $\mathcal{I}_2$ be interpretations. Then the *product* $\mathcal{I}_1 \times \mathcal{I}_2$ of $\mathcal{I}_1$ and $\mathcal{I}_2$ is defined by setting

$$\begin{aligned}
\Delta^{\mathcal{I}_1 \times \mathcal{I}_2} &= \Delta^{\mathcal{I}_1} \times \Delta^{\mathcal{I}_2} \\
A^{\mathcal{I}_1 \times \mathcal{I}_2} &= \{(d_1, d_2) \mid d_1 \in A^{\mathcal{I}_1}, d_2 \in A^{\mathcal{I}_2}\} \\
r^{\mathcal{I}_1 \times \mathcal{I}_2} &= \{((d_1, d_2), (e_1, e_2)) \mid (d_1, e_1) \in r^{\mathcal{I}_1}, (d_2, e_2) \in r^{\mathcal{I}_2}\}
\end{aligned}$$

Lemma 1. *Every Horn-$\mathcal{ALCHIQ}$ TBox $\mathcal{T}$ is* preserved under products*: if $\mathcal{I}_1$ and $\mathcal{I}_2$ are models of $\mathcal{T}$, then $\mathcal{I}_1 \times \mathcal{I}_2$ is a model of $\mathcal{T}$.*

3 Horn Rewritability

In this paper, we aim to understand whether and when a TBox formulated in an expressive DL can be replaced with a TBox formulated in the corresponding Horn-DL (e.g. an $\mathcal{ALCHIF}$ TBox by a Horn-$\mathcal{ALCHIF}$ TBox) without changing the answers to CQs. Following [26, 6, 7], TBoxes $\mathcal{T}_1$ and $\mathcal{T}_2$ are *CQ-inseparable* if for all CQs q, all ABoxes $\mathcal{A}$, and all tuples $\boldsymbol{a}$ of individual names in $\mathcal{A}$ we have $\mathcal{T}_1, \mathcal{A} \models q(\boldsymbol{a})$ iff $\mathcal{T}_2, \mathcal{A} \models q(\boldsymbol{a})$. We say that *rewritability into CQ-inseparable Horn-TBoxes captures* PTime *query evaluation for a DL $\mathcal{L}$* if for every $\mathcal{L}$ TBox $\mathcal{T}$ such that CQ evaluation w.r.t. $\mathcal{T}$ is in PTime, there is a Horn-$\mathcal{L}$ TBox $\mathcal{T}'$ such that $\mathcal{T}$ and $\mathcal{T}'$ are CQ-inseparable. The following example shows that rewritability into a CQ-inseparable Horn-TBox is a weaker notion than rewritability into a Horn-TBox that is logically equivalent.

Example 1. Consider the Horn-$\mathcal{ALC}$ TBox

$$\mathcal{T} = \{\exists r.(A \sqcap \neg B \sqcap \neg E) \sqsubseteq \exists r.(\neg A \sqcap \neg B \sqcap \neg E)\}.$$

It is easy to see that for any CQ $q(\boldsymbol{x})$, ABox $\mathcal{A}$, and tuple $\boldsymbol{a}$ of individuals from $\mathcal{A}$, we have $\mathcal{T}, \mathcal{A} \models q(\boldsymbol{a})$ iff $\emptyset, \mathcal{A} \models q(\boldsymbol{a})$. Thus, CQ evaluation w.r.t. $\mathcal{T}$ is in PTime (actually in AC^0). $\mathcal{T}$ is, however, not equivalent to any TBox preserved under products: consider interpretations $\mathcal{I}_1$ and $\mathcal{I}_2$ with $\Delta^{\mathcal{I}_i} = \{0, 1\}$, $r^{\mathcal{I}_i} = \{(0, 1)\}$, and $A^{\mathcal{I}_i} = \{1\}$ for $i = 1, 2$, but $B^{\mathcal{I}_1} = \{1\}$, $E^{\mathcal{I}_1} = \emptyset$, $B^{\mathcal{I}_2} = \emptyset$, and $E^{\mathcal{I}_2} = \{1\}$. Then $\mathcal{I}_i$ is a model of $\mathcal{T}$ for $i = 1, 2$, but $\mathcal{I}_1 \times \mathcal{I}_2$ is not. Consequently, $\mathcal{T}$ is not equivalent to any Horn-TBox.

We shall make use of several characterizations of CQ evaluation w.r.t. a TBox being in PTime that have been obtained in [15]. A TBox $\mathcal{T}$ is *materializable* if for any ABox $\mathcal{A}$ that is satisfiable w.r.t $\mathcal{T}$ there exists a model $\mathcal{I}$ of $\mathcal{T}$ and $\mathcal{A}$ such that for all CQs $q(\boldsymbol{x})$ and all tuples $\boldsymbol{a}$ of individuals in $\mathcal{A}$, $\mathcal{T}, \mathcal{A} \models q(\boldsymbol{a})$ iff $\mathcal{I} \models q(\boldsymbol{a})$. $\mathcal{T}$ has the *disjunction property for CQs* if for any ABox $\mathcal{A}$, CQs $q_1, \dots, q_n$ and tuples $\boldsymbol{a}_1, \dots, \boldsymbol{a}_n$ of individuals in $\mathcal{A}$, $\mathcal{T}, \mathcal{A} \models \bigvee_{1 \leq i \leq n} q_i(\boldsymbol{a}_i)$ implies that there is a j with $\mathcal{T}, \mathcal{A} \models q_j(\boldsymbol{a}_j)$. An *ontology-mediated query (OMQ)* is a pair $(\mathcal{T}, q)$ with $\mathcal{T}$ a TBox and q a CQ. A Datalog$^{\neq}$ program is a Datalog program that admits inequalities in the body of its rules. We say that $(\mathcal{T}, q)$ is *Datalog$^{\neq}$ rewritable* if there is a Datalog$^{\neq}$-program Π such that for any ABox $\mathcal{A}$ and tuple $\boldsymbol{a}$ of individuals in $\mathcal{A}$, $\mathcal{T}, \mathcal{A} \models q(\boldsymbol{a})$ iff $\mathcal{A} \models \Pi(\boldsymbol{a})$.

Theorem 1. [15] *The following conditions are equivalent for all $\mathcal{ALCHIF}$ TBoxes $\mathcal{T}$ of depth 2 and all $\mathcal{ALCHIQ}$ TBoxes $\mathcal{T}$ of depth 1:*

1. *CQ evaluation w.r.t. $\mathcal{T}$ is in* PTime*;*
2. *$\mathcal{T}$ is materializable;*
3. *$\mathcal{T}$ has the disjunction property for CQs;*
4. *For every CQ q, the OMQ $(\mathcal{T}, q)$ is Datalog$^{\neq}$ rewritable.*

Otherwise, CQ evaluation w.r.t. $\mathcal{T}$ is coNP*-hard.*

Theorem 1 is optimal in many ways. For example, the following is shown in [15]:

– the equivalence of Datalog$^{\neq}$-rewritability and PTime CQ evaluation does not hold for $\mathcal{ALCF}_\ell$ TBoxes of depth 2 and $\mathcal{ALC}$ TBoxes of depth 3.
– there are $\mathcal{ALCIQ}$ TBoxes $\mathcal{T}$ of depth 2 such that CQ-evaluation w.r.t. $\mathcal{T}$ is neither in PTime nor coNP-hard.

4 Negative Results

It follows from results in [15] that there are DLs in which rewritability into CQ-inseparable Horn-TBoxes does not capture PTime query evaluation. In fact, we have seen above that there are $\mathcal{ALCF}_\ell$ TBoxes of depth 2 and $\mathcal{ALC}$ TBoxes of depth 3 such that CQ evaluation w.r.t. $\mathcal{T}$ is in PTime but not every OMQ $(\mathcal{T}, q)$ is Datalog$^{\neq}$-rewritable. On the other hand, it is well known that every OMQ $(\mathcal{T}, q)$ with $\mathcal{T}$ a Horn-$\mathcal{ALCHIQ}$ TBox is Datalog$^{\neq}$-rewritable [15]. We thus obtain the following.

Theorem 2. *For $\mathcal{ALCF}_\ell$ TBoxes of depth 2 and $\mathcal{ALC}$ TBoxes of depth 3, rewritability into CQ-inseparable Horn-TBoxes does not capture* PTime *query evaluation.*

We next prove a negative result for $\mathcal{ALCQ}$ TBoxes of depth 1. Note that by Theorem 1, any $\mathcal{ALCQ}$ TBox $\mathcal{T}$ of depth 1 such that CQ evaluation w.r.t. $\mathcal{T}$ is in PTime is rewritable into Datalog$^{\neq}$. We thus need a different argument as in the proof of Theorem 2. This argument is based on products and actually establishes a stronger statement than aimed at, implying for example that not even rewritability into CQ-inseparable *Horn-FO* TBoxes captures PTime query evaluation.

Theorem 3. *For $\mathcal{ALCQ}$ TBoxes of depth 1, rewritability into CQ-inseparable Horn-TBoxes does not capture* PTIME *query evaluation.*

Proof. Consider the TBox $\mathcal{T} = \{(\geq 3r) \sqsubseteq A\}$. It is easy to show that $\mathcal{T}$ is materializable. Thus, by Theorem 1, CQ evaluation w.r.t. $\mathcal{T}$ is in PTIME. Assume that $\mathcal{T}$ is CQ-inseparable from a TBox $\mathcal{T}'$. We show that $\mathcal{T}'$ is not preserved under products and thus not a Horn-$\mathcal{ALCQ}$ TBox. We have $\mathcal{T}' \models (\geq 3r) \sqsubseteq A$ since, for $\mathcal{A} = \{r(a, b_1), r(a, b_2), r(a, b_3)\}$, we have $\mathcal{T}, \mathcal{A} \models A(a)$ and thus $\mathcal{T}', \mathcal{A} \models A(a)$. We also have $\mathcal{T}' \not\models (\geq 2r) \sqsubseteq A$ since, for $\mathcal{A} = \{r(a, b_1), r(a, b_2)\}$, we have $\mathcal{T}, \mathcal{A} \not\models A(a)$ and thus $\mathcal{T}', \mathcal{A} \not\models A(a)$. Take a model $\mathcal{I}$ of $\mathcal{T}'$ with $d \in (\geq 2r)^{\mathcal{I}}$ but $d \notin A^{\mathcal{I}}$. Then $(d, d) \in (\geq 3r)^{\mathcal{I} \times \mathcal{I}}$ and $(d, d) \notin A^{\mathcal{I} \times \mathcal{I}}$. Thus $\mathcal{I} \times \mathcal{I}$ is not a model of $\mathcal{T}'$. □

We now prove a similar negative result for $\mathcal{ALCHIF}$ TBoxes of depth 1.

Theorem 4. *For $\mathcal{ALCHIF}$ TBoxes of depth 1, rewritability into CQ-inseparable Horn-TBoxes does not capture* PTIME *query evaluation.*

Proof. Let $\mathcal{T}$ be the $\mathcal{ALCHIF}$ TBox that states that role names s_1 and s_2 are functional and contains the RIs $r \sqsubseteq s_1$ and $r \sqsubseteq s_2$ and the CIs

$$\exists s_1.(B_1 \sqcap B_2) \sqsubseteq \exists r.\top$$
$$\exists s_1.\top \sqcap \exists s_2.\top \sqsubseteq \forall s_1.B_1 \sqcap \forall s_2.B_2$$
$$\exists s_1.\top \sqcap \exists s_2.\top \sqsubseteq B \sqcup \exists r.\top$$

We first show that $\mathcal{T}$ is materializable and so, by Theorem 1, CQ evaluation w.r.t. $\mathcal{T}$ is in PTIME. Let $\mathcal{A}$ be an ABox. Due to the disjunction in the final CI, we have to distinguish two kinds of individuals a when constructing a materialization of $\mathcal{A}$. Informally:

– if a has distinct s_1- and s_2-successors, then a having an r-successor contradicts the role inclusions and s_1, s_2 being functional and thus we want to make B true at a;
– if a has a common s_1- and s_2-successor, then $\exists r.\top$ is implied at a.

Formally, we construct the materialization of $\mathcal{A}$ as follows. Whenever a has both an s_1-successor and an s_2-successor, then add B_i to all its s_i-successors, $i \in \{1, 2\}$. Next, for any a that has an s_1-successor b in $B_1 \sqcap B_2$, add $r(a, b)$ and $s_2(a, b)$ to the ABox. Denote the resulting ABox by $\mathcal{A}'$. If s_1 or s_2 are not functional in $\mathcal{A}'$, then $\mathcal{A}$ is not satisfiable w.r.t. $\mathcal{T}$. Otherwise, a materialization is obtained by adding $B(a)$ whenever a has distinct s_1- and s_2-successors.

Assume that $\mathcal{T}'$ is CQ-inseparable from the TBox $\mathcal{T}$. We show that $\mathcal{T}'$ is not preserved under products, thus not a Horn-$\mathcal{ALCHIF}$ TBox. Let

$$\mathcal{A}_1 = \{s_1(a, b_1), s_2(a, b_2)\}, \quad \mathcal{A}_2 = \{s_1(a, b), s_2(a, b)\}.$$

Then $\mathcal{T}, \mathcal{A}_1 \models B(a)$ and $\mathcal{T}, \mathcal{A}_2 \models \exists r.\top(a)$, but $\mathcal{T}, \mathcal{A}_1 \not\models \exists r.\top(a)$ and $\mathcal{T}, \mathcal{A}_2 \not\models B(a)$. Because of CQ-inseparability, the same hold when $\mathcal{T}$ is replaced with $\mathcal{T}'$.

Take a model $\mathcal{I}_1$ of $\mathcal{T}'$ and $\mathcal{A}_1$ with $a \notin (\exists r.\top)^{\mathcal{I}_1}$ and a model $\mathcal{I}_2$ of $\mathcal{T}'$ and $\mathcal{A}_2$ with $a \notin B^{\mathcal{I}_2}$. Then $\mathcal{I}_1 \times \mathcal{I}_2$ is a model of $\mathcal{A}_1$ (for a identified with (a, a), b_1 with (b_1, b) and b_2 with (b_2, b)). However, we do not have $a \in B^{\mathcal{I}_1 \times \mathcal{I}_2}$ and since $\mathcal{T}', \mathcal{A}_1 \models B(a)$ it follows that $\mathcal{I}_1 \times \mathcal{I}_2$ is not a model of $\mathcal{T}'$. $\square$

5 Horn Rewritability in $\mathcal{ALCHIF}$

We introduce a mild restriction on $\mathcal{ALCHIF}$ TBoxes regarding the interaction between RIs and functionality assertions and show that this restriction is sufficient to overcome the negative result stated in Theorem 4. Notably, the restricted form of $\mathcal{ALCHIF}$ TBoxes encompasses both (unrestricted) $\mathcal{ALCHI}$ TBoxes and $\mathcal{ALCIF}$ TBoxes.

An $\mathcal{ALCHIF}^{\sqsubseteq f}$ TBox is an $\mathcal{ALCHIF}$ TBox $\mathcal{T}$ such that whenever $r \sqsubseteq s \in \mathcal{T}$, then neither s nor s^- are functional in $\mathcal{T}$. The aim of this section is to prove the following result.

Theorem 5. *For $\mathcal{ALCHIF}^{\sqsubseteq f}$ TBoxes of depth 1, rewritability into CQ-inseparable Horn-TBoxes captures* PTime *query evaluation.*

We start by introducing a normal form for $\mathcal{ALCHIF}$ TBoxes of depth 1. A *literal* is a concept name or a negation thereof. A CI $C \sqsubseteq D$ is in *normal form* if

1. C is an $\mathcal{ELI}$-concept of depth 1;
2. D is a disjunction of
   - concept names;
   - concepts $\exists r.E$ with E a conjunction of literals;
   - concepts $\forall r.E$ with E a disjunction of literals that contains at least one positive literal.

We set $C = \top$ if C is the empty conjunction and $D = \bot$ if D is the empty disjunction. Given a set $\mathcal{F}$ of functionality assertions $\mathsf{func}(r)$, we can normalise CIs further by demanding that in $C \sqsubseteq D$ for each r such that $\mathsf{func}(r) \in \mathcal{F}$:

- any $\exists r.D'$ in D contains only positive literals;
- if there exists an $\exists r.C'$ in C, then this is the only occurrence of r in $C \sqsubseteq D$;
- if there exists a $\forall r.D'$ in D, then this is the only occurrence of r in $C \sqsubseteq D$.

We then call $C \sqsubseteq D$ in *normal form relative to* $\mathcal{F}$. An $\mathcal{ALCHIF}$ TBox $\mathcal{T}$ is in *normal form* if all its CIs are in normal form relative to the functionality assertions in $\mathcal{T}$. The following result is shown in the appendix.

Lemma 2. *Every $\mathcal{ALCHIF}^{\sqsubseteq f}$ TBox $\mathcal{T}$ of depth 1 can be converted into a logically equivalent $\mathcal{ALCHIF}^{\sqsubseteq f}$ TBox $\mathcal{T}'$ in normal form. $\mathcal{T}'$ is of size at most single exponential in $|\mathcal{T}|$.*

Example 2. The TBox $\mathcal{T}$ from Example 1 can be transformed into the TBox

$$\mathcal{T}' = \{\exists r.A \sqsubseteq \exists r.(\neg A \sqcap \neg B \sqcap \neg E) \sqcup \forall r.(A \rightarrow B \sqcup E)\},$$

which is in normal form. If r is functional, one obtains $\{\top \sqsubseteq \forall r.(A \rightarrow B \sqcup E)\}$.

We now identify a basic yet crucial property of $\mathcal{ALCHIF}^{\sqsubseteq f}$ TBoxes $\mathcal{T}$. For any $\mathcal{ELI}$ concept C of depth 1 that contains at most one conjunct of the form $\exists r.C'$ per functional role r one can define in a straightforward way a tree-shaped ABox $\mathcal{A}_C$ of depth 1 with root ρ_C that corresponds to C. For example, when $C = \exists r.\top \sqcap \exists s.(A \sqcap B)$ then $\mathcal{A}_C = \{r(\rho_C, b_1), s(\rho_C, b_2), A(b_2), B(b_2)\}$. Then one can prove the following:

Lemma 3. *For every $\mathcal{ALCI}$-concept D, $\mathcal{T} \models C \sqsubseteq D$ iff $\mathcal{T}, \mathcal{A}_C \models D(\rho_C)$.*

Note that Lemma 3 fails for unrestricted $\mathcal{ALCHIF}$ TBoxes. Consider for example the TBox $\mathcal{T}$ and ABox $\mathcal{A}_1$ from the proof of Theorem 4 and let $C_{\mathcal{A}_1}$ be $\mathcal{A}_1$ viewed as an $\mathcal{EL}$-concept. Then $\mathcal{T} \not\models C_{\mathcal{A}_1} \sqsubseteq B$, but $\mathcal{T}, \mathcal{A}_1 \models B(a)$.

We next introduce some preliminaries needed for constructing the desired Horn-$\mathcal{ALCHIF}^{\sqsubseteq f}$ TBoxes that are CQ-inseparable from a given $\mathcal{ALCHIF}^{\sqsubseteq f}$ TBox of depth 1. For any conjunction or disjunction of literals E, we use $\mathsf{pos}(E)$ to denote the set of concept names A in E and $\mathsf{neg}(E)$ to denote the set of concept names A such that $\neg A$ in E. If $\forall r.E$ is a universal restriction with E a disjunction of literals that contains at least one positive literal, then a *Horn specialization* of $\forall r.E$ is a concept $\forall r.E'$ where E' is obtained from E by dropping all but one positive literal. Note that any Horn specialization can be written in the form $\forall r.(A_1 \sqcap \cdots \sqcap A_n \rightarrow A)$.

For an $\mathcal{ALCHIF}^{\sqsubseteq f}$ TBox $\mathcal{T}$ in normal form, we use $L_\mathcal{T}$ to denote the set of

- concept names or existential restrictions that occur on top-level on the left-hand side of some CI in $\mathcal{T}$;
- concepts $\exists r.\mathsf{neg}(E)$ such that there is a CI in $\mathcal{T}$ whose right-hand side contains a disjunct $\forall r.E$.

A set $S \subseteq L_\mathcal{T}$ is a *trigger for* a CI $C \sqsubseteq D \in \mathcal{T}$ if S contains all top-level conjuncts of C and all $\exists r.\mathsf{neg}(E)$ with $\forall r.E$ a disjunct of D. For a trigger S, we denote by C_S the conjunction of all concept names in S, all existential restrictions $\exists r.C' \in S$ with r not functional, and all $\exists r.(C_1 \sqcap \cdots \sqcap C_n)$ where r is functional and $C_1, \ldots, C_n$ is the set of all concepts such that $\exists r.C_i \in S$. For each CI $C \sqsubseteq D \in \mathcal{T}$ and trigger S for it, we define a set $\mathsf{Horn}(C \sqsubseteq D, S)$ of Horn-$\mathcal{ALCHIF}^{\sqsubseteq f}$ CIs. As a special case, $\mathsf{Horn}(C \sqsubseteq D, S)$ is $\{C_S \sqsubseteq \bot\}$ if $\mathcal{T} \models C_S \sqsubseteq \bot$. Otherwise, $\mathsf{Horn}(C \sqsubseteq D, S)$ consists of the following CIs whenever they are a consequence of $\mathcal{T}$:

- $C_S \sqsubseteq A$ with A a concept name in D;
- $C_S \sqsubseteq R$ with $R = \forall r.(A_1 \sqcap \cdots \sqcap A_n \rightarrow A)$ a Horn restriction of some universal restriction in D;
- $C_S \sqsubseteq \exists r.\mathsf{pos}(E)$ with $\exists r.E$ an existential restriction in D such that $\mathcal{T} \not\models C_S \sqsubseteq \neg\exists r.E$.

Define a Horn-$\mathcal{ALCHIF}$ TBox $\mathcal{T}_{\mathrm{horn}}$ as the union of all functionality assertions and RIs in $\mathcal{T}$ and

$$\bigcup_{C \sqsubseteq D \in \mathcal{T},\ S \text{ trigger for } C \sqsubseteq D} \mathsf{Horn}(C \sqsubseteq D, S)$$

Observe that, by construction, $\mathcal{T} \models \mathcal{T}_{\text{horn}}$.

Example 3. For the TBox $\mathcal{T}'$ from Example 2 we have $L_{\mathcal{T}'} = \{\exists r.A, \exists r.\top\}$ and thus $\{\exists r.A\}$ is the only trigger for $\exists r.A \sqsubseteq \exists r.(\neg A \sqcap \neg B \sqcap \neg E) \sqcup \forall r.(A \rightarrow B \sqcup E)$ up to logical equivalence. Thus, $\mathcal{T}'_{\text{horn}} = \{\exists r.A \sqsubseteq \exists r.\top\}$ which is logically equivalent to the empty TBox.

Now, Theorem 5 is a consequence of Theorem 1 and the "(1.) $\Rightarrow$ (3.)" part of the following result.

Theorem 6. *Let $\mathcal{T}$ be a $\mathcal{ALCHIF}^{\sqsubseteq f}$ TBox in normal form. Then the following conditions are equivalent:*

1. *$\mathcal{T}$ has the disjunction property for CQs;*
2. *for every $C \sqsubseteq D \in \mathcal{T}$ and trigger S for $C \sqsubseteq D$, $\mathsf{Horn}(C \sqsubseteq D, S) \neq \emptyset$.*
3. *$\mathcal{T}$ and $\mathcal{T}_{horn}$ are CQ-inseparable.*

Proof. (3) $\Rightarrow$ (1.) follows from Theorem 1 since CQ-evaluation is in PTime if $\mathcal{T}$ and $\mathcal{T}_{\text{horn}}$ are CQ-inseparable.

(1.) $\Rightarrow$ (2.). Assume $\mathcal{T}$ has the disjunction property for CQs and assume that there are $C \sqsubseteq D \in \mathcal{T}$ and a trigger S for $C \sqsubseteq D$ such that $\mathsf{Horn}(C \sqsubseteq D, S) = \emptyset$. Let $A_1, \ldots, A_k$, $\exists r_1.E_1, \ldots, \exists r_n.E_n$, and $\forall s_1.F_1, \ldots, \forall s_m.F_m$ be the disjuncts of D such that $\mathcal{T} \not\models C_S \sqsubseteq \neg\exists r_i.E_i$ for $1 \leq i \leq n$. Let $\mathcal{A}_{C_S}$ be C_S viewed as an ABox with root a_0, and let $a_1, \ldots, a_m$ be the individuals in $\mathcal{A}_{C_S}$ introduced for the existential restrictions $\exists s_i.\mathsf{neg}(F_i)$ in C_S. By Lemma 3, we have:

$$\begin{aligned}\mathcal{T}, \mathcal{A}_{C_S} \models & \bigvee_{i=1..k} A_i(a_0) \vee \bigvee_{i=1..n} \exists r_i.\mathsf{pos}(E_i)(a_0) \\ & \vee \bigvee_{i=1..m} \bigvee_{A \in \mathsf{pos}(F_i)} A(a_i)\end{aligned}$$

By the disjunction property for CQs of $\mathcal{T}$ and Lemma 3 at least one of the disjuncts $F(a)$ is entailed. We obtain:

- if $F(a) = A_i(a_0)$ then $\mathcal{T} \models C_S \sqsubseteq A_i$;
- if $F(a) = A(a_i)$ for some $A \in \mathsf{pos}(F_i)$, then $\mathcal{T} \models C_S \sqsubseteq \forall s_i.(\mathsf{neg}(F_i) \rightarrow A)$;
- if $F(a) = \exists r_i.\mathsf{pos}(E_i)(a_0)$, then $\mathcal{T} \models C_S \sqsubseteq \exists r_i.\mathsf{pos}(E_i)$.

In each case, we obtain an $\alpha \in \mathsf{Horn}(C \sqsubseteq D, S)$ and thus derive a contradiction.

(2.) $\Rightarrow$ (3.). Assume (2.) holds. It suffices to show that for every ABox $\mathcal{A}$ the following holds: if $\mathcal{A}$ is satisfiable relative to $\mathcal{T}_{\text{horn}}$, then there exists a materialization $\mathcal{U}$ of $\mathcal{A}$ and $\mathcal{T}_{\text{horn}}$ that is a model of $\mathcal{T}$. $\mathcal{U}$ is constructed using the following set $R_{\mathcal{T}}$ of chase rules:

1. if $(d, e) \in r^{\mathcal{I}}$ and $r \sqsubseteq s \in \mathcal{T}$, then add (d, e) to $s^{\mathcal{I}}$;
2. if $d \in C^{\mathcal{I}}$ and $C \sqsubseteq A \in \mathcal{T}_{\text{horn}}$, then add d to $A^{\mathcal{I}}$;
3. if $d \in C^{\mathcal{I}}$ and $C \sqsubseteq R \in \mathcal{T}_{\text{horn}}$ for $R = \forall r.(A_1 \sqcap \cdots \sqcap A_n \rightarrow A)$, then add e to $A^{\mathcal{I}}$ whenever $(d, e) \in r^{\mathcal{I}}$ and $e \in (A_1 \sqcap \cdots \sqcap A_n)^{\mathcal{I}}$.

4. if $d \in C^{\mathcal{I}}$ and $C \sqsubseteq \exists r.\mathsf{pos}(E) \in \mathcal{T}_{\text{horn}}$ then (i) if r is functional and there exists e with $(d,e) \in r^{\mathcal{I}}$ add e to $F^{\mathcal{I}}$ for all concept names F in $\mathsf{pos}(E)$ and (ii) otherwise take a fresh element e, add (d,e) to $r^{\mathcal{I}}$ and add e to $F^{\mathcal{I}}$ for all concept names in $\mathsf{pos}(E)$. This rule is applied at most once for any pair $(d, \exists r.E)$ with $\exists r.E$ on the right hand side of a CI in $\mathcal{T}_{\text{horn}}$.
   The element e required by this rule is called the *witness for $\exists r.E$ at d*.

The interpretation $\mathcal{U}$ is the limit of the sequence $\mathcal{I}_0, \mathcal{I}_1, \ldots$ obtained from the interpretation $\mathcal{I}_0$ corresponding to the ABox $\mathcal{A}$ by applying the rules in $R_{\mathcal{T}}$ in a fair way. It is standard to prove that $\mathcal{U}$ is a materialization of $\mathcal{A}$ and $\mathcal{T}_{\text{horn}}$ if $\mathcal{A}$ is satisfiable w.r.t. $\mathcal{T}_{\text{horn}}$ (here one uses that $\mathcal{T}$ does not contain any functional roles s with $\mathcal{T} \models r \sqsubseteq s$ for an $r \neq s$).

Our aim now is to prove that $\mathcal{U}$ is a model of $\mathcal{T}$. We proceed in two steps. Let $d \in \Delta^{\mathcal{U}}$ and assume that the chase introduces e as a witness for $\exists r.E$ at d. We say that this witness has been *invalidated* if $e \notin E^{\mathcal{U}}$. Observe that if r is a functional role then for any $C \sqsubseteq \exists r.\mathsf{pos}(E) \in \mathcal{T}_{\text{horn}}$ we have $E = \mathsf{pos}(E)$. Thus for functional roles r no witnesses for $\exists r.E$ can be invalidated. For an interpretation $\mathcal{I}$ and $d \in \Delta^{\mathcal{I}}$ we denote by $\mathsf{eltp}_{\mathcal{U}}^{\mathcal{T}}(d)$ the set of all $F \in L_{\mathcal{T}}$ such that $d \in F^{\mathcal{I}}$.

Claim 1. If the witness for $\exists r.E$ at d was invalidated, then $\mathcal{T}_{\text{horn}} \models C_S \sqsubseteq \neg\exists r.E$ for $S = \mathsf{eltp}_{\mathcal{U}}^{\mathcal{T}}(d)$.

For the proof of Claim 1, denote by $\mathcal{A}_{C_S}$ the ABox with root ρ_S corresponding to C_S. Denote by $\mathcal{A}_{C_S}^{\exists r.\mathsf{pos}(E)}$ the extension of $\mathcal{A}_{C_S}$ with a fresh individual e and fresh assertions $r(\rho_F, e)$ and $A(e)$ for $A \in \mathsf{pos}(E)$. We now apply the chase procedure for $\mathcal{T}_{\text{horn}}$ to $\mathcal{A}_{C_S}^{\exists r.\mathsf{pos}(E)}$ rather that $\mathcal{A}$ and obtain a materialization $\mathcal{U}'$ of $\mathcal{A}_{C_S}^{\exists r.\mathsf{pos}(E)}$ and $\mathcal{T}_{\text{horn}}$. Using the fact that the CIs of $\mathcal{T}_{\text{horn}}$ have depth 1 and the condition that the witness for $\exists r.E$ at d was invalidated in the chase applied to $\mathcal{A}$ one can readily show that it is invalidated in $\mathcal{A}_{C_S}^{\exists r.\mathsf{pos}(E)}$ as well. Thus $e \notin E^{\mathcal{U}'}$ and it follows with Lemma 3 that $\mathcal{T}_{\text{horn}} \models C_S \sqsubseteq \neg\exists r.E$.

Claim 2. $\mathcal{U}$ is a model of $\mathcal{T}$.

Let $C \sqsubseteq D \in \mathcal{T}$ and $d \in C^{\mathcal{U}}$. We show that $d \in D^{\mathcal{U}}$. For a proof by contradiction assume that this is not the case. Then the set $S = \mathsf{eltp}_{\mathcal{U}}^{\mathcal{T}}(d)$ is a trigger for $C \sqsubseteq D$. By (2.) there exists $\alpha \in \text{Horn}(C \sqsubseteq D, S)$. We obtain that at least one of the following holds:

1. there is a concept name A in D such that $\mathcal{T} \models C_S \sqsubseteq A$. Then $C_S \sqsubseteq A \in \mathcal{T}_{\text{horn}}$ and so $d \in A^{\mathcal{U}}$. Thus $d \in D^{\mathcal{U}}$ and we have derived a contradiction.
2. there is a universal restriction $R = \forall r.(A_1 \sqcap \cdots \sqcap A_n \rightarrow A)$ that is a Horn restriction of some universal restriction in D such that $\mathcal{T} \models C_S \sqsubseteq R$. Then $C_S \sqsubseteq R \in \mathcal{T}_{\text{horn}}$ and so $d \in R^{\mathcal{U}}$. Thus $d \in D^{\mathcal{U}}$ and we have derived a contradiction.
3. there is an $\exists r.E$ in D such that $\mathcal{T} \models C_S \sqsubseteq \exists r.\mathsf{pos}(E)$ and $\mathcal{T} \not\models C_S \sqsubseteq \neg\exists r.E$. Then there exists $e \in \Delta^{\mathcal{U}}$ with $(d,e) \in r^{\mathcal{U}}$ such that $e \in \mathsf{pos}(E)^{\mathcal{U}}$. By Claim 1, the witness e for $\exists r.E$ at d was not invalidated. Thus $d \in (\exists r.E)^{\mathcal{U}}$. Hence $d \in D^{\mathcal{U}}$ and we have derived a contradiction.

This finishes the proof of Theorem 6. □

The following example shows that $\mathcal{T}_{\mathrm{horn}}$ can be of exponential size in $|\mathcal{T}|$ even if redundant CIs are removed.

Example 4. Let $\mathcal{T}^n$ contain

$$\begin{aligned}\exists r.\top &\sqsubseteq \exists s.\top \\ B_i &\sqsubseteq \forall s.A_i \\ \hat{B}_i &\sqsubseteq \forall s.A_i \\ \exists r.\top &\sqsubseteq A \sqcup \bigsqcup_{1\leq i\leq n} \exists s.\neg A_i\end{aligned}$$

where $1 \leq i \leq n$. The TBox $\mathcal{T}^n_{\mathrm{horn}}$ is obtained from $\mathcal{T}_n$ by replacing the final CI by the set of CIs

$$\exists r.\top \sqcap \bigsqcap_{1\leq i\leq n} C_i \sqsubseteq A$$

where $C_i \in \{B_i, \hat{B}_i\}$. Then $\mathcal{T}^n_{\mathrm{horn}}$ and $\mathcal{T}^n$ are CQ-inseparable and $\mathcal{T}^n_{\mathrm{horn}}$ is of exponential size in $|\mathcal{T}^n|$.

Clearly, the TBoxes $\mathcal{T}^n_{\mathrm{horn}}$ can be equivalently expressed by Horn-TBoxes of polynomial size in $|\mathcal{T}^n|$ by introducing disjunctions on the left-had side of CIs. In fact, it remains open whether for every $\mathcal{ALCHIF}^{\sqsubseteq f}$ TBoxes in normal form for which CQ-evaluation is in PTIME there exists a CQ-inseparable Horn-$\mathcal{ALCHIF}^{\sqsubseteq f}$ TBox of polynomial size.

6 Conclusion

We have shown that rewritability into CQ-inseparable Horn-TBoxes captures PTIME query evaluation for $\mathcal{ALCHIF}$ TBoxes of depth 1 in which no role is included in a functional role. Interestingly, this result also implies that CQ-inseparable rewritability into Horn-TBoxes is decidable and EXPTIME-complete for such $\mathcal{ALCHIF}$ TBoxes since it is shown in [15] that deciding whether CQ-evaluation is in PTIME for $\mathcal{ALCHIF}$ TBoxes of depth 1 is EXPTIME-complete. We have also shown that for arbitrary $\mathcal{ALCHIF}$ and $\mathcal{ALCQ}$ TBoxes of depth 1, rewritability into CQ-inseparable Horn-TBoxes does not capture PTIME query evaluation. These negative results depend on the unique name assumption we make in this paper. In fact, it is not difficult to see that CQ-evaluation is coNP-hard for the TBoxes used in the proofs of Theorems 3 and 4 and so they cannot serve as counterexamples anymore. We conjecture that without the unique name assumption one can generalize our positive result and show that rewritability into CQ-inseparable Horn-TBoxes captures PTIME query evaluation for arbitrary $\mathcal{ALCHIQ}$ TBoxes of depth 1. Observe that for TBoxes not using functional roles or qualified number restrictions CQ-evaluation does not depend on whether one makes the unique name assumption or not. Thus, the negative result that for $\mathcal{ALC}$ TBoxes of depth 3 CQ-inseparable Horn-TBoxes do not capture PTIME query evaluation does not depend on the unique name assumption. It remains an interesting open question whether rewritability into CQ-inseparable Horn-TBoxes captures PTIME query evaluation for $\mathcal{ALC}$ or $\mathcal{ALCI}$ TBoxes of depth 2.

Acknowledgements André Hernich, Fabio Papacchini, and Frank Wolter were supported by EPSRC UK grant EP/M012646/1. Carsten Lutz acknowledges support by the DFG-funded Collaborative Research Center EASE.

References

1. Artale, A., Calvanese, D., Kontchakov, R., Zakharyaschev, M.: The DL-Lite family and relations. J. of Artifical Intelligence Research 36, 1–69 (2009)
2. Baader, F., Horrocks, I., Lutz, C., Sattler, U.: An Introduction to Description Logics. Cambride University Press (2017)
3. Bienvenu, M., ten Cate, B., Lutz, C., Wolter, F.: Ontology-based data access: A study through disjunctive datalog, csp, and MMSNP. ACM Trans. Database Syst. 39(4), 33:1–33:44 (2014)
4. Bienvenu, M., Hansen, P., Lutz, C., Wolter, F.: First order-rewritability and containment of conjunctive queries in horn description logics. In: IJCAI. pp. 965–971 (2016)
5. Bienvenu, M., Ortiz, M.: Ontology-mediated query answering with data-tractable description logics. In: Proc. of Reasoning Web. pp. 218–307 (2015)
6. Botoeva, E., Konev, B., Lutz, C., Ryzhikov, V., Wolter, F., Zakharyaschev, M.: Inseparability and conservative extensions of description logic ontologies: A survey. In: Reasoning Web: Proceedings of 12th International Summer School. pp. 27–89 (2016)
7. Botoeva, E., Lutz, C., Ryzhikov, V., Wolter, F., Zakharyaschev, M.: Query-based entailment and inseparability for ALC ontologies. In: Proceedings of IJCAI. pp. 1001–1007 (2016)
8. Calì, A., Gottlob, G., Pieris, A.: Towards more expressive ontology languages: The query answering problem. Artif. Intell. 193, 87–128 (2012)
9. Calvanese, D., De Giacomo, G., Lembo, D., Lenzerini, M., Rosati, R.: Data complexity of query answering in description logics. Artificial Intelligence 195, 335–360 (2013)
10. Carral, D., Feier, C., Grau, B.C., Hitzler, P., Horrocks, I.: *EL*-ifying ontologies. In: IJCAR. pp. 464–479 (2014)
11. Chang, C.C., Keisler, H.J.: Model Theory, Studies in Logic and the Foundations of Mathematics, vol. 73. Elsevier (1990)
12. Eiter, T., Gottlob, G., Ortiz, M., Simkus, M.: Query answering in the description logic Horn-$\mathcal{SHIQ}$. In: JELIA. pp. 166–179 (2008)
13. Feier, C., Lutz, C., Kuusisto, A.: Rewritability in monadic disjunctive datalog, mmsnp, and expressive description logics. In: ICDT (2017)
14. Hernich, A., Lutz, C., Ozaki, A., Wolter, F.: Schema.org as a description logic. In: Proceedings of IJCAI. pp. 3048–3054 (2015)
15. Hernich, A., Lutz, C., Papacchini, F., Wolter, F.: Dichotomies on ontology-mediated querying with the guarded fragment. In: Proceedings of PODS (2017)
16. Hustadt, U., Motik, B., Sattler, U.: Reasoning in description logics by a reduction to disjunctive datalog. J. Autom. Reasoning 39(3), 351–384 (2007)
17. Kaminski, M., Grau, B.C.: Computing horn rewritings of description logics ontologies. In: Proceedings of IJCAI. pp. 3091–3097 (2015)

18. Kaminski, M., Nenov, Y., Grau, B.C.: Datalog rewritability of disjunctive datalog programs and non-horn ontologies. Artif. Intell. 236, 90–118 (2016)
19. Kazakov, Y.: Consequence-driven reasoning for Horn-$\mathcal{SHIQ}$ ontologies. In: Boutilier, C. (ed.) IJCAI. pp. 2040–2045 (2009)
20. Konev, B., Lutz, C., Wolter, F., Zakharyaschev, M.: Conservative rewritability of description logic tboxes. In: IJCAI. pp. 1153–1159 (2016)
21. Kontchakov, R., Lutz, C., Toman, D., Wolter, F., Zakharyaschev, M.: The combined approach to query answering in DL-Lite. In: Proc. of KR (2010)
22. Kontchakov, R., Zakharyaschev, M.: An introduction to description logics and query rewriting. In: Proc. of Reasoning Web. pp. 195–244 (2014)
23. Krisnadhi, A., Lutz, C.: Data complexity in the el family of dls. In: Proc. of DL (2007)
24. Krötzsch, M., Rudolph, S., Hitzler, P.: Complexity boundaries for Horn description logics. In: AAAI. pp. 452–457 (2007)
25. Lutz, C., Piro, R., Wolter, F.: Description logic tboxes: Model-theoretic characterizations and rewritability. In: IJCAI (2011)
26. Lutz, C., Wolter, F.: Deciding inseparability and conservative extensions in the description logic EL. J. Symb. Comput. 45(2), 194–228 (2010)
27. Lutz, C., Wolter, F.: Non-uniform data complexity of query answering in description logics. In: Proc. of KR. AAAI Press (2012)
28. Mugnier, M., Thomazo, M.: An introduction to ontology-based query answering with existential rules. In: Proc. of Reasoning Web. pp. 245–278 (2014)
29. Ortiz, M., Rudolph, S., Simkus, M.: Query answering in the Horn fragments of the description logics $\mathcal{SHOIQ}$ and $\mathcal{SROIQ}$. In: IJCAI. pp. 1039–1044 (2011)
30. Poggi, A., Lembo, D., Calvanese, D., Giacomo, G.D., Lenzerini, M., Rosati, R.: Linking data to ontologies. J. Data Semantics 10, 133–173 (2008)
31. Schaerf, A.: On the complexity of the instance checking problem in concept languages with existential quantification. J. of Intel. Inf. Systems 2, 265–278 (1993)

A Normal Form

We state the result to be proved again.

Lemma 2 (restated) *Every $\mathcal{ALCHIF}^{\sqsubseteq f}$ TBox $\mathcal{T}$ of depth 1 can be converted into a logically equivalent $\mathcal{ALCHIF}^{\sqsubseteq f}$ TBox $\mathcal{T}'$ in normal form. $\mathcal{T}'$ is of at most exponential size in $|\mathcal{T}|$.*

Proof. First, we transform $\mathcal{T}$ into a logically equivalent TBox $\mathcal{T}^*$ in which every CI is in normal form and which is of at most exponential size in $|\mathcal{T}|$. The CIs in $\mathcal{T}^*$ are not yet in normal form relative to the functionality assertions in $\mathcal{T}$. We then transform $\mathcal{T}^*$ into a logically equivalent TBox $\mathcal{T}'$ that is of at most polynomial size in $|\mathcal{T}^*|$ and in which all CIs are in normal form with respect to the functionality assertions in $\mathcal{T}$.

For any CIs $C \sqsubseteq D \in \mathcal{T}$:

- for any $\exists r.C$, compute the DNF of C and distribute the existential restrictions over the resulting disjunction; for any $\forall r.C$, compute the CNF of C and distribute the value restrictions over the resulting conjunction;
- convert the left-hand side to DNF, the right-hand side to CNF, remove disjunctions on the left and conjunctions on the right by splitting into multiple CIs; the left-hand side is now a conjunction on top-level, the right-hand side is a disjunction;
- move negative literals left to right and vice versa (negating them), move any $\exists r.C$ on the left-hand side with at least one negative literal in C to the right-hand side while computing the NNF of its negation, move any $\forall r.C$ on the right-hand side with no positive literal in C to the left-hand side while computing the NNF of its negation.

We refer to this first transformation into normal form as Ξ. That is, $\mathcal{T}^* = \Xi(\mathcal{T})$. As Ξ involves CNF and DNF translations, it follows that $\mathcal{T}^*$ is of size at most exponential with respect to $|\mathcal{T}|$. As no new symbol is introduced during the transformation, and each step is based on well-known logical equivalences, $\mathcal{T}^*$ is logically equivalent to $\mathcal{T}$.

The second step of our transformation is based on the following equivalences:

$$\exists r.(C_1 \sqcap C_2) \equiv \exists r.C_1 \sqcap \mathcal{Q}r.C_2, \quad \forall r.(C_1 \sqcup C_2) \equiv \forall r.C_1 \sqcup \mathcal{Q}r.C_2$$

where r is functional, $\mathcal{Q} \in \{\exists, \forall\}$, and C_1, C_2 are $\mathcal{ALCHIF}$ concepts.

The TBox $\mathcal{T}'$ in normal form with respect to $\mathcal{F}$ is obtained as follows. For any $C \sqsubseteq D$ in $\mathcal{T}^*$ and functional role r:

1. substitute any $\exists r.(A_1 \sqcap \ldots \sqcap A_n \sqcap \neg B_1 \sqcap \ldots \sqcap \neg B_m)$ with $\exists r.(A_1 \sqcap \ldots \sqcap A_n) \sqcap \bigsqcap_i \forall r.\neg B_i$, and perform the Ξ transformation again;
2. if there is $\exists r.D'$ or $\forall r.D'$ in D, then move any $\exists r.C'$ from left to right computing the NNF of its negation;
3. if there is a $\forall r.D'$ in D, then let $\mathcal{D} = \{D'' \mid \mathcal{Q}r.D'' \text{ in } D\}$ where $\mathcal{Q} \in \{\exists, \forall\}$, and replace all restrictions for r in D with $\forall r. \bigsqcup_{D'' \in \mathcal{D}} D''$;

4. if there is an $\exists r.C'$ in C, then let $\mathcal{C} = \{C'' \mid \exists r.C'' \text{ in } C\}$ and replace all restrictions for r in C with $\exists r. \bigsqcap_{C'' \in \mathcal{C}} C''$;
5. compute the Ξ transformation again.

As all transformation steps are equivalence-preserving and no new symbol is introduced, it follows that $\mathcal{T}'$ is equivalent to $\mathcal{T}^*$ (and to $\mathcal{T}$). It remains to show that $|\mathcal{T}'|$ is at most polynomial with respect to $|\mathcal{T}^*|$: even though Ξ transformation is needed, one does not obtain an exponential blow-up in size. This is because the application of Ξ in point 1 results in m new axioms per existential restriction where m is number of negative literals in its scope. Similarly for the application of Ξ in point 5, where the number of axioms increases based on the creation of the value restriction in point 3. The CNF transformation of these value restrictions depends on concept of the form $\exists r.(A_1 \sqcap \ldots \sqcap A_n)$ in D, and each such concept contributes n new axioms. $\square$