Finite Model Reasoning in Horn Description Logics

Yazmín Ibañez-García
KRDB Research Centre for Knowledge and Data,
Free University of Bozen-Bolzano, Italy
{ibanezgarcia@inf.unibz.it}

Carsten Lutz and **Thomas Schneider**
Fachbereich Informatik
Universität Bremen, Germany
{clu, tschneider}@informatik.uni-bremen.de

Abstract

We study finite model reasoning in expressive Horn description logics (DLs), starting with a reduction of finite ABox consistency to unrestricted ABox consistency. The reduction relies on reversing certain cycles in the TBox, an approach that originated in database theory, was later adapted to the inexpressive DL *DL-Lite*$_\mathcal{F}$, and is shown here to extend to the expressive Horn DL Horn-$\mathcal{ALCFI}$. The model construction used to establish correctness makes the structure of finite models more explicit than existing approaches to finite model reasoning in expressive DLs that rely on solving systems of inequations over the integers. Since the reduction incurs an exponential blow-up, we then develop a dedicated consequence-based algorithm for finite ABox consistency in Horn-$\mathcal{ALCFI}$ that implements the reduction on-the-fly rather than executing it up-front. The algorithm has optimal worst-case complexity and provides a promising foundation for implementations. We next show that our approach can be adapted to finite (positive existential) query answering relative to Horn-$\mathcal{ALCFI}$ TBoxes, proving that this problem is ExpTime-complete in combined complexity and PTime-complete in data complexity. For finite satisfiability and subsumption, we also show that our techniques extend to Horn-$\mathcal{SHIQ}$.

1 Introduction

Many popular expressive description logics (DLs) include both inverse roles and some form of counting such as functionality restrictions. This combination is well-known to result in a loss of the finite model property (FMP) and, consequently, reasoning w.r.t. the class of finite models (*finite model reasoning*) does not coincide with reasoning w.r.t. the class of all models (*unrestricted reasoning*). On the one hand, this distinction is gaining importance because DLs are increasingly used in database applications, where finiteness of models and databases is a central assumption. On the other hand, finite model reasoning is rarely used when DLs are applied in practice, mainly because for many DLs that lack the FMP, no algorithmic approaches to finite model reasoning are known that lend themselves towards efficient implementation.

Among the most widely-known DLs that include both inverse roles and counting are $\mathcal{ALCFI}$, $\mathcal{ALCQI}$, $\mathcal{SHIF}$, and $\mathcal{SHIQ}$, which are prominent fragments of the OWL2 DL ontology language. While finite model reasoning in these DLs is known to have the same complexity as unrestricted reasoning, namely ExpTime-complete (Lutz, Sattler, and Tendera 2005), the algorithmic approaches are rather different when only finite models are admitted. For unrestricted reasoning, there is a wide range of applicable algorithms such as tableau and resolution calculi, which often perform rather well in practical implementations. For finite model reasoning, all known approaches rely on the construction of some system of inequations (Calvanese 1996; Lutz, Sattler, and Tendera 2005) and then solve this system over the integers; the crux is that the system of inequations is of exponential size in the best case, and consequently it is far from obvious how to come up with efficient implementations. This is also true for the two-variable fragment of first-order logic with counting quantifiers (C2), into which the mentioned DLs can be embedded (Pacholski, Szwast, and Tendera 2000; Pratt-Hartmann 2005), that is, all known approaches to finite model reasoning in C2 rely on solving (at least) exponentially large systems of inequations.

Interestingly, the situation is quite different on the other end of the expressive power spectrum. While $\mathcal{SHIQ}$ et al. belong to the family of expressive DLs, *DL-Lite*$_\mathcal{F}$ is a comparably inexpressive DL that emerged from database applications, but also includes both inverse roles and functionality restrictions and thus lacks the FMP. Building on a technique that was developed in database theory by Cosmadakis, Kanellakis, and Vardi to decide the implication of inclusion dependencies and functional dependencies over finite databases (1990), Rosati has shown that finite model reasoning in *DL-Lite*$_\mathcal{F}$ can be reduced in polynomial time to unrestricted reasoning in *DL-Lite*$_\mathcal{F}$ (2008). The reduction is conceptually simple and relies on completing the TBox by finding certain cyclic inclusions and then 'reversing' them. For example, the cycle

$$\exists r^- \sqsubseteq \exists s \qquad \exists s^- \sqsubseteq \exists r \qquad (\mathsf{funct}\ r^-) \qquad (\mathsf{funct}\ s^-)$$

that consists of existential restrictions in the 'forward direction' and functionality statements in the 'backwards direction' would lead to the addition of the reversed cycle

$$\exists s \sqsubseteq \exists r^- \qquad \exists r \sqsubseteq \exists s^- \qquad (\mathsf{funct}\ r) \qquad (\mathsf{funct}\ s).$$

As a consequence, finite model reasoning in DL-Lite$_\mathcal{F}$ does not require new algorithmic techniques and can be implemented as efficiently as unrestricted reasoning. The reduction also makes explicit the *logical consequences of finite models*; in a sense, it can be viewed as an explicit axiomatization of finiteness.

Given that DL-Lite$_\mathcal{F}$ is only a very small fragment of $\mathcal{ALCFI}$ and $\mathcal{SHIQ}$, this situation raises the question whether the cycle reversion technique extends also to larger fragments of these DLs. In particular, *DL-Lite*$_\mathcal{F}$ is a 'Horn DL', and such logics are well-known to be algorithmically much more well-behaved than non-Horn DLs such as $\mathcal{ALCFI}$ (Baader, Brandt, and Lutz 2005; Calvanese et al. 2007). Maybe, then, this is the reason why cycle reversion works for *DL-Lite*$_\mathcal{F}$?

In this paper, we show that the cycle reversion technique of Cosmadakis et al. extends all the way to the expressive DLs Horn-$\mathcal{ALCFI}$ and Horn-$\mathcal{ALCQI}$. These logics, as well as their extensions Horn-$\mathcal{SHIF}$ and Horn-$\mathcal{SHIQ}$, are popular in ontology-based data access (Hustadt, Motik, and Sattler 2007; Ortiz, Rudolph, and Šimkus 2011; Eiter et al. 2012; Bienvenu, Lutz, and Wolter 2013) and properly extend *DL-Lite*$_\mathcal{F}$ and other relevant Horn DLs such as $\mathcal{ELIF}$ (Krisnadhi and Lutz 2007). We start with showing that finite ABox consistency in Horn-$\mathcal{ALCFI}$ can be reduced to unrestricted ABox consistency in Horn-$\mathcal{ALCFI}$ by cycle reversion; it follows that the same is true for finite satisfiability, finite subsumption, and finite instance checking. While the reduction technique is conceptually similar to that for *DL-Lite*$_\mathcal{F}$, the construction of a finite model in the correctness proof is more demanding. In comparison to approaches to finite model reasoning that rely on solving systems of inequations, though, they make the structure of finite models considerably more explicit.

Another crucial difference to the *DL-Lite*$_\mathcal{F}$ case is that, when completing Horn-$\mathcal{ALCFI}$ TBoxes, the cycles that have to be considered can be of exponential length, and thus the reduction is not polynomial. Consequently, when used in a naive way it can neither be expected to perform well in practice nor be used to (re)prove tight complexity bounds. To address these shortcomings, we develop a dedicated calculus for finite ABox consistency in Horn-$\mathcal{ALCFI}$ that implements the reduction on-the-fly rather than executing it up-front. The calculus is an extension of a consequence-based procedure for unrestricted satisfiability in Horn-$\mathcal{SHIQ}$ that was introduced by Kazakov in (2009) and implemented in the highly performant reasoner CB, first to classify the notoriously difficult Galen ontology. Many other state-of-the art reasoners for Horn-DLs are also based on consequence-based procedures, including ELK (Kazakov, Krötzsch, and Simančík 2011a) and CEL (Baader, Lutz, and Suntisrivaraporn 2006). Our algorithm shares the main feature of other consequence-based procedures to carefully avoid considering 'types' (conjunctions of concept names) that are irrelevant for deciding the problem at hand. We therefore believe that it provides a very promising basis for efficient implementations of finite model reasoning in Horn-$\mathcal{ALCFI}$. It also (re)proves the optimal upper ExpTime complexity bound for finite ABox consistency in this DL. Via a reduction, the cycle reversing reduction and the consequence-based algorithm can be applied also to finite satisfiability and subsumption in Horn-$\mathcal{ALCQI}$.

We then consider the paradigm of ontology-based data access (OBDA), extending our results from finite ABox consistency to answering positive existential queries (PEQs), relative to Horn-$\mathcal{ALCFI}$ TBoxes over finite models. In particular, we show that the reduction based on cycle reversion developed for ABox consistency also works in the case of PEQ answering. The construction of (counter)models in the correctness proofs, however, becomes yet more difficult and technical, and proceeds in two stages. First, we carefully modify the models constructed for finite ABox consistency so that there are no unintended matches of acyclic conjunctive queries (CQs). And second, we take a product with a finite group of high girth to eliminate unintended matches of cyclic CQs. Based on this result, we then prove that finite PEQ entailment (the Boolean version of PEQ answering) in Horn-$\mathcal{ALCFI}$ is ExpTime-complete regarding combined complexity and PTime-complete regarding data complexity. Previously, it was only known that finite CQ answering in (non-Horn) $\mathcal{ALCQI}$ is decidable and in coNP regarding data complexity (Pratt-Hartmann 2009).

Some proof details are deferred to the appendix in the long version: http://tinyurl.com/kr14fmr

2 Preliminaries

We introduce the DLs Horn-$\mathcal{ALCFI}$ and Horn-$\mathcal{ALCQI}$, as well as the reasoning tasks studied in this paper. The original definition of these DLs is based on a notion of polarity and somewhat unwieldy (Hustadt, Motik, and Sattler 2007); alternative and more direct definitions have been proposed later, see for example (Lutz and Wolter 2012). For brevity, we directly introduce Horn-$\mathcal{ALCQI}$ TBoxes in a normal form that is convenient for our purposes and disallows syntactic nesting of operators. It is a minor variation of the normal form proposed in (Kazakov 2009).

Let $\mathsf{N_C}$, $\mathsf{N_R}$, and $\mathsf{N_I}$ be countably infinite and disjoint sets of concept names, role names, and individual names. A *role* is either a role name r or an *inverse role* r^-. A *Horn-$\mathcal{ALCQI}$ TBox* $\mathcal{T}$ is a set of *concept inclusions (CIs)* that can take the following forms:

$$K \sqsubseteq A \qquad K \sqsubseteq \bot \qquad K \sqsubseteq \exists r.K'$$
$$K \sqsubseteq \forall r.K' \qquad K \sqsubseteq (\leqslant 1\; r\; K') \qquad K \sqsubseteq (\geqslant n\; r\; K')$$

where K and K' denote a (possibly empty) conjunction of concept names, A a concept name, r a (potentially inverse) role, and $n \geq 2$. Throughout the paper, we will deliberately confuse conjunctions of concept names and sets of concept names. The empty conjunction is abbreviated by $\top$. As usual, we allow to easily switch between role names and their inverse by identifying $(r^-)^-$ and r. A *Horn-$\mathcal{ALCFI}$ TBox* is a Horn-$\mathcal{ALCQI}$ TBox that does not include CIs of the form $K \sqsubseteq (\geqslant n\; r\; K')$.

The semantics of Horn-$\mathcal{ALCQI}$ is based on interpretations as usual, see (Baader et al. 2003) for details. We write $\mathcal{T} \models C \sqsubseteq D$ if the concept inclusion $C \sqsubseteq D$ is satisfied in all models of the TBox $\mathcal{T}$, and $\mathcal{T} \models_{\mathsf{fin}} C \sqsubseteq D$ if the same holds for all finite models. A concept name A is *(finitely) satisfiable* w.r.t. a TBox $\mathcal{T}$ if $\mathcal{T}$ has a (finite) model $\mathcal{I}$ with $A^\mathcal{I} \neq \emptyset$. If $\mathcal{T} \models A \sqsubseteq B$ (resp. $\mathcal{T} \models_{\mathsf{fin}} A \sqsubseteq B$) with A and B concept names, then we say that B is *(finitely) subsumed* by A.

An *ABox* is a finite set of *concept assertions* $A(a)$ and *role assertions* $r(a, b)$ where A is a concept name, r a role name, and a, b are individual names. For simplicity, we make the *standard names assumption*, that is, every interpretation $\mathcal{I}$ interpretes all individuals as themselves; for example $\mathcal{I}$ satisfies $A(a)$ if $a \in A^{\mathcal{I}}$. The standard names assumption implies the unique name assumption (UNA). The results in this paper, however, do not depend on any of these assumptions. Throughout the paper, we sometimes write $r^-(a,b) \in \mathcal{A}$ for $r(b,a) \in \mathcal{A}$ and use $\mathsf{Ind}(\mathcal{A})$ to denote the set of all individual names that occur in $\mathcal{A}$.

We write $\mathcal{A}, \mathcal{T} \models A(a)$ if the ABox assertion $A(a)$ is satisfied in all common models of the ABox $\mathcal{A}$ and the TBox $\mathcal{T}$, and $\mathcal{A}, \mathcal{T} \models_{\mathsf{fin}} A(a)$ if the same holds for all finite models. We then say that a is a *(finite) instance* of A in $\mathcal{A}$ w.r.t. $\mathcal{T}$. An ABox $\mathcal{A}$ is *(finitely) consistent* w.r.t. $\mathcal{T}$ if there is a (finite) model $\mathcal{I}$ of $\mathcal{T}$ that satisfies all assertions in $\mathcal{A}$.

The above notions give rise to four decision problems studied in this paper, which are *finite satisfiability* (of a concept name w.r.t. a TBox), *finite subsumption* (between two concept names w.r.t. a TBox), *finite ABox consistency* (w.r.t. a TBox) and *finite instance checking* (of an ABox individual and a concept name, w.r.t. an ABox and a TBox). There are easy polynomial time reductions from satisfiability to subsumption to instance checking to ABox consistency, which work both in the finite and in the unrestricted case.

The following examples show that, in Horn-$\mathcal{ALCFI}$, finite and unrestricted reasoning do not coincide.

Example 1

$$\mathcal{T} = \{ \quad A \sqsubseteq \exists r.B, \qquad B \sqsubseteq \exists r.B, \quad B \sqsubseteq (\leqslant 1\, r^- \top), \quad A \sqcap B \sqsubseteq \bot \quad \}$$

A is satisfiable w.r.t. $\mathcal{T}$, but not finitely satisfiable. In fact, when $d \in A^{\mathcal{I}}$ in some model $\mathcal{I}$ of $\mathcal{T}$, then there must be an infinite chain $r(d,d_1), r(d_1,d_2), \ldots$ with $d \in A^{\mathcal{I}}$, and $d_2, d_3, \cdots \in B^{\mathcal{I}}$. Since d cannot be in $B^{\mathcal{I}}$ and r is inverse functional, no two elements on the chain can be identified.

$$\mathcal{T}' = \{ A_1 \sqsubseteq \exists r.A_2, \quad A_2 \sqsubseteq \exists r.(A_1 \sqcap B), \quad \top \sqsubseteq (\leqslant 1\, r^- \top) \quad \}$$

The reader might want to verify that $\mathcal{T}' \not\models A_1 \sqsubseteq B$, but $\mathcal{T}' \models_{\mathsf{fin}} A_1 \sqsubseteq B$.

It follows form the observations in (Kazakov 2009) that, for the purposes of deciding satisfiability of concepts in unrestricted models, the normal form for TBoxes introduced above can be assumed without loss of generality because every Horn-$\mathcal{ALCQI}$ TBox $\mathcal{T}$ can be converted in polynomial time into a TBox $\mathcal{T}'$ in the above form such that every model of $\mathcal{T}'$ is a model of $\mathcal{T}$ and, conversely, every model of $\mathcal{T}$ can be converted into a model of $\mathcal{T}'$ by interpreting the concept names that were introduced during normalization. It follows that normal form can be assumed w.l.o.g. both for unrestricted reasoning and for finite model reasoning, and for all reasoning problems considered in this paper.

3 From Finite Models to Unrestricted Models

We show that finite ABox consistency in Horn-$\mathcal{ALCFI}$ can be reduced to unrestricted ABox consistency by reversing certain cycles in the TBox. The reduction exhibited in this section provides a novel decision procedure for finite ABox consistency in Horn-$\mathcal{ALCFI}$ and Horn-$\mathcal{ALCQI}$ (as well as for finite satisfiability, finite subsumption, and finite instance checking) and is the basis for developing a consequence-based procedure in Section 4. It also highlights the logical consequences of finite models in Horn-$\mathcal{ALCFI}$. The material in this section is an extended and improved version of the workshop paper (Ibáñez-García, Lutz, and Schneider 2013).

Reversing Cycles

Let $\mathcal{T}$ be a Horn-$\mathcal{ALCFI}$ TBox. A *finmod cycle in $\mathcal{T}$* is a sequence $K_1, r_1, K_2, r_2, \ldots, r_{n-1}, K_n$, with $K_1, \ldots, K_n$ conjunctions of concept names and $r_1, \ldots, r_{n-1}$ (potentially inverse) roles such that $K_n = K_1$ and, for $1 \leq i < n$:

$$\mathcal{T} \models K_i \sqsubseteq \exists r_i.K_{i+1} \text{ and } \mathcal{T} \models K_{i+1} \sqsubseteq (\leqslant 1\, r_i^-\ K_i).$$

By *reversing* a finmod cycle $K_1, r_1, K_2, r_2, \ldots, r_{n-1}, K_n$ in a TBox $\mathcal{T}$, we mean to extend $\mathcal{T}$ with the following concept inclusions, for $1 \leq i < n$:

$$K_{i+1} \sqsubseteq \exists r_i^-.K_i \text{ and } K_i \sqsubseteq (\leqslant 1\, r_i\ K_{i+1}).$$

The *completion* $\mathcal{T}_{\mathsf{f}}$ of a TBox $\mathcal{T}$ is obtained from $\mathcal{T}$ by exhaustively reversing finmod cycles. Note that, although there may be infinitely many finmod cycles, only finitely many CIs can be added by cycle reversion (exponentially many in the size of the original TBox, in the worst case). For finding these finitely many CIs, it clearly suffices to consider finmod cycles in which all triples (r_i, K_{i+1}, r_{i+1}) are distinct. Also note that finding finmod cycles requires deciding unrestricted subsumption, which is decidable and ExpTime-complete.

Example 2 *The TBox $\mathcal{T}'$ from Example 1 entails (in unrestricted models)*

$$A_1 \sqcap B \sqsubseteq \exists r.A_2, \qquad A_2 \sqsubseteq \exists r.(A_1 \sqcap B),$$
$$A_2 \sqsubseteq (\leqslant 1\, r^-\ A_1 \sqcap B), \quad A_2 \sqcap B \sqsubseteq (\leqslant 1\, r^-\ A_1).$$

Thus, A_1, r, A_2, r, A_1, is a finmod cycle in $\mathcal{T}'$, which is reversed to

$$A_2 \sqsubseteq \exists r^-.A_1, \qquad A_1 \sqsubseteq \exists r^-.A_2,$$
$$A_2 \sqsubseteq (\leqslant 1\, r\ A_1), \qquad A_1 \sqsubseteq (\leqslant 1\, r\ A_2).$$

From $A_1 \sqsubseteq \exists r^-.A_2$, $A_2 \sqsubseteq \exists r.(A_1 \sqcap B)$, and $A_2 \sqsubseteq (\leqslant 1\ r\ A_1)$, we obtain $\mathcal{T}'_{\mathsf{f}} \models A_1 \sqsubseteq B$, in correspondence with $\mathcal{T}' \models_{\mathsf{fin}} A_1 \sqsubseteq B$. Note that $\mathcal{T}'$ also contains another finmod cycle, which is $(A_1 \sqcap B), r, A_2, r, (A_1 \sqcap B)$.

The following result shows that TBox completion provides a reduction from finite ABox consistency to unrestricted ABox consistency.

Theorem 3 *Let $\mathcal{T}$ be a Horn-$\mathcal{ALCFI}$ TBox and $\mathcal{A}$ an ABox. Then $\mathcal{A}$ is finitely consistent w.r.t. $\mathcal{T}$ iff $\mathcal{A}$ is consistent w.r.t. the completion $\mathcal{T}_{\mathsf{f}}$ of $\mathcal{T}$.*

The "only if" direction of Theorem 3 is an immediate consequence of the observation that all CIs added by cycle reversion are entailed by the original TBox in finite models.

Lemma 4 *Let $K_1, r_1, \ldots, r_{n-1}, K_n$ be a finmod cycle in $\mathcal{T}$. Then $\mathcal{T} \models_{\mathsf{fin}} K_{i+1} \sqsubseteq \exists r_i^-.K_i$ and $\mathcal{T} \models_{\mathsf{fin}} K_i \sqsubseteq (\leqslant 1\ r_i\ K_{i+1})$ for $1 \leq i < n$,.*

Proof. We have to show that if $K_1, r_1, \dots, r_{n-1}, K_n$ is a finmod cycle in $\mathcal{T}$ and $\mathcal{I}$ is a finite model of $\mathcal{T}$, then $K_i^\mathcal{I} \subseteq (\leqslant 1\ r_i\ K_{i+1})^\mathcal{I}$ and $K_{i+1}^\mathcal{I} \subseteq (\exists r_i^-.K_i)^\mathcal{I}$ for $1 \leq i < n$. We first note that, by the semantics of Horn-$\mathcal{ALCFI}$, we must have $|K_1^\mathcal{I}| \leq \dots \leq |K_n^\mathcal{I}|$, thus $K_n = K_1$ yields $|K_1^\mathcal{I}| = \dots = |K_n^\mathcal{I}|$. Fix some i with $1 \leq i < n$. Using $|K_i^\mathcal{I}| = |K_{i+1}^\mathcal{I}|$, $K_i^\mathcal{I} \subseteq (\exists r_i.K_{i+1})^\mathcal{I}$, and $K_{i+1}^\mathcal{I} \subseteq (\leqslant 1\ r_i^-\ K_i)^\mathcal{I}$, it is easy to verify that $K_i^\mathcal{I} \subseteq (\leqslant 1\ r_i\ K_{i+1})^\mathcal{I}$ and $K_{i+1}^\mathcal{I} \subseteq (\exists r_i^-.K_i)^\mathcal{I}$, as required. ❏

We now prove the "if" direction of Theorem 3, which is much more demanding as it requires to explicitly construct finite models.

Constructing Finite Models

Assume that $\mathcal{A}$ is consistent w.r.t. $\mathcal{T}_\mathsf{f}$. Our aim is to construct a finite model $\mathcal{I}$ of $\mathcal{A}$ and $\mathcal{T}_\mathsf{f}$ (and thus also of $\mathcal{T}$). Before we give details of the construction, we introduce some relevant preliminaries.

Let $\mathsf{CN}(\mathcal{T})$ denote the set of concept names used in $\mathcal{T}$ (or, equivalently, in $\mathcal{T}_\mathsf{f}$). A *type for* $\mathcal{T}_\mathsf{f}$ is a subset $t \subseteq \mathsf{CN}(\mathcal{T})$ such that there is a (potentially infinite) model $\mathcal{I}$ of $\mathcal{T}_\mathsf{f}$ and a $d \in \Delta^\mathcal{I}$ such that $\mathsf{tp}_\mathcal{I}(d) = t$, where

$$\mathsf{tp}_\mathcal{I}(d) := \{A \in \mathsf{CN}(\mathcal{T}) \mid d \in A^\mathcal{I}\}$$

is the type *realized* at d in $\mathcal{I}$. We use $\mathsf{TP}(\mathcal{T}_\mathsf{f})$ to denote the set of all types for $\mathcal{T}_\mathsf{f}$. For $t, t' \in \mathsf{TP}(\mathcal{T}_\mathsf{f})$ and r a role, we write

- $t \rightarrow_r t'$ if $\mathcal{T}_\mathsf{f} \models t \sqsubseteq \exists r.t'$ and t' is maximal with this property;
- $t \rightarrow_r^1 t'$ if $t \rightarrow_r t'$ and $\mathcal{T}_\mathsf{f} \models t' \sqsubseteq (\leqslant 1\ r^-\ t)$;
- $t\ {}^1\!\!\leftrightarrow_r^1 t'$ if $t \rightarrow_r^1 t'$ and $t' \rightarrow_{r^-}^1 t$.

Note that when

$$t_1 \rightarrow_{r_1}^1 t_2 \rightarrow_{r_2}^1 \cdots \rightarrow_{r_{n-1}}^1 t_n = t_1 \qquad (*)$$

then $t_1, r_1, \dots, r_{n-1}, t_n$ is a finmod cycle in $\mathcal{T}_\mathsf{f}$ and the fact that it has been reversed means that all '$\rightarrow^1$' in $(*)$ can be replaced with ${}^1\!\!\leftrightarrow^1$. Types related by ${}^1\!\!\leftrightarrow_r^1$ are connected very tightly by the TBox $\mathcal{T}_\mathsf{f}$ and are best considered together when building finite models. This is formalized by the notion of a *type class*, which is a non-empty set $P \subseteq \mathsf{TP}(\mathcal{T}_\mathsf{f})$ such that $t \in P$ and $t\ {}^1\!\!\leftrightarrow_r^1 t'$ implies $t' \in P$, and P is minimal with this condition. Note that the set of all type classes is a partition of $\mathsf{TP}(\mathcal{T}_\mathsf{f})$. We set $P \prec P'$ if there are $t \in P$ and $t' \in P'$ with $t' \subsetneq t$. Let $\prec^+$ be the transitive closure of $\prec$. A proof of the following observation can be found in the appendix.

Lemma 5 *$\prec^+$ is a strict partial order.*

We construct the desired finite model $\mathcal{I}$ of $\mathcal{A}$ and $\mathcal{T}_\mathsf{f}$ by starting with an initial interpretation that essentially consists of the ABox $\mathcal{A}$ and then exhaustively applying three *completion rules* denoted with $(\mathbf{c1})$ to $(\mathbf{c3})$, where $(\mathbf{c1})$ is given preference over $(\mathbf{c2})$. Completion repeatedly introduces elements whose existence is required by CIs $K \sqsubseteq \exists r.C$, carefully distinguishing several cases to ensure that no functionality restrictions are violated. We will prove that rule application terminates after finitely many steps, producing a finite model.

During the construction of $\mathcal{I}$, we will make sure that the following invariants are satisfied:

(i1) $\mathsf{tp}_\mathcal{I}(d) \in \mathsf{TP}(\mathcal{T}_\mathsf{f})$ for all $d \in \Delta^\mathcal{I}$;

(i2) if $(d, d') \in r^\mathcal{I} \setminus (\mathsf{Ind}(\mathcal{A}) \times \mathsf{Ind}(\mathcal{A}))$, then we have $\mathsf{tp}_\mathcal{I}(d) \rightarrow_r \mathsf{tp}_\mathcal{I}(d')$ or $\mathsf{tp}_\mathcal{I}(d') \rightarrow_{r^-} \mathsf{tp}_\mathcal{I}(d)$;

(i3) if $\mathcal{T}_\mathsf{f} \models K \sqsubseteq (\leqslant 1\ r\ K')$, then $\mathcal{I} \models K \sqsubseteq (\leqslant 1\ r\ K')$.

The initial version of $\mathcal{I}$ is defined by introducing an element for every ABox individual, and an element d_t for each $t \in \mathsf{TP}(\mathcal{T}_\mathsf{f})$. In detail, we set

$$\begin{aligned} \Delta^\mathcal{I} &= \mathsf{Ind}(\mathcal{A}) \cup \{\ d_t \mid t \in \mathsf{TP}(\mathcal{T})_\mathsf{f}\ \} \\ A^\mathcal{I} &= \{\ a \in \mathsf{Ind}(\mathcal{A}) \mid A \in \mathsf{tp}_\mathcal{A}(a)\ \} \cup \{d_t \mid A \in t\ \} \\ r^\mathcal{I} &= \{\ (a, b) \mid r(a, b) \in \mathcal{A}\ \} \end{aligned}$$

where

$$\mathsf{tp}_\mathcal{A}(a) := \{A \in \mathsf{CN}(\mathcal{T}) \mid \mathcal{A}, \mathcal{T}_\mathsf{f} \models A(a)\}.$$

The completion rules are described in detail below.

(c1) Choose a $d \in \Delta^\mathcal{I}$ such that $\mathsf{tp}_\mathcal{I}(d) \rightarrow_r^1 t$, $t \not\rightarrow_{r^-}^1 \mathsf{tp}_\mathcal{I}(d)$, and $d \notin (\exists r.t)^\mathcal{I}$. Add a fresh domain element e, and modify the extension of concept and role names such that $\mathsf{tp}_\mathcal{I}(e) = t$ and $(d, e) \in r^\mathcal{I}$.

(c2) Choose a type class P that is minimal w.r.t. the order $\prec^+$, a $\lambda = s\ {}^1\!\!\leftrightarrow_r^1 s'$ with $s \in P$, and an element $d \in s^\mathcal{I} \setminus (\exists r.s')^\mathcal{I}$.

For each $\lambda = s\ {}^1\!\!\leftrightarrow_r^1 s'$ with $s \in P$, set

$$X_{\lambda,1}^\mathcal{I} = s^\mathcal{I} \setminus (\exists r.s')^\mathcal{I} \qquad X_{\lambda,2}^\mathcal{I} = s'^\mathcal{I} \setminus (\exists r^-.s)^\mathcal{I}.$$

Take (i) a fresh set Δ_s for each $s \in P$ such that $|\biguplus_{s \in P} \Delta_s| \leq 2^{|\mathcal{T}|} \cdot |\Delta^\mathcal{I}|$ and (ii) a bijection π_λ between $X_{\lambda,1}^\mathcal{I} \cup \Delta_s$ and $X_{\lambda,2}^\mathcal{I} \cup \Delta_{s'}$ for each $\lambda = s\ {}^1\!\!\leftrightarrow_r^1 s'$ with $s, s' \in P$ and r a role name (the concrete construction is detailed below). Now extend $\mathcal{I}$ as follows:

– add all domain elements in $\biguplus_{s \in P} \Delta_s$;

– extend $r^\mathcal{I}$ with π_λ, for each $\lambda = s\ {}^1\!\!\leftrightarrow_r^1 s'$ with $s, s' \in P$ and r a role name;

– interpret concept names so that $\mathsf{tp}_\mathcal{I}(d) = s$ for all $d \in \Delta_s$, $s \in P$.

(c3) Choose a $d \in \Delta^\mathcal{I}$ such that $\mathsf{tp}_\mathcal{I}(d) \rightarrow_r t$, $\mathsf{tp}_\mathcal{I}(d) \not\rightarrow_r^1 t$, and $d \notin (\exists r.t)^\mathcal{I}$. Add the edge (d, d_t) to $r^\mathcal{I}$, where d_t is the element introduced for type t in the initial version of $\mathcal{I}$.

To complete the description of the rules, we have to show that, in $(\mathbf{c2})$, the sets Δ_s and bijections π_λ indeed exist. Let $n_{\max} = \max\{|s^\mathcal{I}| \mid s \in P\}$. For each $s \in P$, set $\Delta_s := \{d_{s,i} \mid |s^\mathcal{I}| < i \leq n_{\max}\}$ and define the set of *s-instances* $I_s := s^\mathcal{I} \cup \Delta_s$. For each $\lambda = s\ {}^1\!\!\leftrightarrow_r^1 s'$ with $s, s' \in P$, define

$$R_\lambda := \{(d, e) \in r^\mathcal{I} \mid d \in s^\mathcal{I} \text{ and } e \in s'^\mathcal{I}\}.$$

We first note that it is a consequence of invariant $(\mathbf{i3})$ that

$(*)$ the relation R_λ is functional and inverse functional.

In fact, $(d, e_1), (d, e_2) \in R_\lambda$ implies $(d, e_1), (d, e_2) \in r^\mathcal{I}$, $d \in s^\mathcal{I}$, and $e_1, e_2 \in s'^\mathcal{I}$. By λ, $\mathcal{T}_\mathsf{f} \models s \sqsubseteq (\leqslant 1\ r\ s')$. Thus, (**i3**) yields $e_1 = e_2$. Inverse functionality can be shown analogously.

Let R^1_λ be the domain of R_λ, and let R^2_λ be its range. By $(*)$, we have $|R^1_\lambda| = |R^2_\lambda|$. By definition of the sets Δ_s, we have $|I_s| = |I_{s'}|$. Moreover, $R^1_\lambda \subseteq I_s$ and $R^2_\lambda \subseteq I_{s'}$. We can thus choose a bijection π_λ between $I_s \setminus R^1_\lambda$ and $I_{s'} \setminus R^2_\lambda$, which is as required since $I_s \setminus R^1_\lambda = X^\mathcal{I}_{\lambda,1} \cup \Delta_s$ and $I_{s'} \setminus R^2_\lambda = X^\mathcal{I}_{\lambda,2} \cup \Delta_{s'}$. The construction of the sets Δ_s clearly ensures that their union has the required cardinality.

The following theorem summarizes the statements that remain to be proved in order to show that the construction of $\mathcal{I}$ is well-defined and yields a finite model of $\mathcal{A}$ and $\mathcal{T}_\mathsf{f}$.

Theorem 6

1. *Applying* (**c1**) *to* (**c3**) *preserves invariants* (**i1**) *to* (**i3**)*;*
2. *Application of* (**c1**) *to* (**c3**) *terminates;*
3. *$\mathcal{I}$ is a model of $\mathcal{A}$ and $\mathcal{T}_\mathsf{f}$.*

Proof. We refer to the appendix for full proofs and only sketch the central idea in the proof of Point 2 here, going back to (Cosmadakis, Kanellakis, and Vardi 1990). The main issue in the termination proof is to show that no infinite role chain $r_0(d_0, d_1), r_1(d_1, d_2), \ldots$ is generated in which all the elements d_i are pairwise distinct. Since every application of a completion rule generates only finitely many elements, any such chain must be generated by infinitely many rule applications. As there are only finitely many types, we must find elements d_i and d_j with $\mathsf{tp}_\mathcal{I}(d_i) = \mathsf{tp}_\mathcal{I}(d_j)$ and such that d_i and d_j were generated by different rule applications. It can be shown that, w.l.o.g., we can assume that the elements on the chain are ordered so that if $j > i$, then d_j was not generated by an earlier rule application than d_i. Analysing the completion rules, it is easy to see that this implies $\mathsf{tp}_\mathcal{I}(d_i) \rightarrow^1_{r_i} \mathsf{tp}_\mathcal{I}(d_{i+1}) \rightarrow^1_{r_{i+1}} \cdots \rightarrow^1_{r_{j-1}} \mathsf{tp}_\mathcal{I}(d_j)$. Since $\mathsf{tp}_\mathcal{I}(d_i) = \mathsf{tp}_\mathcal{I}(d_j)$, this is a finmod cycle, which has been reversed when constructing $\mathcal{T}_\mathsf{f}$, and thus all arrows $\rightarrow^1_{r_{i+\ell}}$ can be replaced with ${}^1\!\leftrightarrow^1_{r_{i+\ell}}$. By definition of the completion rules, this means that all of $d_i, \ldots, d_j$ were introduced in the same application of (**c2**), which is a contradiction to d_i and d_j being generated by different rule applications. ❑

4 Consequence-Driven Procedure

While completing TBoxes with reversed cycles yields a reduction of finite model reasoning to infinite model reasoning, it blows up the TBox exponentially and is thus not suited for direct implementation. In this section, we build on the results from the previous section to devise a calculus for ABox consistency in Horn-$\mathcal{ALCFI}$ that does not require TBox completion to be carried out up-front, but instead reverses cycles 'on the fly'; moreover, the calculus implicitly groups together cycles that are closely related, potentially reversing a very large number of cycles in only a few steps (see Example 7 below). Our calculus belongs to a family of algorithms that are known as consequence-driven procedures and underly modern and highly efficient reasoners for Horn DLs such as

R1 $$\frac{}{K \sqcap A \sqsubseteq A}$$

R2 $$\frac{}{K \sqsubseteq \top}$$

R3 $$\frac{K \sqsubseteq A_i \quad \bigsqcap A_i \sqsubseteq C}{K \sqsubseteq C}$$

R4 $$\frac{K \sqsubseteq \exists r.K' \quad K' \sqsubseteq \forall r^-.A}{K \sqsubseteq A}$$

R5 $$\frac{K \sqsubseteq \exists r.K' \quad K \sqsubseteq \forall r.A}{K \sqsubseteq \exists r.(K' \sqcap A)}$$

R6 $$\frac{K \sqsubseteq \exists r.K' \quad K' \sqsubseteq \bot}{K \sqsubseteq \bot}$$

R7 $$\frac{\begin{array}{lll} K \sqsubseteq \exists r.K_1 & K \sqsubseteq \exists r.K_2 & K_1 \sqsubseteq A \\ K \sqsubseteq (\leqslant 1\ r\ A) & & K_2 \sqsubseteq A \end{array}}{K \sqsubseteq \exists r.(K_1 \sqcap K_2)}$$

R8 $$\frac{\begin{array}{lll} K \sqsubseteq \exists r.K' & K' \sqsubseteq \exists r^-.K_1 & K \sqsubseteq A \\ K' \sqsubseteq (\leqslant 1\ r^-\ A) & & K_1 \sqsubseteq A \end{array}}{K \sqsubseteq A_1 \quad \text{for any } A_1 \in K_1}$$

R9 $$\frac{\begin{array}{lll} K_i \sqsubseteq \exists r_i.K_{i \oplus_n 1} & & i < n \\ K_{i \oplus_n 1} \sqsubseteq (\leqslant 1\ r_i^-\ A_i) & K_i \sqsubseteq A_i & \end{array}}{K_1 \sqsubseteq \exists r_0^-.K_0 \qquad K_0 \sqsubseteq (\leqslant 1\ r_0\ A_1)}$$

Figure 1: Inference Rules

CEL, CB, and ELK (Baader, Lutz, and Suntisrivaraporn 2006; Kazakov 2009; Kazakov, Krötzsch, and Simančík 2011b). It thus establishes a promising foundation for actual implementations of finite-model reasoning in Horn-$\mathcal{ALCFI}$ and, via the reduction in Section 6, in Horn-$\mathcal{ALCQI}$. For simplicity, we start with a calculus for finite satisfiability and finite subsumption. An expansion to finite ABox consistency (and thus to finite instance checking) is sketched afterwards.

The calculus starts with a given TBox $\mathcal{T}$ and then exhaustively applies a set of inference rules. To ease their presentation, we assume that $\mathcal{T}$ is in a normal form that is slightly stricter than the one introduced in Section 2: in CIs $K \sqsubseteq \forall r.K'$ and $K \sqsubseteq (\leqslant 1\ r\ K)'$, K' must be a concept name A. The inference rules are displayed in Figure 1. They preserve the normal form and are applied in the sense that, if the concept inclusions in the precondition (above the line) are already present, then those in the postcondition (below the line) are added. Recall that K stands for a conjunction of concept names, which we read here modulo commutativity. Rule **R1** is applied only if $K \sqcap A$ *occurs* in the current (partially completed) TBox, that is, there is a CI of the form $K \sqcap A \sqsubseteq C$ or $K' \sqsubseteq \exists r.(K \sqcap A)$. The same is true for rule **R2** with K in place of $K \sqcap A$. In rule **R9**, $\oplus_n$ means addition modulo n.

We point out that rules **R1** to **R8** are minor variations of the corresponding rules in the calculus presented by Kazakov (2009), the main difference being that our language does not include role hierarchies. Rule **R9** is novel and deals with reversing cycles on the fly. Note that only the 'first edge' of each cycle is reversed, and that this is sufficient because the cycle can be rotated to make any edge the 'first' one.

Example 7 *Consider the TBox*

$$\mathcal{T} = \{A \sqsubseteq \exists r.(A \sqcap A_1 \sqcap \cdots \sqcap A_n), \quad (1)$$
$$A \sqsubseteq (\leqslant 1\, r^-\, A)\,\}. \quad (2)$$

Cycle reversion from Section 3 reverses all of the exponentially many cycles K, r, K with $K \subseteq S := \{A, A_1, \ldots, A_n\}$ and $A \in K$, adding $K \sqsubseteq \exists r^-.K$ and $K \sqsubseteq (\leqslant 1\ r\ K)$ for all such K. In contrast, the calculus avoids introducing 'irrelevant' conjunctions K and instead jointly reverses all these cycles by generating $A \sqsubseteq \exists r^-.S$ and $A \sqsubseteq (\leqslant 1\ r\ A)$:

$S \sqsubseteq A$	*from* **R1**		(3)
$A \sqsubseteq A$	*from* **R1**		(4)
$S \sqsubseteq \exists r.S$	*from* (1), (3), **R3**		(5)
$S \sqsubseteq (\leqslant 1\, r^-\, A)$	*from* (2), (3), **R3**		(6)
$S \sqsubseteq \exists r^-.S$	*and*		(7)
$S \sqsubseteq (\leqslant 1\ r\ A)$	*from* (3), (5), (6), **R9**		(8)
$A \sqsubseteq A_i$	*from* (1), (3), (4), (6), (7), **R8**		(9)
$A \sqsubseteq \exists r^-.S$	*from* (7), (9), **R3**		(10)
$A \sqsubseteq (\leqslant 1\ r\ A)$	*from* (8), (9), **R3**		(11)

Note that avoiding to introduce 'irrelevant' conjunctions K as illustrated by Example 7 is a main feature of consequence-based procedures which enables the excellent practical performance typically observed for this class of calculi.

The algorithm terminates after at most exponentially many rule applications since there are only exponentially many different concept inclusions that use the concept and role names of the original TBox. Each rule application can be performed in polynomial time, which is easy to see for the rules **R1**–**R8**. For **R9**, the crucial observation is that it suffices to consider all conjunctions K_0, K_1 and to check whether they are involved in *any* cycle. The latter can easily be done by a variation of directed graph reachability, where the nodes of the graph are the conjunctions that occur in the current TBox and the edges come from inclusions $K \sqsubseteq \exists r.K'$.

The following theorem, which is the main result of this section, states that the calculus is sound and complete.

Theorem 8 *Let $\mathcal{T}$ be a Horn-$\mathcal{ALCFI}$ TBox, $\widehat{\mathcal{T}}$ be obtained by exhaustively applying Rules* **R1**–**R9**, *and let A_0 be a concept name. Then A_0 is finitely satisfiable w.r.t. $\mathcal{T}$ iff $A_0 \sqsubseteq \bot \notin \widehat{\mathcal{T}}$.*

While Theorem 8 is formulated only for finite satisfiability, the algorithm can of course also be used to decide finite subsumption via the usual reduction to finite satisfiability. The following continues Example 7.

Example 9 *Let $\mathcal{T}$ be the TBox from Example 7 and*

$$\mathcal{T}' = \mathcal{T} \cup \{\quad A \sqsubseteq \exists r.(A \sqcap X_1), \quad (12)$$
$$A \sqsubseteq \exists r.(A \sqcap X_2), \quad (13)$$
$$X_1 \sqcap X_2 \sqsubseteq \bot \quad \} \quad (14)$$

The calculus derives $A \sqsubseteq \bot$, thus A is finitely unsatisfiable w.r.t. $\mathcal{T}'$:[1]

$A \sqcap X_i \sqsubseteq A$	*from* **R1**	(15)
$A \sqsubseteq \exists r.(A \sqcap X_1 \sqcap X_2)$	*from* (11)–(13), (15), **R7**	(16)
$A \sqsubseteq \bot$	*from* (14), (16), **R6**	(17)

[1] A is obviously satisfiable w.r.t. $\mathcal{T}'$ in unrestricted models.

We now prove Theorem 8. The "only if" direction (soundness) is straightforward by verifying that each rule is sound in finite models. In contrast, the "if" direction (completeness) turns out to be surprisingly subtle to establish. The proof strategy is as follows. Assume that $A_0 \sqsubseteq \bot \notin \widehat{\mathcal{T}}$. We construct a (possibly infinite) model $\widehat{\mathcal{I}}$ of $\widehat{\mathcal{T}}$ with $A_0^{\widehat{\mathcal{I}}} \neq \emptyset$ and show that $\widehat{\mathcal{I}}$ is actually a model of $\mathcal{T}_\mathsf{f}$. By Theorem 3, it follows that A_0 is finitely satisfiable w.r.t. $\mathcal{T}$. From now on, assume w.l.o.g. that A_0 actually occurs in $\mathcal{T}$.

To construct $\widehat{\mathcal{I}}$, let $\mathsf{KON}(\widehat{\mathcal{T}})$ denote the set of all conjunctions K such that K occurs in $\widehat{\mathcal{T}}$ (in the sense explained above) and $K \sqsubseteq \bot \notin \widehat{\mathcal{T}}$. The domain $\Delta^{\widehat{\mathcal{I}}}$ consists of finite words $d = K_1 K_2 \cdots K_n \in \mathsf{KON}(\widehat{\mathcal{T}})^*$, and we use $\mathsf{tail}(d)$ to denote K_n. Define $\widehat{\mathcal{I}}$ by starting with

$$\Delta^{\widehat{\mathcal{I}}} = \mathsf{KON}(\widehat{\mathcal{T}})$$
$$A^{\widehat{\mathcal{I}}} = \{K \in \mathsf{KON}(\widehat{\mathcal{T}}) \mid K \sqsubseteq A \in \widehat{\mathcal{T}}\}$$
$$r^{\widehat{\mathcal{I}}} = \emptyset$$

Observe that since A_0 occurs in $\widehat{\mathcal{T}}$ and $A_0 \sqsubseteq \bot \notin \widehat{\mathcal{T}}$, $\Delta^{\widehat{\mathcal{I}}}$ contains the conjunction $K = A_0$ and thus $A_0^{\widehat{\mathcal{I}}} \neq \emptyset$. We finish the construction of $\widehat{\mathcal{I}}$ by exhaustively applying the following rule: if there is some $d \in \Delta^{\widehat{\mathcal{I}}}$ with $\mathsf{tail}(d) \sqsubseteq \exists r.K' \in \widehat{\mathcal{T}}$, K' maximal with this property, and $d \notin (\exists r.K')^{\widehat{\mathcal{I}}}$, then add a fresh element $e = dK'$ to $\Delta^{\widehat{\mathcal{I}}}$, add (d, K') to $r^{\widehat{\mathcal{I}}}$, and add dK' to $A^{\widehat{\mathcal{I}}}$ whenever $K' \sqsubseteq A \in \widehat{\mathcal{T}}$.

We first show that $\widehat{\mathcal{I}}$ is a model of $\widehat{\mathcal{T}}$, which amounts to a case distinction over the forms of CIs that can be present in $\widehat{\mathcal{T}}$, in each case relying on the fact that $\widehat{\mathcal{T}}$ is closed under the rules of the calculus. Details are provided in the appendix.

Lemma 10 $\widehat{\mathcal{I}} \models \widehat{\mathcal{T}}$.

It remains to show that $\widehat{\mathcal{I}}$ is a model of $\mathcal{T}_\mathsf{f}$, which is significantly more difficult to prove than Lemma 10 due to the fact that $\mathcal{T}_\mathsf{f}$ is obtained by reversing all cycles in $\mathcal{T}$ whereas the calculus is more careful to reverse only the 'relevant' ones, as explained above. We start with the observation that, when constructing $\mathcal{T}_\mathsf{f}$, it suffices to close only maximal cycles. More precisely, a cycle $K_1, r_1, K_2, \ldots, K_n$ in a TBox $\mathcal{T}$ is *maximal* if K_{j+1} is maximal with $\mathcal{T} \models K_j \sqsubseteq \exists r_j.K_{j+1}$, for $1 \leq j < n$. Let $\mathcal{T}_\mathsf{f}^{\max}$ be the variation of $\mathcal{T}_\mathsf{f}$ that is obtained by reversing only maximal cycles.

Lemma 11 *$\mathcal{T}_\mathsf{f}$ is equivalent to $\mathcal{T}_\mathsf{f}^{max}$.*

To finish the proof of Theorem 8, let $\mathcal{T}_\mathsf{f}^0, \mathcal{T}_\mathsf{f}^1, \ldots$ be the sequence of TBoxes obtained by starting with $\mathcal{T}_\mathsf{f}^0 = \mathcal{T}$ and then exhaustively closing maximal cycles, that is, $\mathcal{T}_\mathsf{f}^{\max}$ is the limit of this sequence. In the appendix, we prove by induction on i that $\widehat{\mathcal{I}}$ is a model of each $\mathcal{T}_\mathsf{f}^i$, thus of $\mathcal{T}_\mathsf{f}$.

We now briefly consider an extension of our algorithm to ABox consistency, with Figure 2 showing the additional rules. Instead of starting with only a TBox $\mathcal{T}$, the algorithm now begins with a set $\mathcal{T} \cup \mathcal{A}$, where $\mathcal{T}$ is a TBox and $\mathcal{A}$ an ABox, and then exhaustively applies rules **R1** to **R12**. In rules **R10**

$$\textbf{R10}\ \frac{K(a) \quad K \sqsubseteq A}{A(a)} \qquad \textbf{R11}\ \frac{K(a) \quad r(a,b) \quad K \sqsubseteq \forall r.K'}{K'(b)}$$

$$\textbf{R12}\ \frac{\begin{array}{c} K_1(a) \quad K_2(a) \quad r(a,b) \quad K(b) \quad K_1 \sqsubseteq (\leqslant 1\ r\ A) \\ K_2 \sqsubseteq \exists r.K' \quad K \sqsubseteq A \quad K' \sqsubseteq A \end{array}}{K'(b)}$$

Figure 2: Additional Inference Rules

to **R12**, $K(a)$ is an abbreviation for $A_1(a) \cdots A_k(a)$ when $K = \{A_1, \dots, A_k\}$. Recall that rules **R1** and **R2** only apply when the conjunction in their precondition occurs in the partially completed TBox. For the extension with ABoxes, an additional way for K to *occur* is that, for some ABox individal a, $K = \{A \mid A(a) \text{ is in the partial completion}\}$. It is easy to see that rule application still terminates after exponentially many steps. Let Γ be the set of concept inclusions and ABox assertions finally generated. The algorithm is sound and complete in the sense that $\mathcal{A}$ is finitely inconsistent w.r.t. $\mathcal{T}$ iff there is an ABox individual a and a conjunction K such that Γ contains both $K(a)$ and $K \sqsubseteq \bot$. To prove this, one updates the construction of $\widehat{\mathcal{I}}$ by starting with an initial interpretation defined by setting $\Delta^{\widehat{\mathcal{I}}} = \mathsf{Ind}(\mathcal{A})$, $r^{\widehat{\mathcal{I}}} = \{(a,b) \mid r(a,b) \in \mathcal{A}\}$, and $A^{\widehat{\mathcal{I}}} = \{a \in \mathsf{Ind}(\mathcal{A}) \mid A(a) \in \Gamma\}$. The rest of the construction of $\widehat{\mathcal{I}}$ is as before. It is not hard to adapt the proof of Lemma 10 to show that $\widehat{\mathcal{I}}$ satisfies all inclusions and assertions in Γ. As in the case of finite satisfiability, it thus remains to prove that $\widehat{\mathcal{I}}$ is a model of $\mathcal{T}_\mathsf{f}$. Fortunately, the proof of goes through without modification.

Apart from providing a basis for practical implementations, our algorithm also yields an EXPTIME upper bound for finite ABox consistency in Horn-$\mathcal{ALCFI}$. This result is known from (Lutz, Sattler, and Tendera 2005), where it is shown that ABox consistency in the non-Horn version of $\mathcal{ALCQI}$ is in EXPTIME. A matching lower bound can be derived from (Baader, Brandt, and Lutz 2008) where an EXPTIME lower bound is established for unrestricted subsumption in (the $\mathcal{ELI}$ fragment of) Horn-$\mathcal{ALCFI}$; the proof can easily be adapted to finite satisfiability.

Theorem 12 *Finite satisfiability and finite ABox consistency in Horn-$\mathcal{ALCQI}$ are* EXPTIME*-complete.*

5 Query Answering in the Finite

In the ontology-based data access (OBDA) paradigm, the central reasoning problem is answering database-style queries over ABoxes in the presence of a DL TBox. In this section, we study the finite model version of this problem, assuming that queries are positive existential queries (PEQs) and that TBoxes are formulated in Horn-$\mathcal{ALCFI}$. We show that, as in the case of ABox consistency, finite PEQ answering can be reduced to unrestricted PEQ answering by reversing finmod cycles in the TBox. This result enables the use of algorithms for unrestricted PEQ answering also in the finite case. It also allows us to show that finite PEQ answering w.r.t. Horn-$\mathcal{ALCFI}$ TBoxes is EXPTIME-complete regarding combined complexity, and PTIME-complete regarding data complexity.

We start with a brief introduction of positive existential queries and of query answering. For simplicity, we concentrate on Boolean queries, that is, queries without answer variables. It is, however, easy to adapt all techniques established in this section to the case of queries with answer variables. A *(Boolean) positive existential query (PEQ)* q takes the form $\exists \mathbf{x}\, \varphi(\mathbf{x})$ where φ is built from atoms of the form $A(x)$ and $r(x,y)$ using conjunction and disjunction, with x, y variables from $\mathbf{x}$, A a concept name, and r a role name. Let $\mathcal{I}$ be an interpretation and $q = \exists \mathbf{x}\, \varphi$ a PEQ. A *match* of q in $\mathcal{I}$ is a mapping $\pi : \mathbf{x} \to \Delta^{\mathcal{I}}$ such that φ evaluates to true unter the valuation that assigns true to an atom $A(x)$ in φ iff $\pi(x) \in A^{\mathcal{I}}$ and true to an atom $r(x,y)$ in φ iff $(\pi(x), \pi(y)) \in r^{\mathcal{I}}$. We write $\mathcal{I} \models q$ if there is a match of q in $\mathcal{I}$. For an ABox $\mathcal{A}$ and a TBox $\mathcal{T}$, we write $\mathcal{A}, \mathcal{T} \models q$ (resp. $\mathcal{A}, \mathcal{T} \models_\mathsf{fin} q$) if $\mathcal{I} \models q$ for all models (resp. finite models) $\mathcal{I}$ of $\mathcal{T}$ and $\mathcal{A}$. We then say that $\mathcal{A}, \mathcal{T}$ *entails* (resp. *finitely entails*) q. The problem that we are interested in is *finite query entailment*, that is, given an ABox $\mathcal{A}$, a TBox $\mathcal{T}$, and a query q, to decide whether $\mathcal{A}, \mathcal{T} \models_\mathsf{fin} q$. We will study both the combined complexity and the data complexity of this problem. When studying combined complexity, all of $\mathcal{A}$, $\mathcal{T}$, and q are considered an input. In the case of data complexity, $\mathcal{T}$ and q are assumed to be fixed and q is the only input.

The main result of this section is the following theorem, where $\mathcal{T}_\mathsf{f}$ is the TBox obtained from $\mathcal{T}$ by exhaustively reversing finmod cycles, exactly as in Section 3.

Theorem 13 *Let $\mathcal{T}$ be a Horn-$\mathcal{ALCFI}$ TBox and $\mathcal{A}$ an ABox that is finitely consistent w.r.t. $\mathcal{T}$. For any PEQ q,*

$$\mathcal{A}, \mathcal{T} \models_\mathsf{fin} q \ \textit{ iff } \ \mathcal{A}, \mathcal{T}_\mathsf{f} \models q$$

The proof of the "$\Leftarrow$" direction is trivial. Indeed, if $\mathcal{A}, \mathcal{T} \not\models_\mathsf{fin} q$, then there is a finite model $\mathcal{I}$ of $\mathcal{A}$ and $\mathcal{T}$ such that $\mathcal{I} \not\models q$. Since every finite model of $\mathcal{T}$ is also a model of $\mathcal{T}_\mathsf{f}$ by Lemma 4, it follows that $\mathcal{A}, \mathcal{T}_\mathsf{f} \not\models q$.

For the proof of the "$\Rightarrow$" direction, we use a well-known (infinite) canonical model $\mathcal{U}$ of $\mathcal{A}$ and $\mathcal{T}_\mathsf{f}$, constructed by starting with the following initial interpretation

$$\begin{aligned} \Delta^{\mathcal{U}} &= \mathsf{Ind}(\mathcal{A}) \\ A^{\mathcal{U}} &= \{a \in \mathsf{Ind}(\mathcal{A}) \mid \mathcal{A}, \mathcal{T}_\mathsf{f} \models A(a)\} \\ r^{\mathcal{U}} &= \{(a,b) \mid r(a,b) \in \mathcal{A}\} \end{aligned}$$

and then exhaustively applying the following completion rule: for all $d \in \Delta^{\mathcal{U}}$ such that $\mathcal{T}_\mathsf{f} \models \mathsf{tp}_{\mathcal{U}}(d) \sqsubseteq \exists r.t'$, where t' is maximal with this property and $d \notin (\exists r.t')^{\mathcal{U}}$, add a fresh element d' to $\Delta^{\mathcal{U}}$, the edge (d, d') to $r^{\mathcal{U}}$, and d' to the interpretation $A^{\mathcal{U}}$ of all concept names $A \in t'$.

The following properties of $\mathcal{U}$ are well-known and the reason for why $\mathcal{U}$ is called canonical (Krisnadhi and Lutz 2007; Eiter et al. 2008; Ortiz, Rudolph, and Šimkus 2011).

Lemma 14

1. *$\mathcal{U}$ is a model of $\mathcal{A}$ and of $\mathcal{T}_\mathsf{f}$;*
2. *For any PEQ q, we have that $\mathcal{A}, \mathcal{T}_\mathsf{f} \models q$ iff $\mathcal{U} \models q$.*

By Point 2 of Lemma 14, we can establish the "⇒" direction of Theorem 13 by showing that $\mathcal{A}, \mathcal{T} \models_{\mathsf{fin}} q$ implies $\mathcal{U} \models q$. The proof makes intense use of homomorphisms. For interpretations $\mathcal{I}_1, \mathcal{I}_2$, a *homomorphism from $\mathcal{I}_1$ to $\mathcal{I}_2$* is a function $h : \Delta^{\mathcal{I}_1} \rightarrow \Delta^{\mathcal{I}_2}$ such that

1. $h(a) = a$ for all $a \in \mathsf{N_I}$;
2. $d \in A^{\mathcal{I}_1}$ implies $h(d) \in A^{\mathcal{I}_2}$ for all concept names A;
3. $(d, e) \in r^{\mathcal{I}_1}$ implies $(h(d), h(e)) \in r^{\mathcal{I}_2}$ for all (possibly inverse) roles r.

For $n > 0$, an *n-substructure* of an interpretation $\mathcal{I}$ is an interpretation $\mathcal{I}'$ obtained from $\mathcal{I}$ by selecting a domain $\Delta^{\mathcal{I}'} \subseteq \Delta^{\mathcal{I}}$ with at most n elements and restricting $\mathcal{I}$ to $\Delta^{\mathcal{I}'}$. To show that $\mathcal{A}, \mathcal{T} \models_{\mathsf{fin}} q$ implies $\mathcal{U} \models q$, it suffices to establish the following.

Proposition 15 *For every $n_0 > 0$, there is a finite model $\mathcal{J}_{n_0}$ of $\mathcal{A}$ and $\mathcal{T}$ such that there is a homomorphism from any n_0-substructure of $\mathcal{J}_{n_0}$ to $\mathcal{U}$.*

In fact, $\mathcal{A}, \mathcal{T} \models_{\mathsf{fin}} q$ implies $\mathcal{J}_{n_0} \models q$ and thus there is an n_0-substructure $\mathcal{J}$ of $\mathcal{J}_{n_0}$ with $\mathcal{J} \models q$, where n_0 is the number of variables in q. The latter is witnessed by a match π. By Proposition 15, there is a homomorphism h from $\mathcal{J}$ to $\mathcal{U}$ and thus a match of q in $\mathcal{U}$ can be found by composing π with h.

We construct the model $\mathcal{J}$ from Proposition 15 by modifying the finite model $\mathcal{I}$ constructed in Section 3. For two reasons, the finite model $\mathcal{I}$ constructed in Section 3 need not satisfy the condition formulated for $\mathcal{J}_{n_0}$ in Proposition 15.

1. $\mathcal{I}$ can contain paths of length $\leq n_0$ that do not exist in $\mathcal{U}$.
2. $\mathcal{I}$ can contain cycles that do not exclusively consist of ABox elements, while no such cycles are present in $\mathcal{U}$.

Let us start with Problem 1 above. There are, in turn, two sources for paths in $\mathcal{I}$ that we cannot reproduce in $\mathcal{U}$.

(i) Application of $(\mathbf{c3})$ can generate a path $(d_1, d) \in r^{\mathcal{I}}$, $(d, d_2) \in s^{\mathcal{I}}$ such that $\mathsf{tp}_{\mathcal{I}}(d_1) \rightarrow_r \mathsf{tp}_{\mathcal{I}}(d) \;{}_s\!\leftarrow \mathsf{tp}_{\mathcal{I}}(d_2)$ and d is not identified by an ABox element. Such situations are not necessarily reproducible in $\mathcal{U}$. As a concrete example, consider

$$\mathcal{A} = \{\ B_1(a), B_2(b)\ \}$$
$$\mathcal{T} = \{\ B_1 \sqsubseteq \exists r.A,\ B_2 \sqsubseteq \exists r.A\ \}.$$

The problematic path is $(a, d_t) \in r^{\mathcal{I}}$, $(d_t, b) \in (r^-)^{\mathcal{I}}$ with $t = \{A\}$.

(ii) Application of $(\mathbf{c2})$ can result in similar a situation as above, but where the middle element d is replaced with a sequence of elements $e_0, \ldots, e_k$ such that $(e_i, e_{i+1}) \in r_i^{\mathcal{I}}$ for all $i < k$ (for some roles $r_0, \ldots, r_{k-1}$) and

$$\mathsf{tp}_{\mathcal{I}}(e_0)\ {}^1\!\leftrightarrow^1_{r_1} \cdots {}^1\!\leftrightarrow^1_{r_{k-1}}\ \mathsf{tp}_{\mathcal{I}}(e_k). \qquad (18)$$

For a very simple example, take

$$\mathcal{A} = \{\ B_1(a),\ B_2(b)\ \}$$

and assume that $\mathcal{T}$ is such that $B_1\ {}^1\!\leftrightarrow^1_r\ B_2$. Then an application of $(\mathbf{c2})$ will simply add $r(a, b)$, an edge that does not exist in $\mathcal{U}$.

To obtain the desired model $\mathcal{J}_{n_0}$ from Proposition 15, we first solve Problems (i) and (ii) above, and then Problem 2. To make precise what we mean by this, we introduce bounded simulations, a weakening of homomorphisms. A *bounded simulation of $\mathcal{I}_1$ in $\mathcal{I}_2$* is a relation $\rho \subseteq \Delta^{\mathcal{I}_1} \times \mathbb{N} \times \Delta^{\mathcal{I}_2}$ such that for all $(d, i, e) \in \rho$, the following conditions are satisfied:

1. if $d \in A^{\mathcal{I}_1}$, then $e \in A^{\mathcal{I}_2}$;
2. if $i > 0$ and $(d, d') \in r^{\mathcal{I}_1}$ for some (possibly inverse) role r, then there is an $e' \in \Delta^{\mathcal{I}_2}$ with $(e, e') \in r^{\mathcal{I}_2}$ and $(d', i-1, e') \in \rho$.

We write $(\mathcal{I}_1, d) \preceq_k (\mathcal{I}_2, e)$, for $d \in \Delta^{\mathcal{I}_1}$ and $e \in \Delta^{\mathcal{I}_2}$, if there is a bounded simulation of $\mathcal{I}_1$ in $\mathcal{I}_2$ such that $(d, k, e) \in \rho$ and for all $a \in \mathsf{N_I} \cap \Delta^{\mathcal{I}_1}$, we have $(a, k, a) \in \rho$. Then $\mathcal{I}_1 \preceq_k \mathcal{I}_2$ denotes that for every $d \in \Delta^{\mathcal{I}_1}$, there is an $e \in \Delta^{\mathcal{I}_2}$ with $(\mathcal{I}_1, d) \preceq_k (\mathcal{I}_2, e)$. We write $(\mathcal{I}_1, d) \sim_k (\mathcal{I}_2, e)$ if $(\mathcal{I}_1, d) \preceq_k (\mathcal{I}_2, e)$ and vice versa.

With solving Problems (i) and (ii), we mean to establish the following intermediate result.

Proposition 16 *For every $n_0 > 0$, there is a finite model $\mathcal{I}_{n_0}$ of $\mathcal{A}$ and $\mathcal{T}$ such that $\mathcal{I}_{n_0} \preceq_{n_0} \mathcal{U}$.*

To remove the undesired paths illustrated in (i) above, we modify the construction of $\mathcal{I}$ by replacing the elements d_t, $t \in \mathsf{TP}(\mathcal{T}_\mathsf{f})$, that are introduced at the beginning of the construction of $\mathcal{I}$ and used as 'targets' for role edges introduced by applications of $(\mathbf{c3})$. In the modified construction, we instead introduce one $(\mathbf{c3})$-target for each *n_0-bounded simulation type*, which is an equivalence class of $\sim_{n_0}$ on the set of all pointed interpretations $(\mathcal{I}_1, d)$. In the example given in (i) above, the result is that the two existential restrictions would no longer be witnessed by the same d_t because the 1-simulation type of the witnesses are different (one has an r-predecessor in B_1, the other in B_2). Since simulations need only to consider symbols that occur in the (fixed) ABox $\mathcal{A}$ and (fixed) TBox $\mathcal{T}$, there are only finitely many n_0-simulation types and thus finiteness of $\mathcal{I}$ is not compromised.

Undesired paths of type (ii) are avoided by modifying the $(\mathbf{c2})$ rule so that the sequences (18) are of length exceeding n_0 and thus the highlighted problem which involves both ends of the sequence is not 'visible' in n_0-substructures. We also include an initial piece of the canonical model $\mathcal{U}$ for $\mathcal{A}$ and $\mathcal{T}_\mathsf{f}$ of depth n_0 in the initial version of $\mathcal{I}$ to avoid the undesired 'shortcuts' between ABox elements illustrated by the example given in (ii) above.

The construction is spelled out in full detail in the appendix. We have actually omitted some aspects in the overview above for the sake of a clearer exposition, such as the fact that we first exhaustively apply rules $(\mathbf{c1})$ and $(\mathbf{c2})$, followed by exhaustive application of $(\mathbf{c3})$ (the latter two in their modified versions), and that we actually cannot include in the initial $\mathcal{I}$ *all* n_0-bounded simulation types, but must select only the 'relevant' ones. This finishes the proof of Proposition 16.

To solve Problem 2 above and thus obtain the model $\mathcal{J}_{n_0}$ stipulated by Proposition 15, we have to eliminate all non-ABox-cycles of size at most n_0 in the model $\mathcal{I}_{n_0}$ delivered by Proposition 16. This is achieved by taking the product with a suitable finite group of high girth, a technique championed by Otto (2012). Details are provided in the appendix. This finishes the proof of Theorem 13.

Apart from enabling the use of algorithms for unrestricted PEQ answering also in the finite case, Theorem 13 yields tight complexity bounds for finite PEQ entailment.

Theorem 17 *Finite PEQ entailment in Horn-$\mathcal{ALCFI}$ is decidable,* ExpTime*-complete in combined complexity, and* PTime*-complete in data complexity.*

Proof.(sketch) For the unrestricted case, an ExpTime lower bound is in (Baader, Brandt, and Lutz 2008) and a PTime one in (Calvanese et al. 2006). Both results can easily be adapted to the finite case. The upper bounds can be proved using the following straightforward algorithm for PEQ entailment, which resembles existing algorithms such as those presented in (Krisnadhi and Lutz 2007; Eiter et al. 2008; Calì, Gottlob, and Lukasiewicz 2009; Ortiz, Rudolph, and Šimkus 2011). Assume that an input ABox $\mathcal{A}$, TBox $\mathcal{T}$, and PEQ q are given, and let n_0 be the number of variables in q. As a consequence of Theorem 3, finite satisfiability w.r.t. $\mathcal{T}$ coincides with unrestricted satisfiability w.r.t. $\mathcal{T}_\mathsf{f}$. Using our algorithm for computing finite satisfiability in Horn-$\mathcal{ALCFI}$ in ExpTime, we can thus compute the set $\mathsf{TP}(\mathcal{T}_\mathsf{f})$ of types for $\mathcal{T}_\mathsf{f}$ without computing $\mathcal{T}_\mathsf{f}$ or explicitly reasoning w.r.t. this exponentially large TBox. Let $\mathcal{A}'$ be the extension of $\mathcal{A}$ with assertions $\{A(a_t) \mid A \in t\}$ for each $t \in \mathsf{TP}(\mathcal{A})$. Now compute an initial piece $\mathcal{U}'$ of the canonical model $\mathcal{U}$ of $\mathcal{A}'$ and $\mathcal{T}_\mathsf{f}$, namely its restriction to depth n_0. Similar to the computation of $\mathsf{TP}(\mathcal{T}_\mathsf{f})$ above, we can do this by using finite subsumption w.r.t. $\mathcal{T}$ instead of unrestricted subsumption w.r.t. $\mathcal{T}_\mathsf{f}$. It is not difficult to prove that $\mathcal{U}' \models q$ iff $\mathcal{U} \models q$. To check whether $\mathcal{U}' \models q$ within the desired time bounds, we can simply enumerate all possible maps of variables in q to elements of $\mathcal{U}'$ and check whether any such map is a match. ❑

Note that decidability of PEQ entailment in Horn-$\mathcal{ALCFI}$ was expected given a result by Pratt-Hartmann which states that finite CQ answering for the two-variable guarded fragment of first-order logic extended with counting quantifiers is decidable (Pratt-Hartmann 2009). We assume that his proof can be extended to unions of conjunctive queries (UCQs), thus to PEQs. Pratt-Hartmann also analyses the data complexity of finite CQ answering in his logic, but finds it to be coNP-complete. He does not analyse combined complexity. Theorem 17 suggests that PEQ entailment in Horn-$\mathcal{ALCFI}$ has the same complexity in finite and in unrestricted models. For the unrestricted case, PTime-completeness in data complexity follows from the results in (Hustadt, Motik, and Sattler 2007), and ExpTime-completeness in combined complexity is proved in (Eiter et al. 2008) for UCQs. We assume that the techniques in that paper extend to PEQs.

6 From Horn-$\mathcal{ALCFI}$ to Horn-$\mathcal{ALCQI}$

Our results for finite satisfiability and finite subsumption (the reasoning tasks that do not involve ABoxes) extend in a straightforward way from Horn-$\mathcal{ALCFI}$ to Horn-$\mathcal{ALCQI}$. In particular, we can convert a Horn-$\mathcal{ALCQI}$ TBox $\mathcal{T}$ into a Horn-$\mathcal{ALCFI}$ TBox $\mathcal{T}'$ such that finite (un)satisfiability is preserved by replacing each CI $K \sqsubseteq (\geqslant n\ r\ K')$ in $\mathcal{T}$ with the following inclusions, for $1 \leq i < j \leq n$:

$$K \sqsubseteq \exists r.B_i, \quad B_i \sqsubseteq K', \quad B_i \sqcap B_j \sqsubseteq \bot \qquad (*)$$

While an easy unraveling argument can be used to prove that this reduction is correct in the presence of infinite models, more care is required in the finite case (see appendix).

Proposition 18 *$\mathcal{T}$ is finitely satisfiable iff $\mathcal{T}'$ is finitely satisfiable.*

It follows from Proposition 18 and Theorem 3 that a Horn-$\mathcal{ALCQI}$ TBox $\mathcal{T}$ is finitely satisfiable iff $(\mathcal{T}')_\mathsf{f}$ is satisfiable. Actually, it is not hard to see that this is the case iff $\mathcal{T}_\mathsf{f}$ (the result of applying cycle reversion directly to the Horn-$\mathcal{ALCQI}$ TBox, ignoring all inclusions $A \sqsubseteq (\geqslant n\ r\ C))$ is satisfiable because if any of the existential restrictions in $\mathcal{T}' \setminus \mathcal{T}$ is involved in a finmod cycle, then a simple semantic argument shows that both $\mathcal{T}_\mathsf{f}$ and $(\mathcal{T}')_\mathsf{f}$ are unsatisfiable. Proposition 18 also enables the use of our consequence-based procedure for deciding finite satisfiability in Horn-$\mathcal{ALCQI}$.

It is not immediately obious how to extend $(*)$ and Proposition 18 to ABox consistency and instance checking. We believe, though, that it is not too hard to modify the proof of Theorem 3 for Horn-$\mathcal{ALCQI}$, to adapt the consequence-based procedure to allow a direct treatment of Horn-$\mathcal{ALCQI}$ TBoxes without prior reduction to Horn-$\mathcal{ALCFI}$, and to extend all model constructions underlying our results about PEQ entailment to Horn-$\mathcal{ALCQI}$. In particular, such a direct approach should yield ExpTime/ PTime upper bounds for PEQ entailment in Horn-$\mathcal{ALCQI}$ even when the numbers in at least restrictions are coded in binary (note that, in this case, the translation $(*)$ incurs an exponential blowup).

7 Future Work

As future research, it would be interesting to extend the results in this paper to Horn-$\mathcal{SHIQ}$, that is, to add role hierarchies and transitive roles. Reducing out role hierarchies does not seem easily possible in the finite,[2] so they would have to be built directly into all constructions. For query entailment, we expect transitive roles to cause significant additional challenges, see for example (Eiter et al. 2009; Mosurovic et al. 2013). In particular, transitive roles result in an additional way in which the finite model property is lost, illustrated by the TBox $\mathcal{T} = \{A \sqsubseteq \exists r.A, \mathsf{trans}(r)\}$ and the conjunctive query $q = \exists x\, r(x, x)$. We have $\{A(a)\}, \mathcal{T} \not\models q$, but $\{A(a)\}, \mathcal{T} \models_\mathsf{fin} q$ although neither counting nor inverse roles are present (the TBox $\mathcal{T}$ is formulated in the DL $\mathcal{EL}_\mathsf{trans}$). Finite model reasoning in versions of Datalog$^\pm$ that extend $\mathcal{EL}_\mathsf{trans}$ has recently been studied in (Gogacz and Marcinkowski 2013b; 2013a).

In this paper, we have not analyzed the size of finite models. It is, however, easy to prove a double exponential lower bound on the size of finite models for satisfiability in Horn-$\mathcal{ALCFI}$ by enforcing a tree of exponential depth in which no two elements can be identical. A matching upper bound follows from Pratt-Hartmann's result that every finitely satisfiable formula in first-order logic with two variables and

[2]In contrast to what we have claimed in the workshop predecessor of this paper (Ibáñez-García, Lutz, and Schneider 2013).

counting quantifiers has a model of at most double exponential size (Pratt-Hartmann 2005). Analyzing the size of finite (counter)models for query entailment is left as future work.

Acknowledgements Carsten Lutz was supported by the SFB/TR8 *Spatial Cognition*.

References

Baader, F.; Calvanese, D.; McGuinness, D.; Nardi, D.; and Patel-Schneider, P., eds. 2003. *The Description Logic Handbook: Theory, Implementation, and Applications*. Cambridge University Press.

Baader, F.; Brandt, S.; and Lutz, C. 2005. Pushing the $\mathcal{EL}$ envelope. In *Proc. of IJCAI-05*, 364–369.

Baader, F.; Brandt, S.; and Lutz, C. 2008. Pushing the $\mathcal{EL}$ envelope further. In *Proc. OWLED-08 DC*, volume 496 of *CEUR-WS.org*.

Baader, F.; Lutz, C.; and Suntisrivaraporn, B. 2006. CEL – a polynomial-time reasoner for life science ontologies. In *Proc. IJCAR-06*, volume 4130 of *LNCS*, 287–291. Springer.

Bienvenu, M.; Lutz, C.; and Wolter, F. 2013. First-order rewritability of atomic queries in Horn description logics. In *Proc. of IJCAI-13*. IJCAI/AAAI.

Calì, A.; Gottlob, G.; and Lukasiewicz, T. 2009. A general datalog-based framework for tractable query answering over ontologies. In *PODS*, 77–86.

Calvanese, D.; Giacomo, G. D.; Lembo, D.; Lenzerini, M.; and Rosati, R. 2006. Data complexity of query answering in description logics. In *Proc. KR-06*, 260–270. AAAI Press.

Calvanese, D.; Giacomo, G. D.; Lembo, D.; Lenzerini, M.; and Rosati, R. 2007. Tractable reasoning and efficient query answering in description logics: The *DL-Lite* family. *J. Autom. Reasoning* 39(3):385–429.

Calvanese, D. 1996. Finite model reasoning in description logics. In *Proc. of KR-96*, 292–303. Morgan Kaufmann.

Cosmadakis, S.; Kanellakis, P.; and Vardi, M. 1990. Polynomial-time implication problems for unary inclusion dependencies. *J. ACM* 37(1):15–46.

Eiter, T.; Gottlob, G.; Ortiz, M.; and Simkus, M. 2008. Query answering in the description logic Horn-$\mathcal{SHIQ}$. In *Proc. JELIA-08*, volume 5293 of *LNCS*, 166–179. Springer.

Eiter, T.; Lutz, C.; Ortiz, M.; and Simkus, M. 2009. Query answering in description logics with transitive roles. In *Proc. IJCAI-09*, 759–764.

Eiter, T.; Ortiz, M.; Šimkus, M.; Tran, T.-K.; and Xiao, G. 2012. Towards practical query answering for Horn-$\mathcal{SHIQ}$. In *Proc. DL-12*, volume 846 of *CEUR-WS.org*.

Gogacz, T., and Marcinkowski, J. 2013a. Converging to the chase - a tool for finite controllability. In *Proc. LICS-13*, 540–549. IEEE Computer Society.

Gogacz, T., and Marcinkowski, J. 2013b. On the BDD/FC conjecture. In *Proc. PODS-13*, 127–138. ACM.

Hustadt, U.; Motik, B.; and Sattler, U. 2007. Reasoning in description logics by a reduction to disjunctive datalog. *J. Autom. Reasoning* 39(3).

Ibáñez-García, Y. A.; Lutz, C.; and Schneider, T. 2013. Finite model reasoning in Horn-$\mathcal{SHIQ}$. In *Proc. DL-13*, volume 1014 of *CEUR-WS.org*, 234–245.

Kazakov, Y.; Krötzsch, M.; and Simančík, F. 2011a. Unchain my $\mathcal{EL}$ reasoner. In *Proc. DL-11*, volume 745 of *CEUR-WS.org*.

Kazakov, Y.; Krötzsch, M.; and Simančík, F. 2011b. Concurrent classification of el ontologies. In *Proc. ISWC-11 (1)*, volume 7031 of *LNCS*, 305–320. Springer.

Kazakov, Y. 2009. Consequence-driven reasoning for Horn-$\mathcal{SHIQ}$ ontologies. In *Proc. IJCAI-09*, 2040–2045.

Krisnadhi, A., and Lutz, C. 2007. Data complexity in the $\mathcal{EL}$ family of DLs. In *Proc. DL-07*, volume 250 of *CEUR-WS.org*.

Lutz, C., and Wolter, F. 2012. Non-uniform data complexity of query answering in description logics. In *Proc. KR-12*. AAAI Press.

Lutz, C.; Sattler, U.; and Tendera, L. 2005. The complexity of finite model reasoning in description logics. *Information and Computation* 199:132–171.

Mosurovic, M.; Krdzavac, N.; Graves, H.; and Zakharyaschev, M. 2013. A decidable extension of $\mathcal{SROIQ}$ with complex role chains and unions. *J. Artif. Intell. Res. (JAIR)* 47:809–851.

Ortiz, M.; Rudolph, S.; and Šimkus, M. 2011. Query answering in the Horn fragments of the description logics $\mathcal{SHOIQ}$ and $\mathcal{SROIQ}$. In *Proc. IJCAI-11*, 1039–1044. IJCAI/AAAI.

Otto, M. 2012. Highly acyclic groups, hypergraph covers, and the guarded fragment. *J. ACM* 59(1):5.

Pacholski, L.; Szwast, W.; and Tendera, L. 2000. Complexity results for first-order two-variable logic with counting. *SIAM J. Comput.* 29(4):1083–1117.

Pratt-Hartmann, I. 2005. Complexity of the two-variable fragment with counting quantifiers. *J. of Logic, Language and Information* 14(3):369–395.

Pratt-Hartmann, I. 2009. Data-complexity of the two-variable fragment with counting quantifiers. *Inf. Comput.* 207(8):867–888.

Rosati, R. 2008. Finite model reasoning in *DL-Lite*. In *Proc. ESWC-08*, volume 5021 of *LNCS*, 215–229. Springer.