

Complete Approximation of Horn DL Ontologies

Anneke Bötcher¹, Carsten Lutz¹, and Frank Wolter²

¹ University of Bremen, Germany

anneke@uni-bremen.de clu@uni-bremen.de

² University of Liverpool, United Kingdom

wolter@liverpool.ac.uk

Abstract. We study the approximation of expressive Horn DL ontologies in less expressive Horn DLs, with completeness guarantees. Cases of interest include Horn-*SRLF*-to- \mathcal{ELR}_\perp , Horn-*SHLF*-to- \mathcal{ELH}_\perp , and others. Since finite approximations almost never exist, we carefully map out the structure of infinite approximations. This provides a solid theoretical foundation for constructing incomplete approximations in practice in a controlled way. Technically, we exhibit a connection to the axiomatization of quasi-equations valid in classes of semilattices with operators and additionally develop a direct proof strategy based on the chase and on homomorphisms that allows us to also deal with approximations of bounded role depth.

1 Introduction

There is a large number of description logics that vary considerably regarding their expressive power and computational properties [2] and despite prominent standardization efforts, many different DLs continue to be used in ontologies from practical applications.³ As a result, it is often necessary to convert an ontology formulated in some DL \mathcal{L} into another DL \mathcal{L}' , a particularly important case being that \mathcal{L}' is a fragment of \mathcal{L} —if it is not, then one could use the fragment $\mathcal{L} \cap \mathcal{L}'$ of \mathcal{L} in place of \mathcal{L}' . For example, this happens in *ontology import* when an engineer who designs an ontology formulated in \mathcal{L}' wants to reuse content from an existing ontology formulated in \mathcal{L} . The problem that emerges from this is *ontology approximation*, a form of knowledge compilation [27, 11].

In this paper, we are interested in approximating an ontology \mathcal{O}_E formulated in an expressive description logic \mathcal{L} by an ontology \mathcal{O}_L formulated in a fragment \mathcal{L}' of \mathcal{L} . We aim to construct \mathcal{O}_L so that it preserves all information from \mathcal{O}_E expressible in \mathcal{L}' , called a *greatest lower bound* in knowledge compilation [27]. Formally, for every \mathcal{L}' concept inclusion $C \sqsubseteq D$ that is formulated in the signature Σ of \mathcal{O}_E , we require that $\mathcal{O}_E \models C \sqsubseteq D$ if and only if $\mathcal{O}_L \models C \sqsubseteq D$, and likewise for role inclusions and any other type of ontology statement supported by \mathcal{L}' . We say that \mathcal{O}_L is *sound* as an approximation if it satisfies the “if” part of this property and *complete* if it satisfies the “only if” part. It is equivalent to

³ See, for example, the BioPortal repository at <https://biportal.bioontology.org/>.

demand that $\mathcal{O}_E \models \mathcal{O}_L$ and for every \mathcal{L} ontology \mathcal{O} in signature Σ , $\mathcal{O}_E \models \mathcal{O}$ implies $\mathcal{O}_L \models \mathcal{O}$. We consider both the case where \mathcal{O}_L is formulated in Σ (*non-projective* approximation) and the case where additional symbols are admitted (*projective* approximation).

In practice, approximations are often constructed in an ad hoc way that is sound but not complete. For example, it is common to simply replace all subconcepts of \mathcal{O}_E that use a constructor which is not available in \mathcal{L}' with top or with bottom. In fact, full completeness is typically not attainable as it brings about infinite ontologies even in simple cases. For example, consider the \mathcal{ELI} ontology $\mathcal{O}_E = \{\exists \text{hasSupervisor}^-. \top \sqsubseteq \text{Manager}\}$ which contains only a single range restriction. There is no finite \mathcal{EL} approximation \mathcal{O}_L since for all $n \geq 1$, \mathcal{O}_E entails the \mathcal{EL} concept inclusion

$$\exists \text{hasSupervisor}^n. \top \sqsubseteq \exists \text{hasSupervisor}. (\text{Manager} \sqcap \exists \text{hasSupervisor}^{n-1}. \top).$$

Also for the \mathcal{ELF} ontology $\mathcal{O}_E = \{\text{func}(\text{hasSupervisor}), \text{func}(\text{reportsTo})\}$, there is no finite \mathcal{EL} approximation \mathcal{O}_L since for all $n, m \geq 1$, \mathcal{O}_E entails the \mathcal{EL} concept inclusion⁴

$$\begin{aligned} \exists \text{reportsTo}. \exists \text{hasSupervisor}^n. \top \sqcap \exists \text{reportsTo}^m. \top &\sqsubseteq \\ \exists \text{reportsTo}. (\exists \text{hasSupervisor}^n. \top \sqcap \exists \text{reportsTo}^{m-1}. \top). & \end{aligned}$$

Nevertheless, ad hoc ways to construct approximations can be a problematic choice since it is often only poorly understood which information has been given up. This is particularly worrying in ontology import where one would like to construct an as meaningful approximation as possible, rather than just *some* approximation. Our aim is to address this problem by carefully studying the structure of infinite approximations. As the expressive DL \mathcal{L} , we consider Horn-*SRI* \mathcal{F} and fragments thereof. As the fragment \mathcal{L}' , we consider \mathcal{ELR}_\perp and corresponding fragments thereof, where \mathcal{ELR}_\perp denotes the extension of \mathcal{ELH}_\perp with role inclusions of the form $r_1 \circ \dots \circ r_n \sqsubseteq r$. For example, we study Horn-*SRI* \mathcal{F} -to- \mathcal{ELR}_\perp approximation, Horn-*SHI* \mathcal{F} -to- \mathcal{ELH}_\perp , \mathcal{ELI} -to- \mathcal{EL} , \mathcal{ELF} -to- \mathcal{EL} , and so on. Subsumption is EXPTIME-complete in all mentioned source DLs and PTIME-complete in all mentioned target DLs [1, 20]. We thus support ontology designers who build an ontology in a tractable DL and want to import in a well-understood way from an existing ontology formulated in a more costly DL, without compromising tractability.

We provide the following results. In Section 3, we present non-projective approximations for the $\mathcal{ELR}^{\mathcal{I}}\mathcal{F}_\perp$ -to- \mathcal{ELR}_\perp case and several subcases such as $\mathcal{ELH}^{\mathcal{I}}_\perp$ -to- \mathcal{ELH}_\perp . Here, superscript \mathcal{I} means that inverse roles are admitted only in role hierarchies of the form $r \sqsubseteq s^-$ but not in concept inclusions and more complex role inclusions, which is actually a very common way to use inverse roles in practice. The presented approximation requires that \mathcal{O}_E is *inverse closed*, meaning that for every role name r in \mathcal{O}_E , there is a role name \hat{r} that is defined via

⁴ A single functionality assertion would also do, but it is convenient for the example to have two role names in the signature of \mathcal{O}_E .

role hierarchies to be the inverse of r . This also yields *projective* approximations for the case where inverse closedness is not assumed and for the Horn- $SRLF$ -to- \mathcal{ELR}_\perp case through a well-known normalization procedure. The completeness proof is based on a novel connection between ontology approximation and the axiomatizations of quasi-equations valid in classes of semilattices with operators (SLOs); note that SLOs have been used before to obtain algorithms for subsumption in extensions of \mathcal{EL} [28, 29].

In Section 4, we construct non-projective $\mathcal{ELH}^T\mathcal{F}_\perp$ -to- \mathcal{ELH}_\perp approximations under the mild assumption that whenever $\mathcal{O}_E \models r \sqsubseteq s^-$, then neither $\text{func}(s) \in \mathcal{O}_E$ nor $\text{func}(s^-) \in \mathcal{O}_E$. This again encompasses relevant subcases such as \mathcal{ELHF} -to- \mathcal{ELH} , without any syntactic assumptions. The main difference to Section 3 is how completeness is established. Here, we find a chase procedure for \mathcal{L} ontologies that is inspired by and closely linked to the proposed approximation scheme, prove that it is complete for consequences formulated in \mathcal{L}' , and then show that consequences computed by the chase can also be derived using \mathcal{O}_L . In contrast to the approach based on SLOs, this enables us to also prove that finite approximations always exist if we restrict the role depth of concepts in consequences to be bounded by a constant; we speak of *depth bounded approximations*.

We then proceed to study \mathcal{ELI}_\perp -to- \mathcal{EL}_\perp approximations in Section 5. In contrast to the cases considered before, where both \mathcal{L} and \mathcal{L}' are based on the concept language \mathcal{EL}_\perp , here the *concept language* of \mathcal{L} (which is \mathcal{ELI}_\perp) different from the one of \mathcal{L}' (which is \mathcal{EL}_\perp). We present non-projective approximations for unrestricted ontologies \mathcal{O}_E and for ontologies \mathcal{O}_E which are in the well-known normal form for \mathcal{ELI}_\perp ontologies that avoids syntactic nesting of concepts. The two approximation schemes are remarkably different, the latter arguably being more informative than the former.

In Section 6, we complement the proposed infinite approximations by showing that finite approximations do not exist even in simple cases and that depth bounded approximations are non-elementary in size even in simple cases.

Proof details are available in the appendix of the long version, available at <http://www.informatik.uni-bremen.de/tdki/research/papers.html>

Related Work. Approximation in a DL context was first studied in [27] where \mathcal{FL} concepts are approximated by \mathcal{FL}^- concepts and in [9] where \mathcal{ALC} concepts are approximated by \mathcal{ALE} concepts. In both cases, the approximation always exists, but ontologies are not considered. An incomplete approach to approximating \mathcal{SHOIN} ontologies in DL-Lite \mathcal{F} is presented in [24] and complete (projective) approximations of \mathcal{SROIQ} ontologies in DL-Lite \mathcal{A} are given in [7]. Such approximations are guaranteed to exist due to the limited expressive power of DL-Lite \mathcal{A} . Approximation of Horn- \mathcal{ALCHIQ} ontologies by DL-Lite \mathcal{R} ontologies in an OBDA context was considered in [8], exploiting the mapping formalism available in OBDA. In [22], approximation of \mathcal{ELU} ontologies in terms of \mathcal{EL} ontologies is studied, the main result being that it is EXPTIME-hard and in 2EXPTIME to decide a finite complete approximation exists. An incomplete approach to approximating \mathcal{SROIQ} ontologies in \mathcal{EL}^{++} is in [25]. There are also

approaches towards efficient DL reasoning that involve computing approximations, which may be greatest lower bounds and/or least upper bounds. Such approximations are intentionally incomplete in order to not compromise efficiency, see for example [26, 14, 10]. Related to approximation is the problem whether a given \mathcal{L} ontology can be equivalently rewritten into the fragment \mathcal{L}' of \mathcal{L} , either non-projectively [21] or projectively [19]; note that this asks whether we have to approximate at all. There are also various approaches to OBDA with expressive DLs that involve forms of approximation such as [30, 31, 12, 15].

2 Preliminaries

Let \mathbb{N}_C and \mathbb{N}_R be disjoint and countably infinite sets of *concept* and *role names*. A *role* is a role name r or an *inverse role* r^- , with r a role name. A *Horn-SRIF concept inclusion (CI)* is of the form $L \sqsubseteq R$, where L and R are concepts defined by the syntax rules

$$\begin{aligned} R, R' &::= \top \mid \perp \mid A \mid \neg A \mid R \sqcap R' \mid \neg L \sqcup R \mid \exists \rho. R \mid \forall \rho. R \\ L, L' &::= \top \mid \perp \mid A \mid L \sqcap L' \mid L \sqcup L' \mid \exists \rho. L \end{aligned}$$

with A ranging over concept names and ρ over roles. The *depth* of a concept R or L is the nesting depth of existential and universal restrictions in it. For example, the depth of $\exists r. B \sqcap \exists r. \exists s. A$ is two. A *Horn-SRIF ontology* \mathcal{O} is a set of Horn-SRIF CIs, *functionality assertions* $\text{func}(\rho)$, *transitivity assertions* $\text{trans}(\rho)$, and *role inclusions (RIs)* $\rho_1 \circ \dots \circ \rho_n \sqsubseteq \rho$. A role inclusion of the special form $\rho_1 \sqsubseteq \rho_2$ is a *role hierarchy (RH)*. We adopt the standard assumption that for any RI $\rho_1 \circ \dots \circ \rho_n \sqsubseteq \rho$ with $n \geq 2$, we neither have $\mathcal{O} \models \text{func}(\rho)$ nor $\mathcal{O} \models \text{trans}(\rho^-)$. Ontologies used in practice of course have to be finite. In this paper, though, we shall frequently consider also infinite ontologies.

An *ELRIF $_{\perp}$ ontology* is a *Horn-SRIF ontology* in which both the left- and right-hand sides of CIs are *ELI $_{\perp}$ concepts*, defined as usual, and that does not contain transitivity assertions. If \mathcal{O} uses neither inverse roles nor functionality assertions, then it is an *ELR $_{\perp}$ ontology*. We assume w.l.o.g. that, in *ELRIF $_{\perp}$* and below, the \perp concept occurs only in the form $C \sqsubseteq \perp$. We assume that other standard DL names such as Horn-SHIF are understood. It should also be clear what we mean by saying that an ontology language is *based* on a concept language. For example, the ontology language *ELHIF $_{\perp}$* is based on the concept language *ELI $_{\perp}$* . We also use a non-standard naming scheme, namely that $\mathcal{H}^{\mathcal{I}}$ indicates the presence of role hierarchies and of inverse roles in role hierarchies, but not in concept inclusions. $R^{\mathcal{I}}$ is similar, but additionally admits role inclusions while still restricting the use of inverse roles to role hierarchies.

For the semantics and more details on the relevant DLs, we refer to [2]. A *signature* Σ is a set of concept and role names. When speaking of *EL(Σ)* concepts, we mean *EL* concepts that only use concept and role names from Σ , and likewise for other concept languages. We use $\text{sig}(\mathcal{O})$ to denote the set of concept and role names used in ontology \mathcal{O} .

Definition 1. Let \mathcal{O}_E be a Horn-SRIF ontology with $\text{sig}(\mathcal{O}_E) = \Sigma$ and let \mathcal{L} be a fragment of Horn-SRIF based on the concept language \mathcal{C} . A (potentially infinite) \mathcal{L} ontology \mathcal{O}_L is an \mathcal{L} approximation of \mathcal{O}_E if the following conditions are satisfied:

1. $\mathcal{O}_E \models C \sqsubseteq D$ iff $\mathcal{O}_L \models C \sqsubseteq D$ for all $\mathcal{C}(\Sigma)$ concepts C, D ;
2. if α is a role inclusion, role hierarchy, or functionality assertion that falls within \mathcal{L} and uses only symbols from Σ , then $\mathcal{O}_E \models \alpha$ iff $\mathcal{O}_L \models \alpha$.

We call \mathcal{O}_L a non-projective \mathcal{EL} approximation if $\text{sig}(\mathcal{O}_L) \subseteq \Sigma$ and projective otherwise.

For $\ell \in \mathbb{N} \cup \{\omega\}$, (non-projective and projective) ℓ -bounded \mathcal{L} approximations are defined analogously, except that only concepts C, D of depth bounded by ℓ are considered in Point 1.

The term ω -bounded approximation, which is in fact the same as an unbounded approximation, is used only for uniformity. In Definition 1 and throughout the paper, we use \mathcal{O}_E to denote ontologies formulated in an Expressive DL and \mathcal{O}_L to denote ontologies formulated in a Lightweight (in the sense of ‘inexpressive’) DL. Note that, trivially, infinite (non-projective and projective) approximations always exist: simply take as \mathcal{O}_L the set of all relevant inclusions and assertions that are entailed by \mathcal{O}_E . Definition 1 speaks about Horn-SRIF ontologies \mathcal{O}_E because this is the most expressive DL considered in this paper. In general, we speak about \mathcal{L}_S -to- \mathcal{L}_T approximation, \mathcal{L}_S a DL and \mathcal{L}_T a fragment thereof, to denote the task of approximating an \mathcal{L}_S ontology in \mathcal{L}_T . We call \mathcal{L}_S the *source DL* and \mathcal{L}_T the *target DL*. One can show that there are \mathcal{ELI} ontologies \mathcal{O}_E that have a finite projective \mathcal{EL} approximation, but no finite non-projective \mathcal{EL} approximation. Details are in the appendix. An alternative definition of approximations is obtained by considering in Point 1 *any* \mathcal{C} concept for C and D rather than only concepts in signature Σ . We choose not to do that because then even in the 1-bounded case, finite approximations do not necessarily exist.

Example 1. Consider the \mathcal{ELI} ontology $\mathcal{O}_E = \{\exists r^-.A \sqsubseteq B\}$. We have as a consequence $A \sqcap \exists r.X \sqsubseteq \exists r.(B \sqcap X)$ for each of the infinitely many concept names $X \in \mathbb{N}_C$. Thus, every (projective or non-projective) 1-bounded \mathcal{EL} approximation of \mathcal{O}_E must be infinite under the alternative definition of approximation.

We now make some basic observations regarding approximations. The proof is straightforward.

Lemma 1. Let \mathcal{O}_E be a Horn-SRIF ontology with $\text{sig}(\mathcal{O}_E) = \Sigma$ and \mathcal{L} a fragment of Horn-SRIF. Then

1. an \mathcal{L} ontology \mathcal{O}_L is an \mathcal{L} approximation of \mathcal{O}_E iff $\mathcal{O}_E \models \mathcal{O}_L$ and for every \mathcal{L} ontology \mathcal{O} with $\mathcal{O}_E \models \mathcal{O}$ and $\text{sig}(\mathcal{O}) \subseteq \Sigma$, $\mathcal{O}_L \models \mathcal{O}$;
2. $\bigcup_{i \geq 0} \mathcal{O}_i$ is an \mathcal{L} approximation of \mathcal{O}_E if for all $\ell \geq 0$, \mathcal{O}_ℓ is an ℓ -bounded \mathcal{L} approximation of \mathcal{O}_E ; the same is true for projective \mathcal{L} approximations provided that $\text{sig}(\mathcal{O}_\ell) \cap \text{sig}(\mathcal{O}_{\ell'}) \subseteq \Sigma$ when $\ell \neq \ell'$.

Point 1 may be viewed as an alternative definition of (non-projective) approximations. Point 2 is important because it allows us to concentrate on bounded approximations in proofs and to then obtain results for unbounded approximations as a byproduct. The following is well-known, see for example [5].

Lemma 2. *Given a Horn-SRIF ontology \mathcal{O}_E with $\text{sig}(\mathcal{O}) = \Sigma$, one can construct in polynomial time an \mathcal{ELRIF}_\perp ontology \mathcal{O}'_E with $\text{sig}(\mathcal{O}'_E) \supseteq \Sigma$ that entails the same Horn-SRIF(Σ) concept inclusions, role inclusions, and functionality assertions.*

Note that the construction of the ontology \mathcal{O}'_E from Lemma 2 requires the introduction of fresh concept names. Still, every ℓ -bounded \mathcal{L} approximation of \mathcal{O}' is a projective ℓ -bounded \mathcal{L} approximation of \mathcal{O} . From now on, we work with \mathcal{ELRIF}_\perp ontologies rather than with Horn-SRIF and thus obtain projective approximations also for the latter. Studying non-projective approximations of Horn-SRIF ontologies is outside the scope of this paper.

3 Unbounded \mathcal{ELRIF}_\perp -to- \mathcal{ELR}_\perp Approximation

We provide (unbounded) approximations of \mathcal{ELRIF}_\perp ontologies in \mathcal{ELR}_\perp . We assume throughout this section that \mathcal{ELRIF}_\perp ontologies \mathcal{O}_E are *inverse closed*, that is, for every role name r used in \mathcal{O}_E , there is a role name, which we denote \hat{r} , such that $r \sqsubseteq \hat{r}^-$ and $\hat{r} \sqsubseteq r^-$ are in \mathcal{O}_E . Thus, \hat{r} is an *explicit name* for the inverse of r . We can clearly additionally assume w.l.o.g. that there are no other occurrences of inverse roles in \mathcal{O}_E , which we shall always do. In other words, it suffices to consider inverse closed $\mathcal{ELR}^{\mathcal{I}}\mathcal{F}_\perp$ ontologies in place of inverse closed \mathcal{ELRIF}_\perp ontologies. We obtain non-projective approximations under this assumption, which clearly also yields projective approximations in the general case. The following theorem summarizes the results from this section.

Theorem 1. *Let \mathcal{O}_E be an inverse closed $\mathcal{ELR}^{\mathcal{I}}\mathcal{F}_\perp$ ontology and $\Sigma = \text{sig}(\mathcal{O}_E)$. Define \mathcal{O}_L to be the \mathcal{ELR}_\perp ontology that contains for all $\mathcal{EL}(\Sigma)$ concepts C, D and role names $r, s \in \Sigma$:*

1. all CIs in \mathcal{O}_E ;
2. $r \sqsubseteq s$ if $\mathcal{O}_E \models r \sqsubseteq s$;
3. $r_1 \circ \dots \circ r_n \sqsubseteq r, \hat{r}_n \circ \dots \circ \hat{r}_1 \sqsubseteq \hat{r}$ if $r_1 \circ \dots \circ r_n \sqsubseteq r \in \mathcal{O}_E$ with $n \geq 2$;
4. $C \sqcap \exists r.D \sqsubseteq \exists r.(D \sqcap \exists \hat{r}.C)$;
5. $\exists r.C \sqcap \exists r.D \sqsubseteq \exists r.(C \sqcap D)$ if $\text{func}(r) \in \mathcal{O}_E$;
6. $\exists r.\exists \hat{r}.C \sqsubseteq C$ if $\text{func}(\hat{r}) \in \mathcal{O}_E$.

Then \mathcal{O}_L is an \mathcal{ELR}_\perp approximation of \mathcal{O}_E .

Note that Points 1 to 3 essentially take over the part of \mathcal{O}_E that is expressible in \mathcal{ELR}_\perp , Point 4 aims at capturing the consequences of inverse roles, Point 5 at functional roles, and Point 6 at the interaction between functional roles and inverse roles. Points 4 to 6 all introduce infinitely many CIs. The following example shows that Point 6 cannot be omitted.

Example 2. Consider $\mathcal{O}_E = \{\text{func}(\hat{r}), r \sqsubseteq \hat{r}^-, \hat{r} \sqsubseteq r^-, A \sqsubseteq A\}$. Then $\mathcal{O}_E \models \exists r. \exists \hat{r}. A \sqsubseteq A$ but $\mathcal{O}'_L \not\models \exists r. \exists \hat{r}. A \sqsubseteq A$ for the ontology \mathcal{O}'_L obtained from \mathcal{O}_L by omitting the CIs of Point 6. To show this, consider the interpretation \mathcal{I} with domain $\{0, 1, \dots\}$ and

$$r^{\mathcal{I}} = \{(2n, 2n+1) \mid n \geq 0\}, \hat{r}^{\mathcal{I}} = \{(2n+1, 2n+2) \mid n \geq 0\}, A^{\mathcal{I}} = \{2n \mid n \geq 1\}$$

Then \mathcal{I} is a model of \mathcal{O}'_L but $0 \in (\exists r. \exists \hat{r}. A)^{\mathcal{I}} \setminus A^{\mathcal{I}}$.

It should be obvious how Point 5 captures the \mathcal{ELF} -to- \mathcal{EL} example from the introduction. Point 4 captures the natural variation of the \mathcal{ELI} -to- \mathcal{EL} example from the introduction obtained by converting the \mathcal{ELI} ontology \mathcal{O}_E used there into an inverse closed $\mathcal{ELH}^{\mathcal{I}}$ ontology, as follows.

Example 3. Let $\mathcal{O}_E = \{\exists \text{supervises}. \top \sqsubseteq \text{Manager}, \text{hasSuper} \sqsubseteq \text{supervises}^-, \text{supervises}^- \sqsubseteq \text{hasSuper}\}$. Point 4 yields

$$\exists \text{hasSuper}^n. \top \sqsubseteq \exists \text{hasSuper}. (\exists \text{supervises}. \top \sqcap \exists \text{hasSuper}^{n-1}. \top)$$

which together with $\exists \text{supervises}. \top \sqsubseteq \text{Manager} \in \mathcal{O}_L$ yields the desired

$$\exists \text{hasSuper}^n. \top \sqsubseteq \exists \text{hasSuper}. (\text{Manager} \sqcap \exists \text{hasSuper}^{n-1}. \top).$$

Theorem 1 also settles several natural subcases of (projective) $\mathcal{ELR}^{\mathcal{I}}\mathcal{F}_{\perp}$ -to- \mathcal{ELR}_{\perp} approximation such as $\mathcal{ELH}^{\mathcal{I}}$ -to- \mathcal{ELH} . For subcases where the source DL does not contain inverse roles such as \mathcal{ELF} -to- \mathcal{EL} , the concept inclusions in Point 4 are still present in the approximation as we still assume inverse closedness. This could also be avoided, as in the results presented in the subsequent section. We find it remarkable that the construction of \mathcal{O}_L is based almost entirely on a purely syntactic analysis of \mathcal{O}_E , rather than involving reasoning. Reasoning is only required to derive the role hierarchies to be included in \mathcal{O}_L , in Point 2 of Theorem 1. Although we do not consider this aspect very important, we mention that this problem is EXPTIME-complete when \mathcal{O}_E is formulated in $\mathcal{ELR}^{\mathcal{I}}\mathcal{F}_{\perp}$ and in PTIME for many of the captures subcases such as when \mathcal{O}_E is formulated in $\mathcal{ELR}^{\mathcal{I}}$.

It is straightforward to show that the ontology \mathcal{O}_L from Theorem 1 is sound as an approximation. To prove completeness, we establish a novel connection between \mathcal{EL}_{\perp} approximations and axiomatizations of the quasi-equations that are valid in classes of semilattices with operators (SLOs) [16, 29, 17]. Roughly speaking, an approximation is obtained from such an axiomatization by instantiating its equations, which correspond (in the sense of modal correspondence theory) to the role inclusions and hierarchies in the original ontology, with arbitrary \mathcal{EL} concepts. A detailed presentation is provided in the appendix.

4 Depth-Bounded $\mathcal{ELH}^{\mathcal{I}}\mathcal{F}_{\perp}$ -to- \mathcal{ELH}_{\perp} Approximation

We pursue an alternative approach to proving that an approximation is complete, based on a suitable version of the chase. This allows us to also treat the

case of depth bounded approximations. Moreover, approximations obtained in this section are non-projective and the assumption of ontologies being inverse closed is not needed. We consider $\mathcal{ELH}^{\perp}\mathcal{F}_{\perp}$ -to- \mathcal{ELH}_{\perp} approximation and subcases thereof. An extension to role inclusions should be possible, but is left as future work.

We assume w.l.o.g. that role hierarchies only take the two forms $r \sqsubseteq s$ and $r \sqsubseteq s^{-}$. We further assume the following syntactic restriction:

(\heartsuit) $\mathcal{O}_E \models r \sqsubseteq s^{-}$ implies that neither $\text{func}(s) \in \mathcal{O}_E$ nor $\text{func}(s^{-}) \in \mathcal{O}_E$.

This assumption is not without loss of generality, it serves to eliminate a subtle interaction between inverse roles, functional roles, and role hierarchies. While this interaction is captured implicitly and gracefully by the unbounded approximation scheme given in the previous section, it causes complications in the depth bounded case. We briefly comment on this at the end of the section.

Let C be an \mathcal{EL}_{\perp} concept and $k \geq 0$. By *decorating C with subconcepts from \mathcal{O}_E at leaves*, we mean to replace any number of occurrences of a quantifier-free subconcept D by a concept $D \sqcap D_1 \sqcap \dots \sqcap D_k$, D_1, \dots, D_k subconcepts of \mathcal{O}_E . The following is the main result of this section.

Theorem 2. *Let \mathcal{O}_E be an $\mathcal{ELFH}_{\perp}^{\perp}$ ontology, $\Sigma = \text{sig}(\mathcal{O}_E)$, and $\ell \in \mathbb{N} \cup \{\omega\}$ a depth bound. Define \mathcal{O}_L to be the \mathcal{ELH}_{\perp} ontology that consists of the following, where $\ell' = \max\{0, \ell - 1\}$ and r, r_1, r_2, s are role names from Σ :*

1. all concept inclusions from \mathcal{O}_E ;
2. $r \sqsubseteq s$ if $\mathcal{O}_E \models r \sqsubseteq s$;
3. $C_1 \sqcap \exists r.C_2 \sqsubseteq \exists r.(C_2 \sqcap \exists s.C_1)$ if $\mathcal{O}_E \models r \sqsubseteq s^{-}$, $\exists s.C_1$ is a subconcept of \mathcal{O}_E or an $\mathcal{EL}(\Sigma)$ concept of depth bounded by ℓ , and C_2 is an $\mathcal{EL}(\Sigma)$ concept of depth bounded by ℓ' decorated with subconcepts of \mathcal{O}_E at leaves;
4. $\exists r_1.C_1 \sqcap \exists r_2.C_2 \sqsubseteq \exists r_1.(C_1 \sqcap C_2)$ if there is a role name s with $\mathcal{O}_E \models r_1 \sqsubseteq s$, $\mathcal{O}_E \models r_2 \sqsubseteq s$, and $\mathcal{O}_E \ni \text{func}(s)$, and C_1, C_2 are $\mathcal{EL}(\Sigma)$ concepts of depth bounded by ℓ' decorated with subconcepts of \mathcal{O}_E at leaves.

Then \mathcal{O}_L is an ℓ -bounded approximation of \mathcal{O}_E .

We tried to be as economic as possible regarding the classes of concepts that have to be considered in Points 3 and 4, which has led to the subtle depth bounds stated there. Note that Theorem 2 also settles the cases of \mathcal{ELFH}_{\perp} -to- \mathcal{ELH}_{\perp} approximation and of $\mathcal{ELH}_{\perp}^{\perp}$ -to- \mathcal{ELH}_{\perp} approximation (both without any syntactic restrictions), as well as the variation of all these cases without \mathcal{H} and/or \perp in both the source and target DL.

Due to Points 3 and 4, the approximation \mathcal{O}_L is of (single) exponential size even when $\ell = 0$. This must necessarily be the case because otherwise we would obtain a subexponential algorithm for the EXPTIME-complete subsumption problem between concept names in \mathcal{ELI} . Note that Theorem 2 also improves the upper bound for this problem: compute the 0-bounded approximation of single exponential size in exponential time and then decide subsumption in \mathcal{EL} in PTIME.

It is again straightforward to verify that the ontology \mathcal{O}_L constructed in Theorem 2 is sound as an approximation, that is, $\mathcal{O}_E \models \mathcal{O}_L$. Completeness is established in two steps. First, we introduce a suitable version of the chase and show that it is sound and complete regarding the consequences relevant for approximation, and second we show that the CIs in \mathcal{O}_L can simulate derivations of this chase.

Let us now drop assumption (\heartsuit) . One can prove that we then need to extend Points 1 to 4 of Theorem 2 with the following:

5. $\exists r.\exists s.C \sqsubseteq C$ if $\mathcal{O}_E \models r \sqsubseteq s^-$, $\mathbf{func}(s) \in \mathcal{O}_E$, and C is a subconcept of \mathcal{O}_E or an $\mathcal{EL}(\Sigma)$ concept of depth bounded by ℓ ;
6. $\exists s.\exists r.C \sqsubseteq C$ if $\mathcal{O}_E \models r \sqsubseteq s^-$, $\mathbf{func}(s^-) \in \mathcal{O}_E$, and C is a subconcept of \mathcal{O}_E or an $\mathcal{EL}(\Sigma)$ concept of depth bounded by ℓ .

However, this is still not sufficient to obtain a complete approximation. Consider the \mathcal{ELH}^I ontology

$$\mathcal{O}_E = \{A \sqsubseteq \exists r_1.\exists r_2.(B \sqcap \exists s.\top), \quad s \sqsubseteq r_1, \quad s \sqsubseteq r_2^-, \quad \mathbf{func}(r_1^-), \quad \mathbf{func}(r_2^-)\}.$$

It can be verified that $\mathcal{O}_E \models A \sqsubseteq B$. However, it can also be proved that even when \mathcal{O}_L is the set of all statements from Points 1 to 6 with $\ell = 0$, $\mathcal{O}_L \not\models A \sqsubseteq B$.

5 \mathcal{ELI}_\perp -to- \mathcal{EL}_\perp Approximation

The previous sections concentrated on cases of \mathcal{L} -to- \mathcal{L}' approximations where both \mathcal{L} and \mathcal{L}' were based on the same concept language \mathcal{EL}_\perp . In this section, we take a look at \mathcal{ELI}_\perp -to- \mathcal{EL}_\perp approximation, thus aiming to approximate away inverse roles in concepts inclusions. We consider bounded and unbounded, projective and non-projective approximations. The proof techniques is based on the chase, as in the previous section.

Constructing informative non-projective approximations appears to be difficult in the \mathcal{ELI}_\perp -to- \mathcal{EL}_\perp case. Informally, concepts of the form $\exists r^-.C$ can be used as a marker invisible to \mathcal{EL}_\perp that is propagated along role edges, resulting in rather complex \mathcal{EL} concept inclusions to be entailed by \mathcal{O}_E .

Example 4. Let $\mathcal{O}_E = \{A \sqsubseteq \exists s^-. \top, \exists r^-. \exists s^-. \top \sqsubseteq \exists s^-. \top, \exists s^-. \top \sqsubseteq B\}$. Then $\mathcal{O}_E \models C \sqsubseteq C'$ for all \mathcal{EL} concepts C, C' where C' is obtained from C by decorating with B any node that is reachable in C from a node decorated with A along an r -path (we view an \mathcal{EL} concept as a tree in the standard way, see for example [18]).

We now give a non-projective approximation that captures the effects demonstrated in Example 4. For an \mathcal{ELI}_\perp ontology \mathcal{O}_E , we use $\mathbf{cl}_{\mathcal{EL}}(\mathcal{O}_E)$ to denote the set of all \mathcal{EL} concepts that can be obtained by starting with a subconcept of a concept from \mathcal{O}_E and then replacing every subconcept of the form $\exists r^-.D$ with \top . Let C be an \mathcal{EL} concept. An \mathcal{EL} concept C' is a $\mathbf{cl}_{\mathcal{EL}}(\mathcal{O}_E)$ decoration of C if it can be obtained from C by conjunctively adding concepts from $\mathbf{cl}_{\mathcal{EL}}(\mathcal{O}_E)$ to a single occurrence of a subconcept in C .

Theorem 3. Let \mathcal{O}_E be an $\mathcal{EL}\mathcal{I}_\perp$ ontology and $\Sigma = \text{sig}(\mathcal{O}_E)$. Define an \mathcal{EL}_\perp ontology \mathcal{O}_L that consists of the following:

1. $C \sqsubseteq C'$ if $\mathcal{O}_E \models C \sqsubseteq C'$, C an $\mathcal{EL}(\Sigma)$ concept and C' a $\text{cl}_{\mathcal{EL}}(\mathcal{O}_E)$ decoration of C ;
2. $C \sqsubseteq \perp$ if $\mathcal{O}_E \models C \sqsubseteq \perp$, C an $\mathcal{EL}(\Sigma)$ concept;

Then \mathcal{O}_L is an approximation of \mathcal{O}_E .

We prove completeness by using the standard chase for $\mathcal{EL}\mathcal{I}_\perp$ that is complete also for $\mathcal{EL}\mathcal{I}$ consequences and then work with homomorphisms that are in a certain sense ‘forwards directed’. Details are given in the appendix.

Arguably, the approximation provided by Theorem 3 is less informative than the ones obtained in the previous sections. We next demonstrate that in the projective case, more informative approximations can be constructed. One way of doing this is to first convert the $\mathcal{EL}\mathcal{I}_\perp$ ontology into an inverse closed $\mathcal{EL}\mathcal{H}_\perp^{\mathcal{I}}$ ontology as in Section 3. Here, we pursue a natural alternative that consists of first converting the $\mathcal{EL}\mathcal{I}_\perp$ ontology into a widely known normal form for such ontologies [2] and then providing non-projective approximations for ontologies in normal form. Note that, in practice, ontologies are sometimes already constructed in this normal form or at least in a form very close to it.

An $\mathcal{EL}\mathcal{I}_\perp$ ontology \mathcal{O} is in *normal form* if all CIs in \mathcal{O} have one of the forms $\top \sqsubseteq A_1$, $A_1 \sqsubseteq \perp$, $A_1 \sqsubseteq \exists\rho.A_2$, $\exists\rho.A_1 \sqsubseteq B$, and $A_1 \sqcap \dots \sqcap A_n \sqsubseteq B$ where A_1, \dots, A_n, B range over concept names and ρ ranges over roles. It is well known that every $\mathcal{EL}\mathcal{I}_\perp$ ontology \mathcal{O} can be converted in an $\mathcal{EL}\mathcal{I}_\perp$ ontology \mathcal{O}' in linear time such that \mathcal{O}' is in normal form and a conservative extension of \mathcal{O} . Clearly, any (projective or non-projective) approximation of \mathcal{O}' is a projective approximation of \mathcal{O} .

Theorem 4. Let \mathcal{O}_E be an $\mathcal{EL}\mathcal{I}_\perp$ ontology in normal form, $\Sigma = \text{sig}(\mathcal{O}_E)$, and $\ell \in \mathbb{N} \cup \{\omega\}$ a depth bound. Define an \mathcal{EL}_\perp ontology \mathcal{O}_L that consists of the following:

1. all concept inclusions from \mathcal{O}_E that are of the form $\top \sqsubseteq A$, $A \sqsubseteq \perp$, $\exists r.A \sqsubseteq B$, or $A \sqsubseteq \exists r.B$;
2. $A_1 \sqcap \dots \sqcap A_n \sqsubseteq B$ if $\mathcal{O}_E \models A_1 \sqcap \dots \sqcap A_n \sqsubseteq B$, $A_1, \dots, A_n, B \in \mathbb{N}_C$ in \mathcal{O}_E ;
3. $A \sqcap \exists r.C \sqsubseteq \exists r.(C \sqcap B)$ if $\exists r^-.A \sqsubseteq B \in \mathcal{O}_E$ and C is an $\mathcal{EL}(\Sigma)$ concept of depth bounded by $\ell - 1$.

Then \mathcal{O}_L is an approximation of \mathcal{O}_E .

It is straightforward to verify that \mathcal{O}_L is sound. To prove completeness, we use the same strategy as for Theorem 2.

6 Size of Approximations

We prove that finite approximations do not necessarily exist and that depth bounded approximations can be non-elementary in size. These results hold both

for projective and non-projective approximations and for all combinations of source and target DL considered in this paper. The ontologies used to prove these results are simple and show that for the vast majority of ontologies that occur in practical applications, neither finite approximations nor depth bounded approximations of elementary size can be expected. We focus on the cases \mathcal{ELIH} -to- \mathcal{ELH} , \mathcal{ELHF} -to- \mathcal{ELH} , and \mathcal{ELH}^\perp -to- \mathcal{ELH} , starting with the non-existence of finite approximations.

Theorem 5. *None of the ontologies*

$$\{\exists r^- . A \sqsubseteq B\}, \quad \{\text{func}(r), A \sqsubseteq A\}, \quad \{r \sqsubseteq s^-, A \sqsubseteq A\}$$

has finite projective \mathcal{ELH} approximations.

We next show that bounded depth approximations can be non-elementary in size. The function $\text{tower} : \mathbb{N} \times \mathbb{N} \rightarrow \mathbb{N}$ is defined as $\text{tower}(0, n) := n$ and $\text{tower}(k + 1, n) := 2^{\text{tower}(k, n)}$. The *size* of a (finite) ontology is the number of symbols needed to write it, with concept and role names counting as one. We use Γ_n to denote a fixed finite tautological set of \mathcal{EL} concept inclusions that contains the symbols $\Sigma_n = \{r_1, r_2, A_1, \hat{A}_1, \dots, A_n, \hat{A}_n\}$. One could take, for example, $\exists r_1. \top \sqsubseteq \top$, $\exists r_2. \top \sqsubseteq \top$, and all $A \sqsubseteq A$ with $A \in \{A_1, \hat{A}_1, \dots, A_n, \hat{A}_n\}$.

Theorem 6. *Let $n \geq 0$ and let \mathcal{O}_n be the union of Γ_n with any of the following sets:*

$$\{\exists r^- . A \sqsubseteq B\}, \quad \{\text{func}(r), A \sqsubseteq A\}, \quad \{r \sqsubseteq s^-, A \sqsubseteq A\}$$

For every $\ell \geq 1$, any ℓ -bounded projective \mathcal{ELH} approximation \mathcal{O}_L of \mathcal{O}_n must be of size at least $\text{tower}(\ell, n)$.

7 Conclusion

There are several questions that emerge from our work. For example, it remains an open problem to develop a convincing approximation for $\mathcal{ELH}^\perp \mathcal{F}_\perp$ -to- \mathcal{ELH}_\perp in the non-projective case, or even for Horn- \mathcal{SRLIF} -to- \mathcal{ELR}_\perp . It would also be useful to consider more expressive target DLs such as the extension of \mathcal{ELH}_\perp or \mathcal{ELR}_\perp with range restrictions, and to add nominals to the picture. Of course, it would also be very interesting to approximate non-Horn DLs such as \mathcal{ALC} , \mathcal{SHIQ} , and \mathcal{SROIQ} in (tractable and intractable) Horn DLs.

From a conceptual perspective, it would be of great interest to understand how approximations can be tailored towards intended applications. In this context, observe that all our bounded depth approximation schemes still work when the set of concepts of depth ℓ is replaced with any set Γ of concepts closed under subconcepts. For example, if one wants to decide subsumption between \mathcal{EL} concepts C and D relative to an \mathcal{ELI} ontology \mathcal{O}_E , one can approximate \mathcal{O}_E in \mathcal{EL} relative to the set Γ of subconcepts of C and D ; the resulting \mathcal{EL} ontology \mathcal{O}_L will entail $C \sqsubseteq D$ iff \mathcal{O}_E does. In a similar spirit, it would be interesting to develop approximations that aim at query answering applications.

Acknowledgments. Supported by the DFG Collaborative Research Center 1320 EASE - Everyday Activity Science and Engineering.

References

1. Baader, F., Brandt, S., Lutz, C.: Pushing the \mathcal{EL} envelope. In: Proc. of IJCAI. Morgan Kaufmann (2005)
2. Baader, F., Horrocks, I., Lutz, C., Sattler, U.: An Introduction to Description Logic. Cambridge University Press (2017)
3. Beklemishev, L.D.: Positive provability logic for uniform reflection principles. *Ann. Pure Appl. Logic* **165**(1), 82–105 (2014)
4. van Benthem, J.: Correspondence theory. In: Gabbay, D., Guenther, F. (eds.) *Handbook of Philosophical Logic*, vol. 2, pp. 167–247. Reidel, Dordrecht (1984)
5. Bienvenu, M., Hansen, P., Lutz, C., Wolter, F.: First order-rewritability and containment of conjunctive queries in horn description logics. In: Proceedings of IJCAI. pp. 965–971 (2016)
6. Blackburn, P., de Rijke, M., Venema, Y.: *Modal Logic*. Cambridge University Press (2001)
7. Botoeva, E., Calvanese, D., Rodriguez-Muro, M.: Expressive approximations in *DL-Lite* ontologies. In: AIMS. pp. 21–31 (2010)
8. Botoeva, E., Calvanese, D., Santarelli, V., Savo, D.F., Solimando, A., Xiao, G.: Beyond OWL 2 QL in OBDA: rewritings and approximations. In: Proc. of AAAI. pp. 921–928. AAAI Press (2016)
9. Brandt, S., Küsters, R., Turhan, A.Y.: Approximation and difference in description logics. In: KR. pp. 203–214 (2002)
10. Carral, D., Feier, C., Grau, B.C., Hitzler, P., Horrocks, I.: \mathcal{EL} -ifying ontologies. In: Proc. of IJCAR. pp. 464–479 (2014)
11. Darwiche, A., Marquis, P.: A knowledge compilation map. *J. Artif. Intell. Res.* **17**, 229–264 (2002)
12. Feier, C., Lutz, C., Kuusisto, A.: Rewritability in monadic disjunctive datalog, MMSNP, and expressive description logics. In: Proceedings of the 20th International Conference on Database Theory (ICDT17) (2017)
13. Goldblatt, R.: Varieties of complex algebras. *Annals of Pure and Applied Logic* **44**, 173–242 (1989)
14. Groot, P., Stuckenschmidt, H., Wache, H.: Approximating description logic classification for semantic web reasoning. In: Proc. of ESWC. pp. 318–332 (2005)
15. Hernich, A., Lutz, C., Papacchini, F., Wolter, F.: Horn-rewritability vs PTime query evaluation in ontology-mediated querying. In: Proceedings of IJCAI-ECAI. AAAI Press (2018)
16. Jackson, M.: Semilattices with closure. *Algebra Universalis* **52**, 1–37 (2004)
17. Kikot, S., Kurucz, Á., Tanaka, Y., Wolter, F., Zakharyashev, M.: Kripke completeness of strictly positive modal logics over meet-semilattices with operators. *CoRR* **abs/1708.03403** (2017), <http://arxiv.org/abs/1708.03403>
18. Konev, B., Lutz, C., Ozaki, A., Wolter, F.: Exact learning of lightweight description logic ontologies. *Journal of Machine Learning Research* **18**(201), 1–63 (2018)
19. Konev, B., Lutz, C., Wolter, F., Zakharyashev, M.: Conservative rewritability of description logic TBoxes. In: Proc. of IJCAI. pp. 1153–1159 (2016)
20. Krötzsch, M., Rudolph, S., Hitzler, P.: Complexities of horn description logics. *ACM Trans. Comput. Log.* **14**(1), 2:1–2:36 (2013)
21. Lutz, C., Piro, R., Wolter, F.: Description logic TBoxes: Model-theoretic characterizations and rewritability. In: IJCAI. pp. 983–988 (2011)
22. Lutz, C., Seylan, I., Wolter, F.: An automata-theoretic approach to uniform interpolation and approximation in the description logic \mathcal{EL} . In: Proc. of KR. AAAI Press (2012)

23. Lutz, C., Wolter, F.: Deciding inseparability and conservative extensions in the description logic \mathcal{EL} . *J. Symb. Comput.* **45**(2), 194–228 (2010)
24. Pan, J.Z., Thomas, E.: Approximating *OWL-DL* ontologies. In: *AAAI*. pp. 1434–1439 (2007)
25. Ren, Y., Pan, J.Z., Zhao, Y.: Soundness preserving approximation for TBox reasoning. In: *AAAI* (2010)
26. Schaerf, M., Cadoli, M.: Tractable reasoning via approximation. *Artificial Intelligence* **74**(2), 249 – 310 (1995)
27. Selman, B., Kautz, H.A.: Knowledge compilation and theory approximation. *J. ACM* **43**(2), 193–224 (1996)
28. Sofronie-Stokkermans, V.: Locality and subsumption testing in \mathcal{EL} and some of its extensions. In: Areces, C., Goldblatt, R. (eds.) *Advances in Modal Logic*, vol. 7. pp. 315–339. College Publications (2008)
29. Sofronie-Stokkermans, V.: Representation theorems and locality for subsumption testing and interpolation in the description logics \mathcal{EL} and \mathcal{EL}^+ and their extensions with n -ary roles and numerical domains. *Fundam. Inform.* **156**(3-4), 361–411 (2017)
30. Tserendorj, T., Rudolph, S., Krötzsch, M., Hitzler, P.: Approximate owl-reasoning with screech. In: *Proc. of RR. LNCS*, vol. 5341, pp. 165–180. Springer (2008)
31. Zhou, Y., Grau, B.C., Nenov, Y., Kaminski, M., Horrocks, I.: Pagoda: Pay-as-you-go ontology query answering using a datalog reasoner. *J. Artif. Intell. Res.* **54**, 309–367 (2015)