A Journey into Ontology Approximation: From Non-Horn to Horn

Anneke Haga[1], **Carsten Lutz**[1], **Johannes Marti**[2] and **Frank Wolter**[3]
[1]Universität Bremen, Germany
[2]Universiteit van Amsterdam, Netherlands
[3]University of Liverpool, UK
{anneke, clu}@uni-bremen.de, johannes.marti@gmail.com, wolter@liverpool.ac.uk

Abstract

We study complete approximations of an ontology formulated in a non-Horn description logic (DL) such as $\mathcal{ALC}$ in a Horn DL such as $\mathcal{EL}$. We provide concrete approximation schemes that are necessarily infinite and observe that in the $\mathcal{ELU}$-to-$\mathcal{EL}$ case finite approximations tend to exist in practice and are guaranteed to exist when the source ontology is acyclic. In contrast, neither of these are the case for $\mathcal{ELU}_\bot$-to-$\mathcal{EL}_\bot$ and for $\mathcal{ALC}$-to-$\mathcal{EL}_\bot$ approximations. We also define a notion of approximation tailored towards ontology-mediated querying, connect it to subsumption-based approximations, and identify a case where finite approximations are guaranteed to exist.

1 Introduction

Despite prominent standardization efforts such as OWL, a large variety of description logics (DLs) continues to be used as ontology languages. In fact, ontology designers choose a DL suitable for their purposes based on many factors including expressive power, computational properties, and tool support [Baader *et al.*, 2017]. Since ontology engineering frequently involves (partial) reuse of existing ontologies, this raises the problem of converting an ontology written in some source DL $\mathcal{L}_S$ into a desired target DL $\mathcal{L}_T$. A particularly important case is ontology approximation where $\mathcal{L}_T$ is a fragment of $\mathcal{L}_S$, studied for example in [Pan and Thomas, 2007; Ren *et al.*, 2010; Botoeva *et al.*, 2010; Carral *et al.*, 2014; Zhou *et al.*, 2015; Bötcher *et al.*, 2019].

In practice, ontology approximation is often done in an ad hoc way by dropping all statements from the source ontology $\mathcal{O}_S$ that are not expressible in $\mathcal{L}_T$, or at least the inexpressible parts of such statements. It is well-known that this results in incomplete approximations, that is, there will be knowledge in $\mathcal{O}_S$ that could be expressed in $\mathcal{L}_T$, but is not contained in the resulting approximated ontology. The degree and nature of the resulting incompleteness is typically neither understood nor analyzed. One reason for this unsatisfactory situation might be the fact that it is by no means easy to construct complete approximations and, even worse, finite complete approximations are not guaranteed to exist. This was studied in depth in [Bötcher *et al.*, 2019] where ontologies formulated in expressive Horn DLs such as Horn-$\mathcal{SHIF}$ and $\mathcal{ELI}$ are approximated in tractable Horn DLs such as $\mathcal{EL}$. For example, it is shown there that finite complete $\mathcal{ELI}$-to-$\mathcal{EL}$ approximations do not exist even in extremely simple cases including those occurring in practice. The authors then lay out a new research program for ontology approximation that consists in mapping out the structure of complete (infinite) ontology approximations as a tool for guiding informed decisions when constructing incomplete (finite) approximations in practice, and also to enable a better understanding of the degree and nature of incompleteness.

In this paper, we consider $\mathcal{L}_S$-to-$\mathcal{L}_T$ ontology approximation where $\mathcal{L}_S$ is a non-Horn DL such as $\mathcal{ALC}$ and $\mathcal{L}_T$ is a tractable Horn DL such as $\mathcal{EL}$. Arguably, these are extremely natural cases of ontology approximation given that Horn vs. non-Horn is nowadays the most important classification criterion for DLs [Baader *et al.*, 2017]. Non-Horn DLs include expressive features such as negation and disjunction and require 'reasoning by cases' which is computationally costly, but also have considerably higher expressive power than Horn DLs. Horn DLs, in contrast, enjoy favourable properties such as the existence of universal models and of 'consequence-based' reasoning algorithms that avoid reasoning by cases [Cucala *et al.*, 2019]. Despite being natural, however, non-Horn-to-Horn approximation turns out to be a challenging endeavour.

We start with the fundamental case of $\mathcal{ELU}$-to-$\mathcal{EL}$ approximation. Given an $\mathcal{ELU}$ ontology $\mathcal{O}_S$, we aim to find a (potentially infinite) $\mathcal{EL}$ ontology $\mathcal{O}_T$ such that for all $\mathcal{EL}$ concepts C, D in the signature of $\mathcal{O}_S$, $\mathcal{O}_S \models C \sqsubseteq D$ iff $\mathcal{O}_T \models C \sqsubseteq D$.

Example 1. *Consider the $\mathcal{ELU}$ ontology*

$$\mathcal{O}_S = \{ \quad \mathtt{Job} \sqsubseteq \mathtt{MainJob} \sqcup \mathtt{SideJob} \quad \exists\mathtt{job.SideJob} \sqsubseteq \exists\mathtt{job.}(\mathtt{MainJob} \sqcap \mathtt{PartTime}) \ \}.$$

Then the following is an $\mathcal{EL}$ approximation of $\mathcal{O}_S$:

$$\mathcal{O}_T = \{ \quad \exists\mathtt{job.SideJob} \sqsubseteq \exists\mathtt{job.}(\mathtt{MainJob} \sqcap \mathtt{PartTime}) \quad \exists\mathtt{job.Job} \sqsubseteq \exists\mathtt{job.MainJob} \quad \exists\mathtt{job.}(\mathtt{Job} \sqcap \mathtt{PartTime}) \sqsubseteq \exists\mathtt{job.}(\mathtt{MainJob} \sqcap \mathtt{PartTime}) \ \}.$$

The last two lines of $\mathcal{O}_T$ illustrate that $\mathcal{EL}$ consequences of $\mathcal{ELU}$ ontologies can be rather non-obvious.

We first prove that finite approximations need not exist in the $\mathcal{ELU}$-to-$\mathcal{EL}$ case and that depth bounded approximations may be non-elementary in size. Our main result is then a concrete approximation scheme that makes explicit the structure of complete infinite approximations and aims to keep as much structure of the source ontology as possible. An interesting and, given the results in [Bötcher *et al.*, 2019], surprising feature of our scheme is that it can be expected to often deliver *finite* approximations in practical cases. We perform a case study based on the Manchester ontology corpus that confirm this expectation. We also show that if $\mathcal{O}_S$ is an acyclic $\mathcal{ELU}$ ontology, then a finite $\mathcal{EL}$ approximation always exists (though it need not be acyclic). The finite approximations that we obtain are too large to be directly used in practice. Nevertheless, we view our results as positive and believe that in practice approximations of reasonable size often exist, as in Example 1. A 'push button technology' for constructing them, however, is outside of the scope of this paper.

We then proceed to the cases of $\mathcal{ELU}_\bot$-to-$\mathcal{EL}_\bot$ and $\mathcal{ALC}$-to-$\mathcal{EL}_\bot$ approximations which turn out to be closely related to each other. They also turn out to be significantly different from the $\mathcal{ELU}$-to-$\mathcal{EL}$ case in that finite approximations do not exist in extremely simple (and practical) cases, much like in the Horn approximation cases studied in [Bötcher *et al.*, 2019]. Also, finite approximations of acyclic ontologies are no longer guaranteed to exist. While this is not good news, it is remarkable that the addition of the $\bot$ symbol has such a dramatic effect. We again provide an (infinite) approximation scheme.

Finally, we propose a notion of approximation that is tailored towards applications in ontology-mediated querying [Calvanese *et al.*, 2009] and show that it is intimately related to the subsumption-based approximations that we had studied before. Remarkably, if we concentrate on atomic queries (AQs), then we obtain finite approximations even in the $\mathcal{ALC}$-to-$\mathcal{EL}_\bot$ case. Compared to the related work presented in [Kaminski *et al.*, 2016], we do not require the preservation of all query answers, but only of a maximal subset thereof, and our method is applicable to all ontologies formulated in the source DL chosen rather than to a syntactically restricted class. We also observe an interesting application to the rewritability of ontology-mediated queries.

All proofs are deferred to the appendix [Haga *et al.*, 2020].

2 Preliminaries

Let $\mathsf{N_C}$ and $\mathsf{N_R}$ be disjoint and countably infinite sets of *concept names* and *role names*. In the description logic $\mathcal{ALC}$, *concepts* C, D are built according to the syntax rule

$$C, D ::= \top \mid \bot \mid A \mid \neg C \mid C \sqcap D \mid C \sqcup D \mid \exists r.C \mid \forall r.C$$

where A ranges over $\mathsf{N_C}$ and r over $\mathsf{N_R}$. The *depth* of a concept is the nesting depth of the constructors $\exists r$ and $\forall r$ in it. For example, the concept $\exists r.B \sqcap \exists r.\exists s.A$ is of depth 2. We introduce other DLs as fragments of $\mathcal{ALC}$. An $\mathcal{ELU}_\bot$ *concept* is an $\mathcal{ALC}$ concept that does not contain negations $\neg C$ and value restrictions $\forall r.C$. An $\mathcal{EL}_\bot$ *concept* is an $\mathcal{ELU}_\bot$ concept that does not contain disjunctions $C \sqcup D$. $\mathcal{ELU}$ *concepts* and $\mathcal{EL}$ *concepts* are defined likewise, but additionally forbid the use of the bottom concept $\bot$.

For any of these DLs $\mathcal{L}$, an $\mathcal{L}$ *ontology* is a set of *concept inclusions (CIs)* $C \sqsubseteq D$ where C and D are $\mathcal{L}$ concepts. While ontologies used in practice have to be finite, we frequently consider also infinite ontologies. W.l.o.g., we assume that all occurrences of $\bot$ in $\mathcal{ELU}_\bot$ ontologies are in CIs of the form $C \sqsubseteq \bot$, where C does not contain $\bot$. An *acyclic ontology* $\mathcal{O}$ is a set of concept inclusions $A \sqsubseteq C$ and *concept equivalences* $A \equiv C$ where A is a concept name (that is, it is not a compound concept), the left-hand sides are unique, and $\mathcal{O}$ does not contain a definitiorial cycle $A_0 \bowtie_1 C_0, \dots, A_n \bowtie_n C_n, \bowtie_i \in \{\sqsubseteq, \equiv\}$, where C_i contains $A_{i+1 \,\mathtt{mod}\, n+1}$ for all $i \leq n$. An equivalence $A \equiv C$ can be viewed as two CIs $A \sqsubseteq C$ and $C \sqsubseteq A$ and thus every acyclic ontology is an ontology in the original sense.

A *signature* Σ is a set of concept and role names, uniformly referred to as *symbols*. We use $\mathtt{sig}(X)$ to denote the set of symbols used in any syntactic object X such as a concept or an ontology. If $\mathtt{sig}(X) \subseteq \Sigma$, we also say that X is *over* Σ. The *size* of a (finite) syntactic object X, denoted $||X||$, is the number of symbols needed to write it, with every occurrence of a concept and role name contributing one.

The semantics of concepts and ontologies is defined in terms of *interpretations* $\mathcal{I} = (\Delta^\mathcal{I}, \cdot^\mathcal{I})$ as usual, see [Baader *et al.*, 2017]. An interpretation $\mathcal{I}$ *satisfies* a CI $C \sqsubseteq D$ if $C^\mathcal{I} \subseteq D^\mathcal{I}$, an equivalence $A \equiv C$ if $A^\mathcal{I} = C^\mathcal{I}$, and it is a *model* of an ontology $\mathcal{O}$ if it satisfies all CIs in $\mathcal{O}$. Concept C is *subsumed by* concept D w.r.t. ontology $\mathcal{O}$, written $\mathcal{O} \models C \sqsubseteq D$, if every model $\mathcal{I}$ of $\mathcal{O}$ satisfies the CI $C \sqsubseteq D$; we then also say that the CI is a *consequence* of $\mathcal{O}$. Subsumption can be decided in polynomial time in $\mathcal{EL}_\bot$ and is ExpTime-complete between $\mathcal{ELU}$ and $\mathcal{ALC}$ [Baader *et al.*, 2017]. We now give our main definition of approximation. With concept of depth bounded by ω, we mean concepts of unrestricted depth.

Definition 1. *Let $\mathcal{O}_S$ be an $\mathcal{ALC}$ ontology, sig$(\mathcal{O}_S) = \Sigma$, $\mathcal{L}_T$ any of the DLs introduced above, and $\ell \in \mathbb{N} \cup \{\omega\}$. A (potentially infinite) $\mathcal{L}_T$ ontology $\mathcal{O}_T$ is an* ℓ-bounded $\mathcal{L}_T$ approximation *of $\mathcal{O}_S$ if*

$$\mathcal{O}_S \models C \sqsubseteq D \textit{ iff } \mathcal{O}_T \models C \sqsubseteq D$$

for all $\mathcal{L}_T$ concepts C, D over Σ of depth bounded by ℓ. $\mathcal{O}_T$ is non-projective *if $\mathtt{sig}(\mathcal{O}_T) \subseteq \Sigma$ and* projective *otherwise. We refer to ω-bounded $\mathcal{L}_T$ approximations as $\mathcal{L}_T$* approximations.

We refer to the "if" direction of the biimplication in Definition 1 as *soundness* of the approximation and to the "only if" direction as *completeness*. Infinite approximations always exist: take as $\mathcal{O}_T$ the set of all $\mathcal{L}$ CIs $C \sqsubseteq D$ with C, D over Σ and $\mathcal{O}_S \models C \sqsubseteq D$. In the same way, finite (non-projective) depth-bounded approximations always exist. With *$\mathcal{L}_S$-to-$\mathcal{L}_T$ approximation*, $\mathcal{L}_S$ a DL and $\mathcal{L}_T$ a fragment of $\mathcal{L}_S$, we mean the task to approximate an $\mathcal{L}_S$ ontology in $\mathcal{L}_T$, possibly using an infinite ontology.

3 $\mathcal{ELU}$-to-$\mathcal{EL}$ Approximation

We consider $\mathcal{ELU}$-to-$\mathcal{EL}$ approximation as the simplest case of approximating non-Horn ontologies in a Horn DL.

Fundamentals. We start with observing that projective approximations are more powerful than non-projective ones.

Proposition 1. *The $\mathcal{ELU}$ ontology*

$$\begin{array}{llll} \mathcal{O}_S = \{ & A & \sqsubseteq & B_1 \sqcup B_2, \\ & \exists r.B_i & \sqsubseteq & B_i, \quad \textit{for } i \in \{1,2\} \\ & B_i \sqcap A' & \sqsubseteq & M \quad \textit{for } i \in \{1,2\} \,\}. \end{array}$$

$$
\begin{array}{rcll}
C & \sqsubseteq & X_C & \\
X_{D_1} \sqcap C & \sqsubseteq & X_{D_2} & \text{if } \mathcal{O}_S \models D_1 \sqcap C \sqsubseteq D_2 \\
X_{D_1} \sqcap X_{D_2} & \sqsubseteq & X_{D_3} & \text{if } \mathcal{O}_S \models D_1 \sqcap D_2 \sqsubseteq D_3 \\
\exists r.X_{D_1} & \sqsubseteq & X_{D_2} & \text{if } \mathcal{O}_S \models \exists r.D_1 \sqsubseteq D_2 \\
X_{D_1} & \sqsubseteq & \exists r.X_{D_2} & \text{if } \mathcal{O}_S \models D_1 \sqsubseteq \exists r.D_2 \\
X_{D_1} & \sqsubseteq & C & \text{if } \mathcal{O}_S \models D_1 \sqsubseteq C
\end{array}
$$

Figure 1: Candidate $\mathcal{EL}$ approximation $\mathcal{O}_T$.

has a finite projective $\mathcal{EL}$ approximation, but every non-projective $\mathcal{EL}$ approximation is infinite.

In fact, a finite projective $\mathcal{EL}$ approximation $\mathcal{O}_T$ of the ontology $\mathcal{O}_S$ from Proposition 1 is obtained from $\mathcal{O}_S$ by replacing the CI in the first line with

$$A \sqsubseteq X_{B_1 \sqcup B_2},\ \exists r.X_{B_1 \sqcup B_2} \sqsubseteq X_{B_1 \sqcup B_2},\ X_{B_1 \sqcup B_2} \sqcap A' \sqsubseteq M.$$

The intuitive reason for why $\mathcal{O}_S$ has no finite non-projective $\mathcal{EL}$ approximation is that $\mathcal{O}_S \models A' \sqcap \exists r^n.A \sqsubseteq M$ for all $n \geq 0$. Proposition 1 indicates that projective approximations are preferable. Since they also seem perfectly acceptable from an application viewpoint, we concentrate on the projective case and from now on mean projective approximations whenever we speak of approximations.

To illustrate the challenges of $\mathcal{ELU}$-to-$\mathcal{EL}$ approximation, it is instructive to consider a candidate approximation scheme that might be suggested by Proposition 1. We use $\mathtt{sub}(\mathcal{O}_S)$ to denote the set of all subconcepts of (concepts in) the ontology $\mathcal{O}_S$ and $\mathtt{sub}^-(\mathcal{O}_S)$ to denote the restriction of $\mathtt{sub}(\mathcal{O}_S)$ to concept names and existential restrictions $\exists r.C$. We use $\mathtt{Con}(\mathcal{O}_S)$ to denote the set of all non-empty conjunctions of concepts from $\mathtt{sub}^-(\mathcal{O}_S)$ without repetitions and $\mathtt{Dis}(\mathcal{O}_S)$ to mean the set of all disjunctions of concepts from $\mathtt{Con}(\mathcal{O}_S)$ without repetitions. Now, a (finite projective) candidate $\mathcal{EL}$ approximation scheme is given in Figure 1 where C ranges over $\mathtt{sub}(\mathcal{O}_S)$ and D_1, D_2, D_3 range over $\mathtt{Dis}(\mathcal{O}_S)$. It indeed yields an approximation when applied to the ontology $\mathcal{O}_S$ in Proposition 1. There are, however, two major problems. First, the syntactic structure of $\mathcal{O}_S$ is lost completely, which is undesirable in practice where ontologies are the result of a careful modeling effort. We could include all $\mathcal{EL}$ concept inclusions from $\mathcal{O}_S$ in the approximation, but this would be purely cosmetic since all such CIs are already implied. Second, the approximation is incomplete in general. In fact, finite approximations need not exist also in the projective case while the approximation scheme in Figure 1 is always finite.

Proposition 2. *The $\mathcal{ELU}$ ontology*

$$
\begin{array}{rrcll}
\mathcal{O}_S = \{ & A & \sqsubseteq & B_1 \sqcup B_2, & \\
& \exists r.B_2 & \sqsubseteq & \exists r.(B_1 \sqcap L), & \\
& L & \sqsubseteq & \exists s.L & \}
\end{array}
$$

has no finite $\mathcal{EL}$ approximation.

The intuitive reason for why $\mathcal{O}_S$ has no finite $\mathcal{EL}$ approximation is that $\mathcal{O}_S \models \exists r.(A \sqcap \exists s^n.\top) \sqsubseteq \exists r.(B_1 \sqcap \exists s^n.\top)$ for all $n \geq 0$.

The ontology in Proposition 2 can be varied to show that even bounded depth approximations can get very large. The function $\mathtt{tower} : \mathbb{N} \times \mathbb{N} \to \mathbb{N}$ is defined as $\mathtt{tower}(0, n) := n$ and $\mathtt{tower}(k+1, n) := 2^{\mathtt{tower}(k,n)}$.

$$
\begin{array}{rcll}
C & \sqsubseteq & \mathtt{DNF}(E)^\uparrow & \text{if } C \sqsubseteq E \in \mathcal{O}_S \\
X_D \sqcap D_1^\uparrow & \sqsubseteq & D_2^\uparrow & \text{if } \mathcal{O}_S \models D \sqcap D_1 \sqsubseteq D_2 \\
\exists r.X_D & \sqsubseteq & D_1^\uparrow & \text{if } \mathcal{O}_S \models \exists r.D \sqsubseteq D_1 \\
F^\uparrow & \sqsubseteq & \exists r.G & \text{if } \mathcal{O}_S \models F \sqsubseteq \exists r.G
\end{array}
$$

where in the last line

- F is an $\mathcal{EL}$ concept over $\mathtt{sig}(\mathcal{O}_S)$ decorated with disjunctions from $\mathtt{Dis}(\mathcal{O}_S)$ at leaves and
- G is an $\mathcal{O}_S$-generatable $\mathcal{EL}$ concept over $\mathtt{sig}(\mathcal{O}_S)$

such that $\mathtt{depth}(F) \leq \mathtt{depth}(G) < \ell$.

Figure 2: ℓ-bounded $\mathcal{EL}$ approximation $\mathcal{O}_T^\ell$.

Proposition 3. *Let $\mathcal{O}_S^n$ be obtained from the ontology $\mathcal{O}_S$ in Proposition 2 by replacing the bottommost CI with*

$$L \sqsubseteq A_1 \sqcap \hat{A}_1 \sqcap \cdots \sqcap A_n \sqcap \hat{A}_n \sqcap \exists r_1.L \sqcap \exists r_2.L$$

Then for all $n, \ell \geq 1$ and any ℓ-bounded $\mathcal{EL}$ approximation $\mathcal{O}_T$ of $\mathcal{O}_S^n$, $||\mathcal{O}_T|| \geq \mathtt{tower}(\ell, n)$.

A Complete Approximation. We present a more careful approximation scheme that aims to preserve the structure of $\mathcal{O}_S$, is complete, and yields a finite approximation in many practical cases. Let $\mathcal{O}_S$ be an $\mathcal{ELU}$ ontology to be approximated. As a preliminary, we assume that for all CIs $C \sqsubseteq D \in \mathcal{O}_S$, C is an $\mathcal{EL}$ concept. If this is not the case, then we can rewrite $\mathcal{O}_S$ by exhaustively replacing every disjunction $C \sqcup D$ that occurs (possibly as a subconcept) on the left-hand side of a concept inclusion in $\mathcal{O}_S$ with a fresh concept name $X_{C \sqcup D}$ and adding the inclusions $C \sqsubseteq X_{C \sqcup D}$ and $D \sqsubseteq X_{C \sqcup D}$. It is not hard to see that the resulting ontology $\mathcal{O}_S'$ is of size polynomial in $||\mathcal{O}_S||$ and a conservative extension of $\mathcal{O}_S$ in the sense that $\mathcal{O}_S \models C \sqsubseteq D$ iff $\mathcal{O}_S' \models C \sqsubseteq D$ for all $\mathcal{ELU}$ concepts C, D over $\mathtt{sig}(\mathcal{O}_S)$. Consequently, every $\mathcal{EL}$ approximation of $\mathcal{O}_S'$ is also a projective $\mathcal{EL}$ approximation of $\mathcal{O}_S$ and we can work with $\mathcal{O}_S'$ in place of $\mathcal{O}_S$.

Let $\ell \in \mathbb{N} \cup \{\omega\}$. The proposed $\mathcal{EL}$ approximation $\mathcal{O}_T^\ell$ of $\mathcal{O}_S$ is given in Figure 2 where D_1, D_2 range over $\mathtt{Dis}(\mathcal{O}_S)$ and D ranges over $\mathtt{Dis}^-(\mathcal{O}_S)$, the set of all disjunctions in $\mathtt{Dis}(\mathcal{O}_S)$ that have at least two disjuncts. We still have to define the notation and terminology used in the figure. For an $\mathcal{ELU}$ concept C such that all disjunctions in C are from $\mathtt{Dis}(\mathcal{O}_S)$, we use $C^\uparrow$ to denote the $\mathcal{EL}$ concept obtained from C by replacing every outermost $D \in \mathtt{Dis}^-(\mathcal{O}_S)$ with a fresh concept name X_D. Set $\mathtt{DNF}(C) = C$ if C is a concept name or of the form $\exists r.D$, $\mathtt{DNF}(C_1 \sqcap C_2) = \mathtt{DNF}(C_1) \sqcap \mathtt{DNF}(C_2)$, and define $\mathtt{DNF}(C_1 \sqcup C_2)$ to be the $\mathcal{ELU}$-concept obtained by converting $C_1 \sqcup C_2$ into disjunctive normal form (DNF), treating existential restrictions $\exists r.D$ as atomic concepts, that is, the argument D is not modified. Note that while $||\mathtt{DNF}(C)||$ may be exponential in $||C||$, we have $||\mathtt{DNF}(C)^\uparrow|| \leq ||C||$. By *decorating an $\mathcal{EL}$ concept C with disjunctions from $\mathtt{Dis}(\mathcal{O}_S)$ at leaves*, we mean to replace subconcepts $\exists r.E$ of C with E of depth 0 by $\exists r.(E \sqcap D)$, $D \in \mathtt{Dis}(\mathcal{O}_S)$. As a special case, we can replace C with $C \sqcap D$, $D \in \mathtt{Dis}(\mathcal{O}_S)$, if C is of depth 0.

Definition 2. *An $\mathcal{EL}$ concept C is* $\mathcal{O}_S$-generatable *if there is an $\exists r.D \in \mathtt{sub}(\mathcal{O}_S)$ that occurs on the right-hand side of a CI in $\mathcal{O}_S$ and satisfies $\mathcal{O}_S \models D \sqsubseteq C$.*

Let us explain the proposed approximation. The first three lines of Figure 2 can be viewed as a more careful version of the first four lines of Figure 1. In the first line, we preserve the structure of $\mathcal{O}_S$ as long as it lies outside the scope of a disjunction operator, thanks to the careful definition of $\mathtt{DNF}(C)$. This is not cosmetic as in the candidate approximation in Figure 1: since we introduce the concept names X_D only when a disjunction is 'derived' (first line) and only for disjunctions $D \in \mathtt{Dis}^-(\mathcal{O}_S)$, $\mathcal{O}_T^\ell$ is no longer guaranteed to be an approximation when the first line in Figure 2 is dropped. The last line of the approximation addresses the effect illustrated by Proposition 2. It is strong enough so that a counterpart of the second last line in Figure 1 is not needed. An example application of our approximation scheme is given in [Haga *et al.*, 2020].

An interesting aspect of our approximation is that it turns out to be finite in many practical cases. In fact, it is easy to see that $\mathcal{O}_T^\ell$ is finite for all $\ell < \omega$ and that $\mathcal{O}_T^\omega$ is finite if and only if there are only finitely many $\mathcal{EL}$ concepts that are $\mathcal{O}_S$-generatable, up to logical equivalence; we then say that $\mathcal{O}_S$ is *finitely generating*. Since ontologies from practice tend to have a simple structure, one might expect that they often enjoy this property. Below, we report about a case study that confirms this expectation.

How does the approximation scheme in Figure 2 relate to the examples given above? For the ontologies $\mathcal{O}_S$ in Example 1 and in Proposition 1, our approximation $\mathcal{O}_T^\omega$ contains all CIs in the approximation $\mathcal{O}_T$ given in place. Of course, $\mathcal{O}_T^\omega$ also contains a lot of additional CIs that, however, do not result in any new consequences $C \sqsubseteq D$ with C, D $\mathcal{EL}$ concepts over $\mathtt{sig}(\mathcal{O}_S)$. It seems very difficult to identify up front those CIs that are really needed. We can remove them after constructing $\mathcal{O}_T^\omega$ by repeatedly deciding conservative extensions [Lutz and Wolter, 2010], but this is not practical given the size of $\mathcal{O}_T^\omega$. Nevertheless, both ontologies $\mathcal{O}_S$ are finitely generating and thus in both cases $\mathcal{O}_T^\omega$ is finite. In Example 1, the $\mathcal{O}_S$-generatable concepts are $\top$, $\mathtt{MainJob}$, $\mathtt{PartTime}$, and $\mathtt{MainJob} \sqcap \mathtt{PartTime}$ (up to logical equivalence) while there are no $\mathcal{O}_S$-generatable concepts for Proposition 1. For Proposition 2, there are infinitely many $\mathcal{O}_S$-generatable concepts such as $\exists s^n.\top$ for all $n \geq 0$.

Case Study. We have considered the seven non-trivial $\mathcal{ELU}$ ontologies that are part of the Manchester OWL corpus.[1] The size of the ontologies ranges from 113 to 813 concept inclusions and equalities. All ontologies use disjunction on the right-hand side of CIs (thus in a non-trivial way) and none of them is acyclic. We have been able to prove that all these ontologies are finitely generating and thus the approximation $\mathcal{O}_T^\omega$ is finite. Our proof relies on the following observation.

Lemma 1. *$\mathcal{O}_S$ is not finitely generating iff for every $n \geq 0$, there is an $\exists r.D \in \mathtt{sub}(\mathcal{O}_S)$ that occurs on the right-hand side of a CI and a sequence $r_1, \dots, r_n$ of role names from $\mathcal{O}_S$ such that $\mathcal{O}_S \models D \sqsubseteq \exists r_1.\cdots.\exists r_n.\top$.*

In our implementation, we use role inclusions to avoid going through all of the exponentially many sequences $r_1, \dots, r_n$. Lemma 1 can also be used to show the following.

[1] http://owl.cs.manchester.ac.uk/publications/supporting-material/owlcorpus/

Theorem 1. *It is decidable whether a given $\mathcal{ELU}$-ontology $\mathcal{O}_S$ is finitely generating.*

By what was said above, this implies that it is decidable whether the approximation $\mathcal{O}_T^\omega$ from Figure 2 is finite.

Soundness and Completeness. We now establish soundness and completeness of the proposed approximation, the main result in this section.

Theorem 2. *For every $\ell \in \mathbb{N} \cup \{\omega\}$, $\mathcal{O}_T^\ell$ is an ℓ-bounded $\mathcal{EL}$ approximation of $\mathcal{O}_S$.*

While soundness is easy to show, completeness is remarkably subtle to prove. It is stated by the following lemma which shows that our approximation $\mathcal{O}_T^\ell$ is actually stronger than required in that it preserves all $\mathcal{EL}$ subsumptions $C \sqsubseteq D$ with D of depth bounded by ℓ and C of unrestricted depth.

Lemma 2. *Let $\ell \in \mathbb{N} \cup \{\omega\}$. Then $\mathcal{O}_S \models C_0 \sqsubseteq D_0$ implies $\mathcal{O}_T^\ell \models C_0 \sqsubseteq D_0$ for all $\mathcal{EL}$ concepts C_0, D_0 over $\mathtt{sig}(\mathcal{O}_S)$ such that the role depth of D_0 is bounded by ℓ.*

The proof of Lemma 2 is the most substantial one in this paper. It uses a chase procedure for $\mathcal{ELU}$ ontologies that is specifically tailored towards proving completeness in that it is deterministic rather than disjunctive and mimics the concept inclusions in Figure 2. Showing that this chase is complete is far from trivial.

Fewer Symbols. The number of fresh concept names X_D in $\mathcal{O}_T^\ell$ is double exponential in $||\mathcal{O}_S||$ since the number of disjunctions in $\mathtt{Dis}^-(\mathcal{O})$ is. However, $\mathcal{O}_T^\ell$ can be rewritten into an ontology $\widehat{\mathcal{O}}_T^\ell$ that uses only single exponentially many fresh concept names and is still an ℓ-bounded approximation of $\mathcal{O}_S$. The idea is to transition from disjunctive normal form to conjunctive normal form, that is, to replace each concept name X_D, $D \in \mathtt{Dis}^-(\mathcal{O})$, with a conjunction of concept names $Y_{D'}$ where D' is a disjunction of concepts from $\mathtt{sub}^-(\mathcal{O})$, rather than conjunctions thereof. Details are in [Haga *et al.*, 2020].

Theorem 3. *For every $\ell \in \mathbb{N} \cup \{\omega\}$, $\widehat{\mathcal{O}}_T^\ell$ is an ℓ-bounded $\mathcal{EL}$ approximation of $\mathcal{O}_S$.*

Acyclic Ontologies. Using Lemma 1, one can show that $\mathcal{O}_T^\omega$ is finite whenever $\mathcal{O}_S$ is an acyclic $\mathcal{ELU}$ ontology. In fact, the length n of role sequences with the properties stated in the lemma is bounded by $||\mathcal{O}_S||$ if $\mathcal{O}_S$ is acyclic.

Theorem 4. *Every acyclic $\mathcal{ELU}$ ontology has a finite $\mathcal{EL}$ approximation.*

There is, however, more that we can say about acyclic ontologies. We first observe that there are acyclic $\mathcal{ELU}$ ontologies that have finite $\mathcal{EL}$ approximations, but no $\mathcal{EL}$ approximation that is an acyclic ontology.

Example 2. *Consider the acyclic $\mathcal{ELU}$ ontology*

$$\mathcal{O}_S = \{A \equiv (B_1 \sqcap B_2) \sqcup (B_1 \sqcap B_3)\}.$$

Then $\mathcal{O}_T = \{B_1 \sqcap B_2 \sqsubseteq A, B_1 \sqcap B_3 \sqsubseteq A, A \sqsubseteq B_1\}$ is an $\mathcal{EL}$ approximation of $\mathcal{O}_S$, but $\mathcal{O}_S$ has no $\mathcal{EL}$ approximation that is an acyclic ontology, finite or infinite.

Further, our approximations $\mathcal{O}_T^\ell$ can be simplified for acyclic $\mathcal{ELU}$ ontologies $\mathcal{O}_S$. Let $\widetilde{\mathcal{O}}_T^\ell$ be defined like $\mathcal{O}_T^\ell$

in Figure 2, except that in the last line, F ranges only over concept names (not decorated with disjunctions) rather than over compound concepts, a significant simplification.

Theorem 5. *Let $\ell \in \mathbb{N} \cup \{\omega\}$ and let $\mathcal{O}_S$ be an acyclic $\mathcal{ELU}$ ontology. Then $\widetilde{\mathcal{O}}^\ell_T$ is an ℓ-bounded $\mathcal{EL}$ approximation of $\mathcal{O}_S$.*

Based on this observation, constructing finite $\mathcal{EL}$ approximations of acyclic $\mathcal{ELU}$ ontologies does not seem infeasible in practice.

4 $\mathcal{ALC}$-to-$\mathcal{EL}_\bot$ Approximation

We consider $\mathcal{ELU}_\bot$-to-$\mathcal{EL}_\bot$ and $\mathcal{ALC}$-to-$\mathcal{EL}_\bot$ approximation which turn out to be closely related to each other and significantly different from $\mathcal{ELU}$-to-$\mathcal{EL}$ approximation.

It immediately follows from the results in Section 3 that finite approximations are guaranteed to exist neither in the $\mathcal{ELU}_\bot$-to-$\mathcal{EL}_\bot$ nor in the $\mathcal{ALC}$-to-$\mathcal{EL}_\bot$ case. However, while we have argued that finite $\mathcal{ELU}$-to-$\mathcal{EL}$ approximations can be expected to exist in many practical cases, this does not appear to be true for $\mathcal{ELU}_\bot$-to-$\mathcal{EL}_\bot$ and $\mathcal{ALC}$-to-$\mathcal{EL}_\bot$. The following example illustrates the problem.

Example 3. *Consider the $\mathcal{ELU}_\bot$ ontology*

$$\mathcal{O}_S = \{ \quad A_1 \sqsubseteq M \sqcup N_1, \quad A_2 \sqsubseteq M \sqcup N_2, \quad \exists r.N_1 \sqcap \exists r.N_2 \sqsubseteq \bot \quad \}.$$

There are no $\mathcal{O}_S$-generatable $\mathcal{EL}$ concepts. Yet, there is no finite $\mathcal{EL}_\bot$ approximation of $\mathcal{O}_S$. Informally, this is because

$$\mathcal{O}_S \models \exists r.(A_1 \sqcap \exists r^n.\top) \sqcap \exists r.(A_2 \sqcap \exists r^n.\top) \sqsubseteq \exists r.(M \sqcap \exists r^n.\top)$$

for all $n \geq 1$.[2]

While the above example is for $\mathcal{ELU}_\bot$-to-$\mathcal{EL}_\bot$, there is an additional effect in $\mathcal{ALC}$-to-$\mathcal{EL}_\bot$ that already occurs for very simple ontologies $\mathcal{O}_S$.

Example 4. *The $\mathcal{ALC}$ ontology $\mathcal{O}_S = \{A \sqsubseteq \forall r.B\}$ has no finite $\mathcal{EL}_\bot$ approximation. This is shown in [Bötcher* et al.*, 2019] for the equivalent $\mathcal{ELI}$ ontology $\{\exists r^-.A \sqsubseteq B\}$. Informally, this is because $\mathcal{O}_S \models A \sqcap \exists r^{n+1}.\top \sqsubseteq \exists r.(B \sqcap \exists r^n.\top)$ for all $n \geq 1$.*

Note that the ontology $\mathcal{O}_S$ in Example 4 is acyclic and thus in contrast to the $\mathcal{ELU}$-to-$\mathcal{EL}$ case, finite $\mathcal{EL}_\bot$ approximations of acyclic $\mathcal{ALC}$ ontologies need not exist. In a sense, Example 3 shows the same negative result for the $\mathcal{ELU}_\bot$-to-$\mathcal{EL}_\bot$ case. While the ontology used there is not strictly acyclic, acyclic ontologies do not make much sense in the case of $\mathcal{ELU}_\bot$ and additionally admitting CIs $C_1 \sqcap C_2 \sqsubseteq \bot$ as used in Example 3 seems to be the most modest extension possible that incorporates $\bot$ in a meaningful way.

Despite these additional challenges, we can extend the approximation given in Section 3 to $\mathcal{ELU}_\bot$-to-$\mathcal{EL}_\bot$ and to $\mathcal{ALC}$-to-$\mathcal{EL}_\bot$ when we are willing to drop $\mathcal{O}_S$-generatability and, as a consequence, accept the fact that approximations are infinite unless they are depth bounded. Note that the latter is also the case in $\mathcal{L}$-to-$\mathcal{EL}$ approximation where $\mathcal{L}$ is an expressive Horn DL such as $\mathcal{ELI}$ [Bötcher *et al.*, 2019].

[2] A formal proof is analogous to that of Proposition 2.

$$\begin{array}{rcll} C & \sqsubseteq & \mathtt{DNF}(E)^\uparrow & \text{if } C \sqsubseteq E \in \mathcal{O}_S \\ X_D \sqcap D_1^\uparrow & \sqsubseteq & D_2^\uparrow & \text{if } \mathcal{O}_S \models D \sqcap D_1 \sqsubseteq D_2 \\ \exists r.X_D & \sqsubseteq & D_1^\uparrow & \text{if } \mathcal{O}_S \models \exists r.D \sqsubseteq D_1 \\ F^\uparrow & \sqsubseteq & \exists r.G & \text{if } \mathcal{O}_S \models F \sqsubseteq \exists r.G \end{array}$$

where in the last line F is an $\mathcal{EL}$ concept over $\mathtt{sig}(\mathcal{O}_S)$ decorated with disjunctions from $\mathtt{Dis}(\mathcal{O}_S)$ at leaves and G is an $\mathcal{EL}$ concept over $\mathtt{sig}(\mathcal{O}_S)$ such that

1. F has no top-level conjunct $\exists r.F'$ s.t. $\mathcal{O}_S \models F' \sqsubseteq G$;
2. $\mathtt{depth}(F) \leq \mathtt{depth}(G) < \ell$.

Figure 3: ℓ-bounded $\mathcal{EL}_\bot$ approximation $\mathcal{O}^\ell_T$.

We first reduce $\mathcal{ALC}$-to-$\mathcal{EL}_\bot$ approximations to $\mathcal{ELU}_\bot$-to-$\mathcal{EL}_\bot$ approximations. Let $\mathcal{O}_S$ be an $\mathcal{ALC}$ ontology. We can transform $\mathcal{O}_S$ into an $\mathcal{ELU}_\bot$ ontology as follows:

1. replace each subconcept $\forall r.C$ with $\neg\exists r.\neg C$;
2. select a concept $\neg C$ such that C contains no negation, replace all occurrences of $\neg C$ with the fresh concept name $A_{\neg C}$, and add the CIs $\top \sqsubseteq C \sqcup A_{\neg C}$ and $C \sqcap A_{\neg C} \sqsubseteq \bot$; repeat until no longer possible.

The resulting ontology $\mathcal{O}'_S$ is of size polynomial in $||\mathcal{O}_S||$ and a conservative extension of $\mathcal{O}_S$ in the sense that $\mathcal{O}_S \models C \sqsubseteq D$ iff $\mathcal{O}'_S \models C \sqsubseteq D$ for all $\mathcal{ALC}$ concepts C, D over $\mathtt{sig}(\mathcal{O}_S)$. Consequently, every $\mathcal{EL}_\bot$ approximation of $\mathcal{O}'_S$ is also a (projective) $\mathcal{EL}_\bot$ approximation of $\mathcal{O}_S$.

It thus suffices to consider $\mathcal{ELU}_\bot$-to-$\mathcal{EL}_\bot$ approximations. Thus let $\mathcal{O}_S$ be an $\mathcal{ELU}_\bot$ ontology. For each $\ell \in \mathbb{N} \cup \{\omega\}$, the $\mathcal{EL}_\bot$ approximation $\mathcal{O}^\ell_T$ of $\mathcal{O}_S$ is given in Figure 2 where again D ranges over $\mathtt{Dis}^-(\mathcal{O}_S)$ and D_1, D_2 range over $\mathtt{Dis}(\mathcal{O}_S)$; both $\mathtt{Dis}(\mathcal{O}_S)$ and $\mathtt{Dis}^-(\mathcal{O}_S)$ are defined exactly as for $\mathcal{ELU}$ ontologies and in $\mathtt{DNF}(C)$ we drop all disjuncts that contain $\bot$ as a conjunct, possibly resulting in the empty disjunction (which represents $\bot$). Point 1 can be viewed as an optimization that sometimes helps to avoid the expensive last line. There, a *top-level conjunct* means a concept F_i if F takes the form $F_1 \sqcap \cdots \sqcap F_n$, $n \geq 1$. In [Haga *et al.*, 2020], we point out another non-trivial such optimization.

Theorem 6. *$\mathcal{O}^\ell_T$ is an ℓ-bounded $\mathcal{EL}_\bot$ approximation of $\mathcal{O}_S$.*

The proof of Theorems 2 and 6 also establishes another result that will turn out to be interesting in the context of ontology-mediated queries in Section 5. We use $\mathcal{O}^-_T$ to denote the restriction of $\mathcal{O}^\omega_T$ to the (instantiations) of the first three lines in Figure 3 (equivalently: Figure 2). Clearly, $\mathcal{O}^-_T$ is always finite.

Theorem 7. *Let C_0, D_0 be $\mathcal{EL}_\bot$ concepts with $D_0 \in \mathtt{sub}(\mathcal{O}_S)$. Then $\mathcal{O}_S \models C_0 \sqsubseteq D_0$ iff $\mathcal{O}^-_T \models C_0 \sqsubseteq D_0$.*

5 Approximations and Query Evaluation

The notion of approximations given in Section 2 is tailored towards preserving subsumptions. In ontology-mediated querying, in contrast, the main aim of approximation is to preserve as many query answers as possible. We propose a suitable notion of approximation and show that the results obtained in the previous sections have interesting applications also in ontology-mediated querying.

Let $\mathsf{N_I}$ be a countably infinite set of *individual names* disjoint from $\mathsf{N_C}$ and $\mathsf{N_R}$. An *ABox* is a finite set of *concept assertions* $A(a)$ and *role assertions* $r(a,b)$ where $A \in \mathsf{N_C}$, $r \in \mathsf{N_R}$, and $a,b \in \mathsf{N_I}$. We use $\mathtt{Ind}(\mathcal{A})$ to denote the set of individual names in the ABox $\mathcal{A}$. An interpretation $\mathcal{I}$ *satisfies* a concept assertion $A(a)$ if $a \in A^{\mathcal{I}}$ and a role assertion $r(a,b)$ if $(a,b) \in r^{\mathcal{I}}$. It is a *model* of an ABox if it satisfies all assertions in it. A Σ-*ABox* is an ABox $\mathcal{A}$ with $\mathtt{sig}(\mathcal{A}) = \Sigma$.

An *ontology-mediated query (OMQ)* is a triple $Q = (\mathcal{O}, \Sigma, q)$ with $\mathcal{O}$ an ontology, Σ an ABox signature, and q an actual query. While conjunctive queries (CQs) and unions of CQs are a popular choice for formulating q and our central Definition 3 below makes sense also for these richer query languages, for simplicity we concentrate on *atomic queries (AQs)* $A(x)$ where A is a concept name and on $\mathcal{EL}$ *queries (ELQs)* $C(x)$ where C an $\mathcal{EL}$ concept. We also mention $\mathcal{ALC}$ *queries (ALCQs)* $C(x)$ where C is an $\mathcal{ALC}$ concept. Note that all such queries are unary. We use ELQ(Σ) to denote the language of all ELQs that use only symbols from signature Σ. Let $(\mathcal{L}, \mathcal{Q})$ denote the *OMQ language* that contains all OMQs Q in which $\mathcal{O}$ is formulated in DL $\mathcal{L}$ and q in query language $\mathcal{Q}$, such as in $(\mathcal{EL}, \text{AQ})$.

Let $Q = (\mathcal{O}, \Sigma, C(x))$ be an OMQ and $\mathcal{A}$ a Σ-ABox. Then $a \in \mathtt{Ind}(\mathcal{A})$ is an *answer* to Q on $\mathcal{A}$, written $\mathcal{A} \models Q(a)$, if $a \in C^{\mathcal{I}}$ for all models $\mathcal{I}$ of $\mathcal{O}$ and $\mathcal{A}$. For OMQs Q_1 and Q_2, $Q_i = (\mathcal{O}_i, \Sigma, q_i)$, we say that Q_1 is *contained* in Q_2 and write $Q_1 \subseteq Q_2$ if for every Σ-ABox $\mathcal{A}$ and $a \in \mathtt{Ind}(\mathcal{A})$, $\mathcal{A} \models Q_1(a)$ implies $\mathcal{A} \models Q_2(a)$. We say that Q_1 is *equivalent* to Q_2 and write $Q_1 \equiv Q_2$ if $Q_1 \subseteq Q_2$ and $Q_2 \subseteq Q_1$.

A natural definition of ontology approximation in the context of OMQs is as follows.

Definition 3. *Let $\mathcal{O}_S$ be an $\mathcal{ALC}$ ontology, $\mathcal{L}_T$ one of the DLs from Section 2, and $\mathcal{Q}$ a query language. An $\mathcal{L}_T$ ontology $\mathcal{O}_T$ is an $\mathcal{L}_T$* approximation of $\mathcal{O}_S$ w.r.t $\mathcal{Q}$ *if for all queries $q \in \mathcal{Q}$ and all signatures Σ with $\Sigma \cap \mathtt{sig}(\mathcal{O}_T) \subseteq \mathtt{sig}(\mathcal{O}_S)$,*

1. *$(\mathcal{O}_S, \Sigma, q) \supseteq (\mathcal{O}_T, \Sigma, q)$ and*
2. *$(\mathcal{O}_S, \Sigma, q) \supseteq Q$ implies $(\mathcal{O}_T, \Sigma, q) \supseteq Q$ for all OMQs $Q = (\mathcal{O}'_T, \Sigma, q) \in (\mathcal{L}_T, \mathcal{Q})$.*

$\mathcal{O}_T$ might use fresh symbols and thus approximations are projective. Informally, Point 1 is a soundness condition and Point 2 formalizes 'to preserve as many query answers as possible'. It is not guaranteed that the OMQs $(\mathcal{O}_S, \Sigma, q)$ and $(\mathcal{O}_T, \Sigma, q)$ are equivalent for all relevant queries q and signatures Σ, and the following example shows that this is in fact impossible to achieve. However, Point 2 of Definition 3 ensures that $(\mathcal{O}_T, \Sigma, q)$ is the best approximation of $(\mathcal{O}_S, \Sigma, q)$ from below among all OMQs from $(\mathcal{L}_T, \mathcal{Q})$.

Example 5. *Let $\mathcal{O}_S$ be the $\mathcal{ELU}$ ontology*

$$\mathcal{O}_S = \{\top \sqsubseteq B_1 \sqcup B_2\} \cup \{B_i \sqcap \exists r.B_i \sqsubseteq A \mid i \in \{1,2\}\}$$

Then an $\mathcal{EL}$ approximation of $\mathcal{O}_S$ w.r.t. ELQ is

$$\mathcal{O}_T = \{B_1 \sqcap B_2 \sqcap \exists r.\top \sqsubseteq A,\ \exists r.(B_1 \sqcap B_2) \sqsubseteq A\} \cup \{B_i \sqcap \exists r.B_i \sqsubseteq A \mid i \in \{1,2\}\}.$$

However, there is no OMQ in $(\mathcal{EL}, \text{ELQ})$ that is equivalent to $(\mathcal{O}_S, \{r\}, A(x))$ since it would have to return a as an answer on the ABox $\{r(a,a)\}$, but not on the ABox $\{r(a,b), r(b,a)\}$. No OMQ from $(\mathcal{EL}, \text{ELQ})$ has this property.

It turns out that the approximations from Sections 3 and 4 are also useful in the context of Definition 3 when we choose ELQ or AQ as the query language. In particular, it follows from Theorem 7 that every $\mathcal{ALC}$ ontology $\mathcal{O}_S$ has a finite $\mathcal{EL}_\bot$ approximation w.r.t. AQ, namely the fragment $\mathcal{O}_T^-$ of the approximation scheme proposed in Section 4.

Theorem 8. *Let $\mathcal{O}_S$ be an $\mathcal{ALC}$ ontology, $\mathtt{sig}(\mathcal{O}_S) = \Sigma$. Then*

1. *the ontology $\mathcal{O}_T^\omega$ from Section 4 is an $\mathcal{EL}_\bot$ approximation of $\mathcal{O}_S$ w.r.t. ELQ(Σ);*
2. *the ontology $\mathcal{O}_T^-$ from Section 4 is a (finite) $\mathcal{EL}_\bot$ approximation of $\mathcal{O}_S$ w.r.t. AQ;*
3. *if $\mathcal{O}_S$ falls within $\mathcal{ELU}$, then the ontology $\mathcal{O}_T^\omega$ from Section 3 is an $\mathcal{EL}$ approximation of $\mathcal{O}_S$ w.r.t. ELQ(Σ).*

Point 2 also implies that $\mathcal{O}_T^-$ is an $\mathcal{EL}$ approximation of $\mathcal{O}_S$ w.r.t. AQ whenever $\mathcal{O}_S$ is an $\mathcal{ELU}$ ontology. We close with an interesting application of Theorem 8.

The topic of rewriting an OMQ into a simpler query language has received a lot of interest in the literature, see for example [Calvanese *et al.*, 2007; Gottlob *et al.*, 2014; Kaminski *et al.*, 2016; Feier *et al.*, 2019] and references therein. An OMQ Q is $(\mathcal{L}, \mathcal{Q})$-*rewritable* if there is an OMQ Q' in the OMQ language $(\mathcal{L}, \mathcal{Q})$ such that $Q \equiv Q'$. By virtue of Theorem 8, we can decide whether an OMQ $Q = (\mathcal{O}, \Sigma, A(x))$ from $(\mathcal{ALC}, \text{AQ})$ is $(\mathcal{EL}_\bot, \text{AQ})$-rewritable: construct the finite approximation $\mathcal{O}_T^-$ of $\mathcal{O}$ and check whether $Q \equiv (\mathcal{O}_T^-, \Sigma, A(x))$. The latter is decidable [Bienvenu *et al.*, 2014] and by Condition 2 of Definition 3, the answer is 'yes' if and only if Q is $(\mathcal{EL}_\bot, \text{AQ})$-rewritable. We can extend this result to $(\mathcal{ALC}, \text{ALCQ})$ since every OMQ from this language is equivalent to one from $(\mathcal{ALC}, \text{AQ})$. Via the results in [Feier *et al.*, 2018], this result even extends further to a certain class of conjunctive queries.

Theorem 9. *Given an OMQ $Q \in (\mathcal{ALC}, \text{ALCQ})$, it is decidable whether Q is $(\mathcal{EL}_\bot, \text{AQ})$-rewritable.*

6 Conclusion

We have investigated the structure and finiteness of ontology approximations when transitioning from non-Horn DLs to Horn DLs. We believe that our results shed significant light on the situation. It remains, however, an important and challenging topic for future work to push our techniques further towards practical applicability. Also, there are many other relevant cases of approximation. As a first step, one might think about extending the DLs considered in this paper with role inclusions. It might further be interesting to study the problem to decide whether a given (finite) candidate is an approximation of a given ontology. We expect this to be quite non-trivial. A related result in [Lutz *et al.*, 2012] states that it is between ExpTime and 2ExpTime to decide whether a given $\mathcal{ELU}$ ontology $\mathcal{O}_S$ of a restricted syntactic form has a finite complete $\mathcal{EL}$ approximation. Without the restriction, even decidability is open.

Acknowledgements

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

References

[Baader *et al.*, 2017] Franz Baader, Ian Horrocks, Carsten Lutz, and Ulrike Sattler. *An Introduction to Description Logic*. Cambridge University Press, 2017.

[Bienvenu *et al.*, 2014] Meghyn Bienvenu, Balder ten Cate, Carsten Lutz, and Frank Wolter. Ontology-based data access: A study through disjunctive datalog, CSP, and MMSNP. *ACM Trans. Database Syst.*, 39(4):33:1–33:44, 2014.

[Bötcher *et al.*, 2019] Anneke Bötcher, Carsten Lutz, and Frank Wolter. Ontology approximation in Horn description logics. In *Proc. of IJCAI*, pages 1574–1580. ijcai.org, 2019.

[Botoeva *et al.*, 2010] Elena Botoeva, Diego Calvanese, and Mariano Rodriguez-Muro. Expressive approximations in *DL-Lite* ontologies. In *Proc. of AIMSA*, volume 6304 of *LNCS*, pages 21–31. Springer, 2010.

[Calvanese *et al.*, 2007] Diego Calvanese, Giuseppe De Giacomo, Domenico Lembo, Maurizio Lenzerini, and Riccardo Rosati. Tractable reasoning and efficient query answering in description logics: The *DL-Lite* family. *J. Autom. Reasoning*, 39(3):385–429, 2007.

[Calvanese *et al.*, 2009] Diego Calvanese, Giuseppe De Giacomo, Domenico Lembo, Maurizio Lenzerini, Antonella Poggi, Mariano Rodriguez-Muro, and Riccardo Rosati. Ontologies and databases: The DL-Lite approach. In *Reasoning Web*, volume 5689 of *LNCS*, pages 255–356. Springer, 2009.

[Carral *et al.*, 2014] David Carral, Cristina Feier, Bernardo Cuenca Grau, Pascal Hitzler, and Ian Horrocks. $\mathcal{EL}$-ifying ontologies. In *Proc. of IJCAR*, pages 464–479, 2014.

[Cucala *et al.*, 2019] David Tena Cucala, Bernardo Cuenca Grau, and Ian Horrocks. 15 years of consequence-based reasoning. In *Description Logic, Theory Combination, and All That - Essays Dedicated to Franz Baader on the Occasion of His 60th Birthday*, volume 11560 of *LNCS*, pages 573–587. Springer, 2019.

[Feier *et al.*, 2018] Cristina Feier, Carsten Lutz, and Frank Wolter. From conjunctive queries to instance queries in ontology-mediated querying. In *Proc. of IJCAI*, pages 1810–1816. ijcai.org, 2018.

[Feier *et al.*, 2019] Cristina Feier, Antti Kuusisto, and Carsten Lutz. Rewritability in monadic disjunctive datalog, MMSNP, and expressive description logics. *Logical Methods in Computer Science*, 15(2), 2019.

[Gottlob *et al.*, 2014] Georg Gottlob, Stanislav Kikot, Roman Kontchakov, Vladimir V. Podolskii, Thomas Schwentick, and Michael Zakharyaschev. The price of query rewriting in ontology-based data access. *Artif. Intell.*, 213:42–59, 2014.

[Haga *et al.*, 2020] Anneke Haga, Carsten Lutz, Johannes Marti, and Frank Wolter. A journey into ontology approximation: From Non-Horn to Horn. *CoRR*, abs/2001.07754, 2020.

[Kaminski *et al.*, 2016] Mark Kaminski, Yavor Nenov, and Bernardo Cuenca Grau. Datalog rewritability of disjunctive datalog programs and non-Horn ontologies. *Artif. Intell.*, 236:90–118, 2016.

[Lutz and Wolter, 2010] Carsten Lutz and Frank Wolter. Deciding inseparability and conservative extensions in the description logic $\mathcal{EL}$. *J. Symb. Comput.*, 45(2):194–228, 2010.

[Lutz *et al.*, 2012] Carsten Lutz, Inanç Seylan, and Frank Wolter. An automata-theoretic approach to uniform interpolation and approximation in the description logic $\mathcal{EL}$. In *Proc. of KR*. AAAI Press, 2012.

[Pan and Thomas, 2007] Jeff Z. Pan and Edward Thomas. Approximating *OWL-DL* ontologies. In *AAAI*, pages 1434–1439, 2007.

[Ren *et al.*, 2010] Yuan Ren, Jeff Z. Pan, and Yuting Zhao. Soundness preserving approximation for tbox reasoning. In *Proc. of AAAI*. AAAI Press, 2010.

[Zhou *et al.*, 2015] Yujiao Zhou, Bernardo Cuenca Grau, Yavor Nenov, Mark Kaminski, and Ian Horrocks. Pagoda: Pay-as-you-go ontology query answering using a datalog reasoner. *J. Artif. Intell. Res.*, 54:309–367, 2015.

A Proofs for Propositions 1, 2, and 4

We state the results to be proved again.

Proposition 1. *The $\mathcal{ELU}$ ontology*

$$\begin{array}{rrcll} \mathcal{O}_S = \{ & A & \sqsubseteq & B_1 \sqcup B_2, & \\ & \exists r.B_i & \sqsubseteq & B_i, & \text{for } i \in \{1,2\} \\ & B_i \sqcap A' & \sqsubseteq & M & \text{for } i \in \{1,2\}\ \}. \end{array}$$

has a finite projective $\mathcal{EL}$ approximation, but every non-projective $\mathcal{EL}$ approximation is infinite.

Proof. We show that $\mathcal{O}_S$ has no finite non-projective $\mathcal{EL}$ approximation. Observe that the ontology $\mathcal{O}$ obtained from $\mathcal{O}_S$ by replacing the topmost CI with the infinite set

$$M = \{A' \sqcap \exists r^n.A \sqsubseteq M \mid n \geq 0\}$$

is an infinite non-projective $\mathcal{EL}$ approximation of $\mathcal{O}_S$. Now assume for a proof by contradiction that there exists a finite non-projective $\mathcal{EL}$ approximation of $\mathcal{O}_S$. Then, by compactness of reasoning in $\mathcal{EL}$, there exists a finite subset $\mathcal{O}'$ of $\mathcal{O}$ that is an $\mathcal{EL}$ approximation of $\mathcal{O}_S$. Let n be maximal such that $A' \sqcap \exists r^n.A \sqsubseteq M \in \mathcal{O}'$. Then $\mathcal{O}' \not\models A' \sqcap \exists r^{n+1}.A \sqsubseteq M$ and we have derived a contradiction. ❑

To prove Proposition 2 and 4, we use the following lemma from [Lutz and Wolter, 2010]. If C is an $\mathcal{EL}$ concept of the form $C_1 \sqcap \cdots \sqcap C_n$, $n \geq 1$, then the *top-level conjuncts* of C are $C_1, \ldots, C_n$.

Lemma 3. *Let $\mathcal{O}$ be an $\mathcal{EL}$ ontology and C, D be $\mathcal{EL}$ concepts. Then $\mathcal{O} \models C \sqsubseteq \exists r.D$ implies that*

1. *there exists a top-level conjunct $\exists r.C'$ of C such that $\mathcal{O} \models C' \sqsubseteq D$ or*
2. *there exists a $C' \in \mathtt{sub}(\mathcal{O})$ such that $\mathcal{O} \models C \sqsubseteq \exists r.C'$ and $\mathcal{O} \models C' \sqsubseteq D$.*

Proposition 2. *The $\mathcal{ELU}$ ontology*

$$\begin{array}{rrcll} \mathcal{O}_S = \{ & A & \sqsubseteq & B_1 \sqcup B_2, & \\ & \exists r.B_2 & \sqsubseteq & \exists r.(B_1 \sqcap L), & \\ & L & \sqsubseteq & \exists s.L & \} \end{array}$$

has no finite $\mathcal{EL}$ approximation.

Proof. Let $\mathcal{O}_T$ be a (potentially projective) $\mathcal{EL}$ approximation of $\mathcal{O}_S$. Then for all $n \geq 0$ and $m > n$, we have

(a) $\mathcal{O}_T \models \exists r.(A \sqcap \exists s^n.\top) \sqsubseteq \exists r.(B_1 \sqcap \exists s^n.\top)$ and

(b) $\mathcal{O}_T \not\models \exists r.(A \sqcap \exists s^n.\top) \sqsubseteq \exists r.(B_1 \sqcap \exists s^m.\top)$

since the same is true for $\mathcal{O}_S$. To establish the desired result, it suffices to argue that for every $n \geq 0$, there is a $C_n \in \mathtt{sub}(\mathcal{O}_T)$ such that $\mathcal{O}_T \models C_n \sqsubseteq B_1 \sqcap \exists s^n.\top$ and $\mathcal{O}_T \not\models C_n \sqsubseteq B_1 \sqcap \exists s^m.\top$ for any $m > n$. In fact, if this is the case, then $\mathcal{O}_T$ has infinitely many subconcepts and is thus infinite.

Let $n \geq 0$. First note that

(c) $\mathcal{O}_T \not\models A \sqcap \exists s^n.\top \sqsubseteq B_1$.

because the same is true for $\mathcal{O}_S$. It follows from (a), (c), and Lemma 3 that there exists a $C \in \mathtt{sub}(\mathcal{O}_T)$ such that $\mathcal{O}_T \models \exists r.(A \sqcap \exists s^n.\top) \sqsubseteq \exists r.C$ and $\mathcal{O}_T \models C \sqsubseteq B_1 \sqcap \exists s^n.\top$. Set $C_n = C$. By choice and by (b), C_n is as desired. ❑

Proposition 4. *Let $\mathcal{O}_S^n$ be obtained from the ontology $\mathcal{O}_S$ in Proposition 2 by replacing the bottommost CI with*

$$L \sqsubseteq A_1 \sqcap \hat{A}_1 \sqcap \cdots \sqcap A_n \sqcap \hat{A}_n \sqcap \exists r_1.L \sqcap \exists r_2.L$$

Then for all $n, \ell \geq 1$ and any ℓ-bounded $\mathcal{EL}$ approximation $\mathcal{O}_T$ of $\mathcal{O}_S^n$, $||\mathcal{O}_T|| \geq \mathtt{tower}(\ell, n)$.

Proof. Assume that a depth bound $\ell \geq 1$ is given. $\mathcal{EL}$ concepts C_1, C_2 are *incomparable* w.r.t. $\mathcal{O}_S$ if neither $\mathcal{O}_S \models C_1 \sqsubseteq C_2$ nor $\mathcal{O}_S \models C_2 \sqsubseteq C_1$. Take a set Ω of $\mathcal{EL}$ concepts of depth bounded by $\ell - 1$ that are pairwise incomparable w.r.t. $\mathcal{O}_S$ and use only the symbols $r_1, r_2, A_1, \hat{A}_1, \ldots, A_n, \hat{A}_n$. It is straightforward to construct such a set Ω and that has size at least $\mathtt{tower}(\ell, n)$. It then suffices to show that for every $E \in \Omega$ there exists a $C_E \in \mathtt{sub}(\mathcal{O}_T)$ such that $\mathcal{O}_T \models C_E \sqsubseteq E$ and $\mathcal{O}_T \not\models C_E \sqsubseteq E'$ for any $E' \in \Omega$ with $E' \neq E$.

Let $E \in \Omega$. Then

(a) $\mathcal{O}_T \models \exists r.(A \sqcap E) \sqsubseteq \exists r.(B_1 \sqcap E)$,

(b) $\mathcal{O}_T \not\models \exists r.(A \sqcap E) \sqsubseteq \exists r.(B_1 \sqcap E')$ for any $E' \in \Omega$ with $E' \neq E$, and

(c) $\mathcal{O}_T \not\models A \sqcap E \sqsubseteq B_1$.

since the same is true for $\mathcal{O}_S$. Thus, similarly to the proof of Proposition 2 we can show that must exist a $C \in \mathtt{sub}(\mathcal{O}_T)$ such that $\mathcal{O}_T \models \exists r.(A \sqcap E) \sqsubseteq \exists r.C$ and $\mathcal{O}_T \models C \sqsubseteq B_1 \sqcap E$ and use C as C_E. ❑

B Proof of Theorem 1

Lemma 1. *$\mathcal{O}_S$ is not finitely generating iff for every $n \geq 0$, there is an $\exists r.D \in \mathtt{sub}(\mathcal{O}_S)$ that occurs on the right-hand side of a CI and a sequence $r_1, \ldots, r_n$ of role names from $\mathcal{O}_S$ such that $\mathcal{O}_S \models D \sqsubseteq \exists r_1. \cdots .\exists r_n.\top$.*

Proof. Observe that the number of non-logically equivalent $\mathcal{EL}$ concepts over $\Sigma = \mathtt{sig}(\mathcal{O}_S)$ and of depth bounded by n is finite, for any natural number $n \geq 0$. Moreover, any two $\mathcal{EL}$ concepts of distinct depth are not logically equivalent. Thus, there are infinitely many non-logically equivalent $\mathcal{O}_S$-generatable $\mathcal{EL}$ concepts if, and only if, for every $n \geq 0$ there exists an $\mathcal{O}_S$-generatable $\mathcal{EL}$ concept of depth n. The latter holds if, and only if, for every $n \geq 0$ there exist role names $r_1, \ldots, r_n$ in $\mathcal{O}_S$ such that $\exists r_1. \cdots \exists r_n.\top$ is $\mathcal{O}_S$-generatable. ❑

Theorem 1. *It is decidable whether a given $\mathcal{ELU}$-ontology $\mathcal{O}_S$ is finitely generating.*

Proof. It follows from Lemma 1 that it suffices to decide whether there exists a bound $\ell \geq 0$ such that for every $\exists r.D \in \mathtt{sub}(\mathcal{O}_S)$ on the right hand side of a CI in $\mathcal{O}_S$ and any sequence $r_1, \ldots, r_n$ of role names in $\mathcal{O}_S$, if $\mathcal{O}_S \models D \sqsubseteq \exists r_1. \cdots \exists r_n.\top$, then $n \leq \ell$. We show that there exists such an ℓ if, and only if, there exists such an ℓ with $\ell \leq |\mathtt{sig}(\mathcal{O}_S)| \times 2^{2^{||\mathcal{O}_S||}}$. Then decidability follows directly. We use a straightforward pumping argument to show the claim. Assume that there are $n > |\mathtt{sig}(\mathcal{O}_S)| \times 2^{2^{||\mathcal{O}_S||}}$, $\exists r.D \in \mathtt{sub}(\mathcal{O}_S)$ on the right hand side of a CI in $\mathcal{O}_S$, and role names $r_1, \ldots, r_n$ in $\mathcal{O}_S$ with $\mathcal{O}_S \models D \sqsubseteq \exists r_1. \cdots \exists r_n.\top$.

We show that then there exists such a concept $\exists r.D$ and sequence of role names of length $n' > n$. An $\mathcal{O}_S$-*type* is a subset t of the closure under single negation of $\mathtt{sub}(\mathcal{O}_S)$ such that for any $C \in \mathtt{sub}(\mathcal{O}_S)$ either $C \in t$ or $\neg C \in t$ and there exists a model $\mathcal{I}$ of $\mathcal{O}_S$ and $d \in \Delta^\mathcal{I}$ with $d \in (\bigsqcap_{C \in t} C)^\mathcal{I}$. We identify an $\mathcal{O}_S$-type t with the concept $\bigsqcap_{C\in t} C$ and let $D(\mathcal{O}_S)$ be the set of disjunctions of $\mathcal{O}_S$-types (without repetitions). We show that there exists a sequence $X_1, \dots, X_n \in D(\mathcal{O}_S)$ such that

$$\mathcal{O}_S \models D \sqsubseteq \exists r_1.X_1, \quad \mathcal{O}_S \models X_i \sqsubseteq \exists r_{i+1}.X_{i+1},$$

for all $i < n$. The proof is as follows. Let X_1 be the set of all $\mathcal{O}_S$-types t such that there exist a model $\mathcal{I}$ of $\mathcal{O}_S$ and $d, e \in \Delta^\mathcal{I}$ with $(d, e) \in r_1^\mathcal{I}$, $d \in D^\mathcal{I}$ and $e \in t^\mathcal{I}$. Assume that X_i has been defined. Then X_{i+1} is the set of all $\mathcal{O}_S$-types t such that there exist a model $\mathcal{I}$ of $\mathcal{O}_S$ and $d, e \in \Delta^\mathcal{I}$ with $(d, e) \in r_{i+1}^\mathcal{I}$, $d \in X_i^\mathcal{I}$ and $e \in t^\mathcal{I}$. By definition

$$\mathcal{O}_S \models D \sqsubseteq \forall r_1.X_1, \quad \mathcal{O}_S \models X_i \sqsubseteq \forall r_{i+1}.X_{i+1},$$

and now one can readily show by induction on i, and using that $\mathcal{O}_S \models D \sqsubseteq \exists r_1. \cdots \exists r_n.\top$, that

$$\mathcal{O}_S \models D \sqsubseteq \exists r_1.X_1, \quad \mathcal{O}_S \models X_i \sqsubseteq \exists r_{i+1}.X_{i+1},$$

for all $i < n$. Thus, as $n > |\mathtt{sig}(\mathcal{O}_S)| \times 2^{2^{||\mathcal{O}_S||}}$, there exist $1 < i < j \leq n$ such that $r_i = r_j$ and $X_i = X_j$. But then

$$\mathcal{O}_S \models D \sqsubseteq \exists r_1. \cdots \exists r_{j-1}.\exists r_i. \cdots .\exists r_n.\top,$$

and we have found the sequence of role names of length $n + (j - i) > n$ we wanted. ❑

C Proof of Theorem 2

C.1 Preliminaries

We write $\mathcal{A} \models C(a)$ if $a \in C^\mathcal{I}$ where $\mathcal{I}$ is $\mathcal{A}$ viewed as an interpretation in the obvious way. An ABox $\mathcal{A}$ is *ditree-shaped* if the directed graph $G_\mathcal{A} = (\mathtt{Ind}(\mathcal{A}), \{(a, b) \mid r(a, b) \in \mathcal{A}\})$ is a tree and there are no multi-edges, that is, $r(a, b), s(a, b) \in \mathcal{A}$ implies $r = s$. Every $\mathcal{EL}$ concept C can be viewed as a ditree-shaped ABox $\mathcal{A}_C$ in an obvious way.

We will sometimes also use *extended ABoxes*, that is, ABoxes $\mathcal{A}$ that can also contain concept assertions of the form $C(a)$, C a compound concept. If all concepts that occur in such assertions are formulated in a description logic $\mathcal{L}$, we speak of *extended* $\mathcal{L}$*-ABoxes*. If $\mathcal{A}$ is an extended ABox, then we use $\mathcal{A}^-$ to denote the non-extended ABox obtained from $\mathcal{A}$ by removing all assertions $C(a)$ where C is not a concept name.

We next introduce a standard chase procedure for $\mathcal{EL}$ ontologies. The procedure uses ABoxes as a data structure. Let $\mathcal{O}$ be an $\mathcal{EL}$ ontology. There is a single chase rule that can be applied to an ABox $\mathcal{A}$:

- if $C \sqsubseteq D \in \mathcal{O}$ and $\mathcal{A} \models C(a)$, then a copy of $\mathcal{A}_D$ whose individuals are disjoint from those in $\mathcal{A}$ and replace $\mathcal{A}$ with the union of $\mathcal{A}$ and $\mathcal{A}_D$, identifying the root of the latter with a.

The chase starts with an ABox $\mathcal{A}_0$ and exhaustively applies the above rule in a fair way, resulting in sequence of ABoxes $\mathcal{A}_0, \mathcal{A}_1, \dots$. The *result* of the chase is the (potentially infinite) ABox $\bigcup_{i\geq 0} \mathcal{A}_i$ obtained in the limit, denoted $\mathtt{ch}_\mathcal{O}(\mathcal{A})$. The result is unique since the chase is oblivious, that is, a rule can applied to $C \sqsubseteq D$ and $C(a)$ even if $\mathcal{A} \models D(a)$ already holds. A proof of the following is standard and omitted.

Lemma 4. *$\mathcal{O} \models C \sqsubseteq D$ iff $\mathtt{ch}_\mathcal{O}(\mathcal{A}_C) \models D(a_0)$, for all $\mathcal{EL}$ concepts C and D.*

C.2 Main Proof

We start with soundness.

Lemma 5. *$\mathcal{O}_T^\omega \models C_0 \sqsubseteq D_0$ implies $\mathcal{O}_S \models C_0 \sqsubseteq D_0$ for all $\mathcal{EL}$ concepts C_0, D_0 over $\mathtt{sig}(\mathcal{O}_S)$.*

Proof. Assume that $\mathcal{O}_T^\omega \models C_0 \sqsubseteq D_0$ where C_0, D_0 are $\mathcal{EL}$ concepts over $\mathtt{sig}(\mathcal{O}_S)$. Then $\mathtt{ch}_{\mathcal{O}_T^\omega}(\mathcal{A}_{C_0}) \models D_0(a_0)$, a_0 the root of $\mathcal{A}_{C_0}$. Let $\mathcal{A}_{C_0} = \mathcal{A}_0, \mathcal{A}_1, \dots$ be a sequence of ABoxes produced by the $\mathcal{EL}$ chase of $\mathcal{A}_{C_0}$ with $\mathcal{O}_T^\omega$. Clearly, all ABoxes $\mathcal{A}_0, \mathcal{A}_1, \dots$ are ditree-shaped and can thus be viewed as an $\mathcal{EL}$ concept C_i. For an $\mathcal{EL}$ concept C over $\mathtt{sig}(\mathcal{O}_T^\omega)$, let $C^\downarrow$ be the $\mathcal{ELU}$ concept obtained from C by replacing every X_D with $D \in \mathtt{Dis}(\mathcal{O}_S)$. We prove the following by induction on i.

Claim. $\mathcal{O}_S \models C_i^\downarrow \sqsubseteq C_{i+1}^\downarrow$ for all $i \geq 0$.

To prove the claim, let $i \geq 0$. $\mathcal{A}_{i+1}$ was obtained from $\mathcal{A}_i$ by applying the chase rule. Thus let $C \sqsubseteq D \in \mathcal{O}_T^\omega$, $\mathcal{A}_i \models C(a)$, and let $\mathcal{A}_{i+1}$ be obtained from $\mathcal{A}_i$ by taking a copy of $\mathcal{A}_D$ whose individuals are disjoint from those in $\mathcal{A}_i$ and defining $\mathcal{A}_{i+1}$ as the union of $\mathcal{A}_i$ and $\mathcal{A}_D$, identifying the root of the latter with a. By definition of $\mathcal{O}_T^\omega$, we have $\mathcal{O}_S \models C^\downarrow \sqsubseteq D^\downarrow$. By construction of $\mathcal{A}_{i+1}$, we thus have $\mathcal{O}_S \models C_i^\downarrow \sqsubseteq C_{i+1}^\downarrow$ as required and thus the claim is proved.

From $\mathtt{ch}_{\mathcal{O}_T^\omega}(\mathcal{A}_{C_0}) \models D_0$, we obtain $\mathcal{A}_i \models D_0$ for some i. Since D_0 is over $\mathtt{sig}(\mathcal{O}_S)$, $\mathcal{A}_i \models D_0$ implies $\emptyset \models C_i^\downarrow \sqsubseteq D_0$. Together with the claim and since $C_0 = C_0^\downarrow$, this gives $\mathcal{O}_S \models C_0 \sqsubseteq D_0$. ❑

We now address completeness, starting with the essential Lemma 6 below. Preparing for the case of $\mathcal{ELU}_\bot$-to-$\mathcal{EL}_\bot$ approximations, we state and prove the lemma directly for this case. This requires a few preliminaries.

Let $\mathcal{O}$ be an $\mathcal{ELU}_\bot$ ontology. For every $\mathcal{EL}$ concept C, we define $\mathtt{Dis}_\mathcal{O}(C)$ as in the case without $\bot$. This can now be the empty disjunction, which we identify with $\bot$. In fact, C is satisfiable w.r.t. $\mathcal{O}$ if and only if $\mathtt{Dis}_\mathcal{O}(C) = \bot$. We set $\bot^\uparrow = \bot$. We further associate with every $\mathcal{EL}$ concept C a disjunction $\mathtt{Dis}_\mathcal{O}^{\mathcal{EL}}(C)$ that contains a disjunct $\bigsqcap S$ for every set $S \subseteq \mathtt{sub}^-(\mathcal{O})$ such that there is a model $\mathcal{I}$ of $\mathcal{O}$ and a $d \in C^\mathcal{I}$ with

$$S = \{E \in \mathtt{sub}^-(\mathcal{O}) \mid d \in E^\mathcal{I} \text{ and } E \text{ is an } \mathcal{EL} \text{ concept}\}$$

while this is not true for any proper subset of S. If $\mathtt{Dis}_\mathcal{O}^{\mathcal{EL}}(C)$ consists of a single disjunct that is the empty conjunction, we identify it with $\top$. The empty disjunction is again identified with $\bot$.

For the following lemma, we assume that $\mathcal{O}_S$ is an $\mathcal{ELU}_\bot$ ontology. The lemma refers to $\mathcal{O}_T^-$. Note that when $\mathcal{O}_S$ is formulated in $\mathcal{ELU}$, then $\mathcal{O}_T^-$ consists of all instantiations of the first three lines of Figure 2 and that for $\mathcal{ELU}_\bot$, the same is true for Figure 3. However, the first three lines of these figures are identical.

Lemma 6. *$\mathcal{O}_T^- \models C_0 \sqsubseteq \mathtt{Dis}^{\mathcal{EL}}_{\mathcal{O}_S}(C_0)^\uparrow$ for every $\mathcal{EL}$ concept C_0 over $\mathtt{sig}(\mathcal{O}_S)$.*

We prove Lemma 6 by first introducing a special chase procedure for $\mathcal{ELU}_\bot$ ontologies that is specifically tailored towards our approximations. Unlike more standard chase procedures for $\mathcal{ELU}_\bot$, our chase is deterministic rather than disjunctive.

We define an entailment notion $\mathcal{A} \vdash C(a)$ between extended $\mathcal{ELU}_\bot$ ABoxes $\mathcal{A}$ and $\mathcal{ELU}_\bot$ concepts C as follows:

- $\mathcal{A} \vdash \top(a)$ always holds;
- $\mathcal{A} \vdash \bot(a)$ if $\bot(b) \in \mathcal{A}$ for some b;
- $\mathcal{A} \vdash A(a)$ if $A(a) \in \mathcal{A}$;
- $\mathcal{A} \vdash C_1 \sqcap C_2(a)$ if $\mathcal{A} \vdash C_1(a)$ and $\mathcal{A} \vdash C_2(a)$;
- $\mathcal{A} \vdash C_1 \sqcup C_2(a)$ if $C_1 \sqcup C_2(a) \in \mathcal{A}$;
- $\mathcal{A} \vdash \exists r.C(a)$ if there is b such that $r(a,b) \in \mathcal{A}$ and $\mathcal{A} \vdash C(b)$.

Note that if C is an $\mathcal{EL}$ concept, then $\mathcal{A} \vdash C(a)$ if $a \in C^{\mathcal{I}_{\mathcal{A}^-}}$ where $\mathcal{I}_{\mathcal{A}^-}$ is $\mathcal{A}^-$ viewed as an interpretation in the obvious way. Let $\mathtt{Dis}_{\mathcal{O}}(C)$ be defined in the same way as $\mathtt{Dis}^{\mathcal{EL}}_{\mathcal{O}}(C)$ except that all concepts in $\mathtt{sub}(\mathcal{O})$ that are concept names or of the form $\exists r.E$ are considered instead of only $\mathcal{EL}$ concepts of this form.

Let $\mathcal{A}$ be an ABox and $\mathcal{O}$ an $\mathcal{ELU}_\bot$ ontology. The chase produces produces a sequence of ABoxes $\mathcal{A} = \mathcal{A}_0, \mathcal{A}_1, \mathcal{A}_2, \cdots$ such that $\mathcal{A}_i \subseteq \mathcal{A}_{i+1}$ for all $i \geq 0$. Although different sequences can be produced, the limit $\bigcup_{i\geq 0} \mathcal{A}_i$ will be unique and we call it the *result* of chasing $\mathcal{A}$ with $\mathcal{O}$. We call an individual in $\bigcup_{i\geq 0} \mathcal{A}_i$ *original* if it already occurs in $\mathcal{A}$ and *anonymous* otherwise. In the ABoxes $\mathcal{A}_i$, anonymous individuals can be marked or not. Each ABox $\mathcal{A}_{i+1}$ is obtained from $\mathcal{A}_i$ by *chasing a single step* with $\mathcal{O}$, that is, $\mathcal{A}_{i+1}$ is obtained from $\mathcal{A}_i$ in one of the following ways:

1. choose $C \sqsubseteq D \in \mathcal{O}$ and $a \in \mathtt{Ind}(\mathcal{A})$ with $\mathcal{A} \vdash C(a)$ and add $\mathtt{DNF}(D)(a)$;
2. choose $C_1 \sqcap C_2(a) \in \mathcal{A}$ and add $C_1(a), C_2(a)$;
3. choose $\exists r.C(a) \in \mathcal{A}$ and add $r(a,b), C(b)$ for a fresh b; we say that b was *introduced for* C;
4. choose $D_1(a) \in \mathcal{A}$ with $D_1 \in \mathtt{Dis}^-(\mathcal{O})$ and $D_2, D_3 \in \mathtt{Dis}(\mathcal{O})$ such that $\mathcal{A} \vdash D_2(a)$ and $\mathcal{O} \models D_1 \sqcap D_2 \sqsubseteq D_3$, and add $D_3(a)$;
5. choose $r(a,b), D_1(b) \in \mathcal{A}$ with $D_1 \in \mathtt{Dis}^-(\mathcal{O})$ and $D_2 \in \mathtt{Dis}(\mathcal{O})$ such that $\mathcal{O} \models \exists r.D_1 \sqsubseteq D_2$ and add $D_2(a)$;
6. choose $D(a) \in \mathcal{A}$ with $D \in \mathtt{Dis}^-(\mathcal{O})$ and a anonymous and introduced for C, and add $\mathtt{Dis}_{\mathcal{O}}(C)(a)$; mark a;
7. choose $r(a,b) \in \mathcal{A}$ with a anonymous and introduced for C_a and b marked, anonymous, and introduced for C_b such that $\mathtt{Dis}_{\mathcal{O}}(\exists r.C_b) \in \mathtt{Dis}^-(\mathcal{O}_S)$; add $\mathtt{Dis}_{\mathcal{O}}(C_a)(a)$; mark a if it is anonymous.

Note that Rules 1-3 implement Line 1 of our $\mathcal{EL}$ approximations $\mathcal{O}_T^\ell$ while Rules 4 and 5 correspond to Lines 2 and 3 of the approximation. Rules 6-7 are there to deal with anonymous individuals which behave in a different way than original ones.

We require that the chase is *fair*, that is, every possible way to chase a single step is eventually used. Note that our chase is oblivious, that is, a chase rule can be applied even if its 'consequence' is already there. This implies that the results of the chase, which we denote with $\mathtt{ch}^{\mathrm{sp}}_{\mathcal{O}}(\mathcal{A})$, is unique up to isomorphism.

The main property that we require of the chase is the following completeness property.

Lemma 7. *Let $\mathcal{O}$ be an $\mathcal{ELU}_\bot$ ontology and C_0 an $\mathcal{EL}$ concept over $\mathtt{sig}(\mathcal{O}_S)$. Then $\mathtt{ch}^{sp}_{\mathcal{O}}(\mathcal{A}_{C_0}) \vdash \mathtt{Dis}^{\mathcal{EL}}_{\mathcal{O}}(C_0)(a_0)$.*

Since the proof of Lemma 7 is rather laborious, we defer it to Section C.3.

We now return to the proof of Lemma 6. Let $\mathcal{A}_0, \mathcal{A}_1, \ldots$ be a sequence of ABoxes generated by chasing $\mathcal{A}_{C_0} = \mathcal{A}_0$ with $\mathcal{O}_S$ using the special chase introduced above. It is easy to see that all extended ABoxes $\mathcal{A}_0, \mathcal{A}_1, \ldots$ are ditree-shaped and can thus be viewed as $\mathcal{ELU}_\bot$ concepts $C_0, C_1, \ldots$ in which all disjunctions are from $\mathtt{Dis}(\mathcal{O}_S)$. Note that also the ABox assertions $C(a) \in \mathcal{A}_i$ with C a compound concept or $\bot$ give raise to subconcepts in C_i.

Claim. $\mathcal{O}_T^- \models C_i^\uparrow \sqsubseteq C_{i+1}^\uparrow$, for all $i \geq 0$.

To prove the claim, let $i \geq 0$. We make a case distinction according to the chase rule with which $\mathcal{A}_{i+1}$ is obtained from $\mathcal{A}_i$:

1. Then there is a $C \sqsubseteq D \in \mathcal{O}_S$ and an $a \in \mathtt{Ind}(\mathcal{A})$ such that $\mathcal{A}_i \vdash C(a)$ and $\mathcal{A}_{i+1} = \mathcal{A}_i \cup \{\mathtt{DNF}(D)(a)\}$. Let E_a be the subconcept of C_i that corresponds to the subtree rooted at a in $\mathcal{A}_i$ and let F_a be the subconcept of C_{i+1} that corresponds to the subtree rooted at a in $\mathcal{A}_{i+1}$. Since C is an $\mathcal{EL}$ concept, $\mathcal{A}_i \vdash C(a)$ implies $\mathcal{A}_i^- \vdash C(a)$. Consequently, $\emptyset \models E_a^\uparrow \sqsubseteq C$. Moreover, $F_a = E_a \sqcap \mathtt{DNF}(D)$ and $\mathcal{O}_T^-$ contains the CI $C \sqsubseteq \mathtt{DNF}(D)^\uparrow$, thus $\mathcal{O}_T^- \models C_i^\uparrow \sqsubseteq C_{i+1}^\uparrow$ as required.
2. Trivial.
3. Trivial.
4. Then there are $D_1(a) \in \mathcal{A}_i$ with $D_1 \in \mathtt{Dis}^-(\mathcal{O}_S)$ and $D_2, D_3 \in \mathtt{Dis}(\mathcal{O}_S)$ such that $\mathcal{A}_i \vdash D_2(a)$, $\mathcal{O}_S \models D_1 \sqcap D_2 \sqsubseteq D_3$, and $\mathcal{A}_{i+1} = \mathcal{A}_i \cup \{D_3(a)\}$. Let E_a be the subconcept of C_i that corresponds to the subtree rooted at a in $\mathcal{A}_i$ and let F_a be the subconcept of C_{i+1} that corresponds to the subtree rooted at a in $\mathcal{A}_{i+1}$. Then $F_a = E_a \sqcap D_3$. From $D_1(a) \in \mathcal{A}$ and $D_1 \in \mathtt{Dis}^-(\mathcal{O}_S)$, we obtain that X_{D_1} is a top-level conjunct of $E_a^\uparrow$. From $\mathcal{A}_i \vdash D_2(a)$, we obtain that $\emptyset \models E_a^\uparrow \sqsubseteq D_2^\uparrow$. Moreover, $\mathcal{O}_T^-$ contains the CI $X_{D_1} \sqcap D_2^\uparrow \sqsubseteq D_3^\uparrow$, and thus $\mathcal{O}_T^- \models C_i^\uparrow \sqsubseteq C_{i+1}^\uparrow$ as required.

5. Similar to the previous case, using the third line of $\mathcal{O}_T^-$.
6. Then there is a $D(a) \in \mathcal{A}_i$ with $D \in \mathtt{Dis}^-(\mathcal{O}_S)$, a anonymous and introduced for C, and $\mathcal{A}_{i+1} = \mathcal{A}_i \cup \{\mathtt{Dis}_{\mathcal{O}_S}(C)(a)\}$. Let E_a be the subconcept of C_i that corresponds to the subtree rooted at a in $\mathcal{A}_i$ and let F_a be the subconcept of C_{i+1} that corresponds to the subtree rooted at a in $\mathcal{A}_{i+1}$. Since a was introduced for C, $C(a) \in \mathcal{A}_i$ and thus $\emptyset \models E_a \sqsubseteq C$. Since $C \in \mathtt{Dis}(\mathcal{O}_S)$ and $\mathcal{O}_S \models C \sqsubseteq \mathtt{Dis}_{\mathcal{O}_S}(C)$, $\mathcal{O}_T^-$ contains the CI $X_D \sqcap C^\uparrow \sqsubseteq \mathtt{Dis}_{\mathcal{O}_S}(C)$. Moreover, $F_a = E_a \sqcap \mathtt{Dis}_{\mathcal{O}_S}(C)$. It follows that $\mathcal{O}_T^- \models C_i^\uparrow \sqsubseteq C_{i+1}^\uparrow$.
7. Similar to the previous case.

This finishes the proof of the claim.

By Lemma 7, $\mathtt{ch}^{\mathrm{sp}}_{\mathcal{O}_S}(\mathcal{A}_{C_0}) \vdash \mathtt{Dis}^{\mathcal{EL}}_{\mathcal{O}_S}(C_0)(a_0)$ and thus $\mathcal{A}_i \vdash \mathtt{Dis}^{\mathcal{EL}}_{\mathcal{O}_S}(C_0)(a_0)$ for some i. First assume that $\mathtt{Dis}^{\mathcal{EL}}_{\mathcal{O}_S}(C_0)$ contains more than one disjunct. Then, by definition of $\vdash$, $\mathtt{Dis}^{\mathcal{EL}}_{\mathcal{O}_S}(C_0)(a_0) \in \mathcal{A}_i$ and thus $X_{\mathtt{Dis}^{\mathcal{EL}}_{\mathcal{O}_S}(C_0)} = \mathtt{Dis}^{\mathcal{EL}}_{\mathcal{O}_S}(C_0)^\uparrow$ is a top-level conjunct of $C_i^\uparrow$ implying $\emptyset \models C_i^\uparrow \sqsubseteq X^{\mathcal{EL}}_{\mathtt{Dis}_{\mathcal{O}_S}(C_0)}$. From the claim and $C_0 = C_0^\uparrow$, we obtain $\mathcal{O}_T^- \models C_0 \sqsubseteq C_i^\uparrow$ and are done. Now assume that $\mathtt{Dis}^{\mathcal{EL}}_{\mathcal{O}_S}(C_0)$ contains a single disjunct. Then $\mathcal{A}_i \vdash K$ for each conjunct K of $\mathtt{Dis}^{\mathcal{EL}}_{\mathcal{O}_S}(C_0)$. By definition of '$\vdash$' and $C_i^\uparrow$, it follows that $\emptyset \models C_i^\uparrow \sqsubseteq K^\uparrow$ for each such K, and thus $\emptyset \models C_i^\uparrow \sqsubseteq \mathtt{Dis}^{\mathcal{EL}}_{\mathcal{O}_S}(C_0)^\uparrow$. It again remains to apply the claim. Finally assume that $\mathtt{Dis}^{\mathcal{EL}}_{\mathcal{O}_S}(C_0) = \bot$. Then $\bot(b) \in \mathcal{A}_i$ for some b and thus $\emptyset \models C_i^\uparrow \sqsubseteq \bot$. We can once more apply the claim. This finishes the proof of Lemma 6.

Now back to the proof of completeness, that is, of Lemma 2. We need some more preliminaries.

Lemma 8. *Let $\mathcal{O}$ be an $\mathcal{ELU}$ ontology and $C, \exists r.D$ $\mathcal{EL}$ concepts. If $\mathcal{O} \models C \sqsubseteq \exists r.D$ and C contains no top-level conjunct $\exists r.C'$ such that $\mathcal{O} \models C' \sqsubseteq D$, then D is $\mathcal{O}$-generatable.*

Proof. Assume $\mathcal{O} \models C \sqsubseteq \exists r.D$ and C contains no top-level conjunct $\exists r.C'$ such that $\mathcal{O} \models C' \sqsubseteq D$. Assume D is not $\mathcal{O}$-generatable. Let

$$C = A_1 \sqcap \cdots \sqcap A_n \sqcap \exists r_1.E_1 \sqcap \cdots \sqcap \exists r_m.E_m.$$

In $\mathcal{A}_C$, the root a_0 has outgoing edges $r_1(a_0, b_1), \ldots, r_m(a_0, b_m)$. Extend $\mathcal{A}_C$ to a model $\mathcal{I}$ as follows:

1. add for any b_i, $1 \leq i \leq m$, a ditree-shaped model $\mathcal{I}_{b_i}$ of $\mathcal{O}$ with root b_i such that $b_i \in E_i^{\mathcal{I}_{b_i}}$ and $b_i \notin D^{\mathcal{I}_{b_i}}$;
2. add for any $\mathcal{ELU}$ concept $\exists r.E$ such that there is a CI $C' \sqsubseteq D$ in $\mathcal{O}$ such that D contains $\exists r.E$ as a top-level conjunct an r-successor $a_{r,E}$ of a_0 and a ditree-shaped model $\mathcal{I}_{r,E}$ of $\mathcal{O}$ with root $a_{r,E}$ such that $a_{r,E} \in E^{\mathcal{I}_{r,E}}$ and $a_{r,E} \notin D^{\mathcal{I}_{b_i}}$;
3. a_0 to $A^\mathcal{I}$ for any concept name A.

Note that the interpretations $\mathcal{I}_{b_i}$ exist since C contains no top-level conjunct $\exists r.C'$ such that $\mathcal{O} \models C' \sqsubseteq D$ and the interpretations $\mathcal{I}_{r,E}$ exist since we assume that D is not $\mathcal{O}$-generatable. By construction, $a_0 \notin (\exists r.D)^\mathcal{I}$ and $\mathcal{I}$ is a model of $\mathcal{O}$ as all nodes distinct from a_0 clearly satisfy all CIs in $\mathcal{O}$ and a_0 satisfies all CIs in $\mathcal{O}$ by construction. We have derived a contradiction to $\mathcal{O} \models C \sqsubseteq \exists r.D$ as $a_0 \in C^\mathcal{I}$. ❑

For a ditree-shaped ABox $\mathcal{A}$ and $k \geq 0$, we use $\mathcal{A}|_k$ to denote the result of removing from $\mathcal{A}$ all individuals on levels larger than k and $C^a_\mathcal{A}$ to denote the subABox of $\mathcal{A}$ rooted at a viewed as an $\mathcal{EL}$ concept. To prepare for the case of $\mathcal{ELU}_\bot$-to-$\mathcal{EL}$ approximations, we establish the following lemma directly for $\mathcal{ELU}_\bot$ instead of for $\mathcal{ELU}$.

Lemma 9. *Let $\mathcal{O}$ be an $\mathcal{ELU}_\bot$ ontology such that all concepts on the left hand side of CIs in $\mathcal{O}$ are $\mathcal{EL}$ concepts. Let $\mathcal{A}$ be a ditree-shaped ABox with root a_0 such that $\mathcal{O}, \mathcal{A} \models \exists r.C(a_0)$, C an $\mathcal{EL}$ concept of depth k. Let $\mathcal{A}^\pm$ be the extended ABox obtained from $\mathcal{A}|_k$ by adding $\mathtt{Dis}^{\mathcal{EL}}_\mathcal{O}(C^a_\mathcal{A})$ whenever a is a leaf in $\mathcal{A}|_k$. Then $\mathcal{O}, \mathcal{A}^\pm \models \exists r.C(a_0)$.*

Proof. Assume that $\mathcal{O}, \mathcal{A}^\pm \not\models \exists r.C(a_0)$. Take a ditree shaped model $\mathcal{I}$ of $\mathcal{O}$ and $\mathcal{A}^\pm$ with $a_0 \notin (\exists r.C)^\mathcal{I}$. Let a be a node of depth k in $\mathcal{A}$. We have $a \in \mathtt{Dis}^{\mathcal{EL}}_\mathcal{O}(C^a_\mathcal{A})^\mathcal{I}$ and thus there is a disjunct D of $\mathtt{Dis}^{\mathcal{EL}}_\mathcal{O}(C^a_\mathcal{A})$ with $a \in D^\mathcal{I}$. Let $\mathcal{A}_a$ be the subABox of $\mathcal{A}$ rooted at a. Observe that $\mathcal{A}_a$ is satisfiable w.r.t. $\mathcal{O}$: otherwise $\bot$ is the only disjunct of $\mathtt{Dis}^{\mathcal{EL}}_\mathcal{O}(C^a_\mathcal{A})$ and so $\mathcal{A}^\pm$ is not satisfiable. Thus $\mathcal{O}, \mathcal{A}^\pm \models \exists r.C(a_0)$, and we have derived a contraction. As $\mathcal{A}_a$ is satisfiable w.r.t. $\mathcal{O}$ we obtain by definition of $\mathtt{Dis}^{\mathcal{EL}}_\mathcal{O}(C^a_\mathcal{A})$ that there is a model $\mathcal{J}_a$ of $\mathcal{O}$ and $\mathcal{A}_a$ such that whenever $a \in E^{\mathcal{J}_a}$ for some $\mathcal{EL}$ concept $E \in \mathtt{sub}^-(\mathcal{O})$, then $a \in E^\mathcal{I}$. Construct a new interpretation $\mathcal{I}'$ by adding to $\mathcal{I}$ the interpretation $\mathcal{J}_a$, for all nodes a of depth k in $\mathcal{A}$ (where $\mathcal{I}$ and $\mathcal{J}_a$ only share a). $\mathcal{I}'$ is a model of $\mathcal{O}$ and $\mathcal{A}$ since $a \in E^\mathcal{I}$ if $a \in E^{\mathcal{I}'}$, for all $\mathcal{EL}$ concepts $E \in \mathtt{sub}^-(\mathcal{O})$ and a of depth k in $\mathcal{A}$. Moreover, $a_0 \notin (\exists r.C)^{\mathcal{I}'}$, as required. ❑

We are now in a position to prove Lemma 2.

Lemma 2. *Let $\ell \in \mathbb{N} \cup \{\omega\}$. Then $\mathcal{O}_S \models C_0 \sqsubseteq D_0$ implies $\mathcal{O}_T^\ell \models C_0 \sqsubseteq D_0$ for all $\mathcal{EL}$ concepts C_0, D_0 over $\mathtt{sig}(\mathcal{O}_S)$ such that the role depth of D_0 is bounded by ℓ.*

Proof. Assume that $\mathcal{O}_S \models C_0 \sqsubseteq D_0$ with C_0, D_0 $\mathcal{EL}$ concepts over $\mathtt{sig}(\mathcal{O}_S)$ such that the role depth of D_0 is bounded by ℓ. It clearly suffices to consider the cases where D_0 is a concept name and where it is of the form $\exists r.E_0$.

We start with the former, so let $D_0 = A$. Clearly, $\mathcal{O}_S \models C_0 \sqsubseteq A$ implies $\mathcal{O}_S \models \mathtt{Dis}^{\mathcal{EL}}_{\mathcal{O}_S}(C_0) \sqsubseteq A$. It thus follows from Lemma 6 that $\mathcal{O}_T^\ell \models C_0 \sqsubseteq A$. To see this, first assume that $\mathtt{Dis}^{\mathcal{EL}}_{\mathcal{O}_S}(C_0)$ contains a single disjunct. Then A must be a conjunct of $\mathtt{Dis}^{\mathcal{EL}}_{\mathcal{O}_S}(C_0)$ and it suffices to apply Lemma 6. Now assume that $\mathtt{Dis}^{\mathcal{EL}}_{\mathcal{O}_S}(C_0)$ has more than one disjunct, that is, it is in $\mathtt{Dis}^-(\mathcal{O}_S)$. Then $\mathcal{O}_T^\ell$ contains the CI $X_{\mathtt{Dis}^{\mathcal{EL}}_{\mathcal{O}_S}(C_0)} \sqcap X_{\mathtt{Dis}^{\mathcal{EL}}_{\mathcal{O}_S}(C_0)} \sqsubseteq A$ (second line of Figure 2) and thus it again suffices to apply Lemma 6.

The case where $D_0 = \exists r.E_0$ is a consequence of the following claim. For each $a \in \mathtt{Ind}(\mathcal{A}_{C_0})$, we write C_0^a as an abbreviation for $C^a_{\mathcal{A}_{C_0}}$.

Claim. For all $a \in \mathtt{Ind}(\mathcal{A}_{C_0})$ and $\mathcal{EL}$ concepts $\exists r.E$ of depth $\ell - \mathtt{depth}(a)$, $\mathcal{O}_S \models C_0^a \sqsubseteq \exists r.E$ implies $\mathcal{O}_T^\ell \models C_0^a \sqsubseteq \exists r.E$.

Proof of claim. The proof is by induction on the co-depth of a.

Induction start. Then a is a leaf in $\mathcal{A}_{C_0}$ and thus C_0^a does not have any top-level conjuncts of the form $\exists r.E'$. Lemma 8 thus yields that E is $\mathcal{O}_\mathcal{S}$-generatable. Thus $C_0^a \sqsubseteq \exists r.E$ is a CI in $\mathcal{O}_T^\ell$.

Induction step. Then a is a non-leaf in $\mathcal{A}_{C_0}$. We distinguish two cases.

Case 1. There is a top-level conjunct $\exists r.E'$ in C_0^a such that $\mathcal{O}_\mathcal{S} \models E' \sqsubseteq E$. Then a has an r-successor b in $\mathcal{A}_{C_0}$ such that $C_0^b = E'$. Let

$$E = A_1 \sqcap \cdots \sqcap A_n \sqcap \exists r_1.E_1 \sqcap \cdots \sqcap \exists r_m.E_m.$$

Since we have already shown Lemma 2 for the case where D_0 is a concept name, we obtain $\mathcal{O}_T^\ell \models C_0^b \sqsubseteq A_i$ for $1 \leq i \leq n$. From the induction hypothesis, we further obtain $\mathcal{O}_T^\ell \models C_0^b \sqsubseteq \exists r_i.E_i$ for $1 \leq i \leq m$. Thus $\mathcal{O}_T^\ell \models C_0^b \sqsubseteq E$ and consequently $\mathcal{O}_T^\ell \models C_0^a \sqsubseteq \exists r.E$ as required.

Case 2. There is no top-level conjunct $\exists r.E'$ in C_0^a such that $\mathcal{O}_\mathcal{S} \models E' \sqsubseteq E$. Then Lemma 8 yields that E is $\mathcal{O}_\mathcal{S}$-generatable. Let $\mathcal{A}$ be the ditree-shaped subABox of $\mathcal{A}_{C_0}$ rooted at a and let $\mathcal{A}^\pm$ be the extended ABox obtained from $\mathcal{A}|_k$, with k the depth of $\exists r.E$, by adding $\mathtt{Dis}^{\mathcal{EL}}_{\mathcal{O}_S}(C^c_\mathcal{A})^\uparrow(c)$ whenever c is a leaf in $\mathcal{A}|_k$. Applying Lemma 9 to $\mathcal{A}$ and $\mathcal{A}^\pm$ and with $\exists r.E$ in place of $\exists r.C$, we obtain $\mathcal{O}_S, \mathcal{A}^\pm \models \exists r.E(a)$. Let $C^\pm$ be $\mathcal{A}^\pm$ viewed as an $\mathcal{ELU}$ concept. Then $\mathcal{O}_S \models C^\pm \sqsubseteq \exists r.E$. Since E is $\mathcal{O}_S$-generatable, $(C^\pm)^\uparrow \sqsubseteq \exists r.E$ is thus a CI in $\mathcal{O}_T^\ell$. We next observe that, by Lemma 6, $\mathcal{O}_T^\ell \models C^c_\mathcal{A} \sqsubseteq \mathtt{Dis}^{\mathcal{EL}}_{\mathcal{O}_S}(C^c_\mathcal{A})^\uparrow$ and thus $\mathcal{O}_T^\ell, \mathcal{A} \models \mathtt{Dis}^{\mathcal{EL}}_{\mathcal{O}_S}(C^c_\mathcal{A})^\uparrow(c)$ for all leaves c in $\mathcal{A}|_k$. Together with the construction of $\mathcal{A}^\pm$ and $C^\pm$, this yields that $\mathcal{O}_T^\ell \models C_0^a \sqsubseteq (C^\pm)^\uparrow$. Together with $(C^\pm)^\uparrow \sqsubseteq \exists r.E$ being a CI in $\mathcal{O}_T^\ell$, we obtain $\mathcal{O}_T^\ell \models C_0^a \sqsubseteq \exists r.E$ as required. ❏

C.3 Soundness and Completeness of the Special Chase

Our main aim is to establish Lemma 7. We start, however, with proving soundness of the chase. While this is interesting in its own right, we are not going to use it directly in the context of approximations. It is, however, an ingredient to the subsequent completeness proof.

Lemma 10. *Let C_0 be an $\mathcal{EL}$ concept and $\mathcal{O}$ an $\mathcal{ELU}_\bot$ ontology. Then $\mathtt{ch}^{sp}_\mathcal{O}(\mathcal{A}_{C_0}) \vdash D(a_0)$ implies $\mathcal{O} \models C_0 \sqsubseteq D$ for all $D \in \mathtt{Dis}(\mathcal{O})$.*

Proof. Let $\mathcal{A}_{C_0} = \mathcal{A}_0, \mathcal{A}_1, \dots$ be a sequence generated by chasing $\mathcal{A}_{C_0}$ with $\mathcal{O}$ using the special chase. Further, let $\mathcal{I}$ be a model of $\mathcal{O}$ and let $d \in C_0^\mathcal{I}$. An *extended homomorphism* from $\mathcal{A}_i$ to $\mathcal{I}$ is a function $h : \mathtt{Ind}(\mathcal{A}) \to \Delta^\mathcal{I}$ such that

1. $C(a) \in \mathcal{A}_i$, C potentially compound, implies $h(a) \in C^\mathcal{I}$ and
2. $r(a, b) \in \mathcal{A}_i$ implies $(h(a), h(b)) \in r^\mathcal{I}$.

We next observe the following.

Claim. if $\mathcal{A}_i \vdash D(a)$, $a \in \mathtt{Ind}(\mathcal{A}_i)$ and $D \in \mathtt{Dis}(\mathcal{O})$, and h is an extended homomorphism from $\mathcal{A}_i$ to $\mathcal{I}$, then $h(a) \in D^\mathcal{I}$.

The claim can be proved by induction on the structure of D. If D takes the form $D_1 \sqcap D_2$ or $\exists r.D_1$, then this is straightforward using the semantics and induction hypothesis. If D is $\top$, $\bot$, a concept name, or of the form $D_1 \sqcup D_2$ (note that in the latter case $\mathcal{A}_i \vdash D(a)$ implies $D(a) \in \mathcal{A}_i$), then this is immediate by definition of extended homomorphisms.

We show by induction on i that for each $i \geq 0$, there is an extended homomorphism h_i from $\mathcal{A}_i$ to $\mathcal{I}$ with $h(a_0) = d$. This is trivial for $i = 0$ since $d \in C_0^\mathcal{I}$. For $i \geq 0$, we make a case distinction according to the rule that was applied to obtain $\mathcal{A}_{i+1}$ from $\mathcal{A}_i$:

1. Then there is a $C \sqsubseteq D \in \mathcal{O}$ and an $a \in \mathtt{Ind}(\mathcal{A})$ such that $\mathcal{A}_i \vdash C(a)$ and $\mathcal{A}_{i+1} = \mathcal{A}_i \cup \{\mathtt{DNF}(D)(a)\}$. By the claim, $\mathcal{A}_i \vdash C(a)$ implies $h_i(a) \in C^\mathcal{I}$. Since $\mathcal{I}$ is a model of $\mathcal{O}$, $h_i(a) \in D^\mathcal{I} = \mathtt{DNF}(D)^\mathcal{I}$ and consequently h_i can be extended to an extended homomorphism h_{i+1} from $\mathcal{A}_{i+1}$ to $\mathcal{I}$.
2. Trivial.
3. Trivial.
4. Then there are $D_1(a) \in \mathcal{A}_i$ with $D_1 \in \mathtt{Dis}^-(\mathcal{O})$ and $D_2, D_3 \in \mathtt{Dis}(\mathcal{O})$ such that $\mathcal{A}_i \vdash D_2(a)$, $\mathcal{O} \models D_1 \sqcap D_2 \sqsubseteq D_3$, and $\mathcal{A}_{i+1} = \mathcal{A}_i \cup \{D_3(a)\}$. From $D_1(a) \in \mathcal{A}_i$, $\mathcal{A}_i \vdash D_2(a)$, and the claim, we get $h_i(a) \in (D_1 \sqcap D_2)^\mathcal{I}$. Since $\mathcal{I}$ is a model of $\mathcal{O}$, $h_i(a) \in D_3(a)$. Thus, we can choose $h_{i+1} = h_i$.
5. Similar to the previous case.
6. Then there is a $D(a) \in \mathcal{A}_i$ with $D \in \mathtt{Dis}^-(\mathcal{O})$, a anonymous and introduced for C, and $\mathcal{A}_{i+1} = \mathcal{A}_i \cup \{\mathtt{Dis}_\mathcal{O}(C)(a)\}$. We have $C(a) \in \mathcal{A}_i$, and thus $h(a) \in C^\mathcal{I}$. Since $\mathcal{I}$ is a model of $\mathcal{O}$, this implies $h(a) \in \mathtt{Dis}_\mathcal{O}(C)^\mathcal{I}$ and thus we can choose $h_{i+1} = h_i$.
7. Similar to the previous case.

We now finish the proof of Lemma 10. Let $\mathtt{ch}^{\mathrm{sp}}_\mathcal{O}(\mathcal{A}_{C_0}) \vdash D(a_0)$ with $D \in \mathtt{Dis}(\mathcal{O})$. Then there is an $\mathcal{A}_i$ with $\mathcal{A}_i \vdash D(a_0)$. From the claim, we obtain $d = h_i(a_0) \in C^\mathcal{I}$. Since this holds for all $\mathcal{I}$ and d, we have shown that $\mathcal{O} \models C_0 \sqsubseteq D$, as required. ❏

Lemma 7. *Let $\mathcal{O}$ be an $\mathcal{ELU}_\bot$ ontology and C_0 an $\mathcal{EL}$ concept over $\mathtt{sig}(\mathcal{O}_S)$. Then $\mathtt{ch}^{sp}_\mathcal{O}(\mathcal{A}_{C_0}) \vdash \mathtt{Dis}^{\mathcal{EL}}_\mathcal{O}(C_0)(a_0)$.*

Proof. We start with a special case, which is that $\bot(b) \in \mathtt{ch}^{\mathrm{sp}}_\mathcal{O}(\mathcal{A}_{C_0})$ for some b. By Lemma 10, $\mathcal{O} \models C_0 \sqsubseteq \bot$ and thus $\mathtt{Dis}^{\mathcal{EL}}_\mathcal{O}(C_0) = \bot$. By definition of '$\vdash$', $\mathtt{ch}^{\mathrm{sp}}_\mathcal{O}(\mathcal{A}_{C_0}) \vdash \bot(a_0)$ and thus we are done. In what follows, we can thus assume that $\bot(b) \notin \mathtt{ch}^{\mathrm{sp}}_\mathcal{O}(\mathcal{A}_{C_0})$ for all b.

To deal with the general case, assume to the contrary of what we have to prove that $\mathtt{ch}^{\mathrm{sp}}_\mathcal{O}(\mathcal{A}_{C_0}) \not\vdash \mathtt{Dis}^{\mathcal{EL}}_\mathcal{O}(C_0)(a_0)$. We are going to construct from $\mathtt{ch}^{\mathrm{sp}}_\mathcal{O}(\mathcal{A}_{C_0})$ a model $\mathcal{I}$ of $\mathcal{O}$ with an element d such that $d \in C_0^\mathcal{I} \setminus (\mathtt{Dis}^{\mathcal{EL}}_\mathcal{O}(C_0))^\mathcal{I}$, in contradiction to the definition of $\mathtt{Dis}^{\mathcal{EL}}_\mathcal{O}(C_0)$.

An original $a \in \mathtt{Ind}(\mathtt{ch}^{\mathrm{sp}}_{\mathcal{O}}(\mathcal{A}_{C_0}))$ is *disjunctive* if $\mathtt{ch}^{\mathrm{sp}}_{\mathcal{O}}(\mathcal{A}_{C_0})$ contains at least one assertion $D(a)$ with $D \in \mathtt{Dis}^{-}(\mathcal{O})$. With each original disjunctive a, we associate a disjunction

$$D_a = \mathtt{Dis}_{\mathcal{O}}(\bigsqcap\{D \in \mathtt{Dis}(\mathcal{O}) \mid \mathtt{ch}^{\mathrm{sp}}_{\mathcal{O}}(\mathcal{A}_{C_0}) \vdash D(a)\}).$$

For the above definition, it is important to note that $\mathtt{sub}^{-}(\mathcal{O}) \subseteq \mathtt{Dis}(\mathcal{O})$ and thus also all $C \in \mathtt{sub}^{-}(\mathcal{O})$ with $\mathtt{ch}^{\mathrm{sp}}_{\mathcal{O}}(\mathcal{A}_{C_0}) \vdash C(a)$ contribute to the definition of D_a. We observe the following:

(P1) If $\mathtt{ch}^{\mathrm{sp}}_{\mathcal{O}}(\mathcal{A}_{C_0}) \vdash D(a)$ with $D \in \mathtt{Dis}(\mathcal{O})$, then $\emptyset \models D_a \sqsubseteq D$.
$\mathtt{ch}^{\mathrm{sp}}_{\mathcal{O}}(\mathcal{A}_{C_0}) \vdash D(a)$ implies

$$\emptyset \models \bigsqcap\{D' \in \mathtt{Dis}(\mathcal{O}) \mid \mathtt{ch}^{\mathrm{sp}}_{\mathcal{O}}(\mathcal{A}_{C_0}) \vdash D'(a_0)\} \sqsubseteq D.$$

By definition of $\mathtt{Dis}_{\mathcal{O}}$,

$$\emptyset \models D_a \sqsubseteq \bigsqcap\{D' \in \mathtt{Dis}(\mathcal{O}) \mid \mathtt{ch}^{\mathrm{sp}}_{\mathcal{O}}(\mathcal{A}_{C_0}) \vdash D'(a_0)\}$$

and thus $\emptyset \models D_a \sqsubseteq D$.

(P2) If $\emptyset \models D_a \sqsubseteq D \in \mathtt{Dis}(\mathcal{O})$, then $D(a) \in \mathtt{ch}^{\mathrm{sp}}_{\mathcal{O}}(\mathcal{A}_{C_0})$.
Since a is disjunctive, there is some $D_1(a) \in \mathtt{ch}^{\mathrm{sp}}_{\mathcal{O}}(\mathcal{A}_{C_0})$ with $D_1 \in \mathtt{Dis}^{-}(\mathcal{O})$. Let $D_1, \dots, D_k$ be all disjunctions from $\mathtt{Dis}(\mathcal{O})$ with $\mathtt{ch}^{\mathrm{sp}}_{\mathcal{O}}(\mathcal{A}_{C_0}) \vdash D_i(a)$. Consider $\mathtt{Dis}_{\mathcal{O}}(D_1 \sqcap D_2)$. In the very special case that this disjunction consists of a single disjunct that contains all concepts from $\mathtt{sub}^{-}(\mathcal{O})$ as conjuncts, $\mathtt{Dis}_{\mathcal{O}}(D_1 \sqcap D_2) = D_a$ and Rule 4 applied to D_1 and D_2 yields $D_a(a) \in \mathtt{ch}^{\mathrm{sp}}_{\mathcal{O}}(\mathcal{A}_{C_0})$ as required. Otherwise, we find some D' with at least two disjuncts such that $\emptyset \models \mathtt{Dis}_{\mathcal{O}}(D_1 \sqcap D_2) \sqsubseteq D'_2$. We can apply Rule 4 again to show $D'_2(a) \in \mathtt{ch}^{\mathrm{sp}}_{\mathcal{O}}(\mathcal{A}_{C_0})$. Since D'_2 has at least two disjuncts, we can proceed in the same way applying Rule 4 to D'_2, D_3, then to D'_3, D_4, and so on. In the last step, we can clearly choose D_a as D'_k. Finally, another application of Rule 4 with $D_1 = D_2 = D_a$ and $D_3 = D$ yields $D(a) \in \mathtt{ch}^{\mathrm{sp}}_{\mathcal{O}}(\mathcal{A}_{C_0})$.

Note that it follows from (P2) and the assumption that $\bot(a) \notin \mathtt{ch}^{\mathrm{sp}}_{\mathcal{O}}(\mathcal{A}_{C_0})$ that $D_a \neq \bot$, that is, D_a has at least one disjunct.

We now consider each original disjunctive $a \in \mathtt{Ind}(\mathtt{ch}^{\mathrm{sp}}_{\mathcal{O}}(\mathcal{A}_{C_0}))$, identify a disjunct E_a of D_a and extend $\mathtt{ch}^{\mathrm{sp}}_{\mathcal{O}}(\mathcal{A}_{C_0})$ with $D(a)$ for each $D \in \mathtt{Dis}(\mathcal{O})$ with $\emptyset \models E_a \sqsubseteq D$. We then show that no new applications of chase rules are possible afterwards, with the possible exception of applications of Rule 3 to original disjunctive individuals a. We also select an E_a for each original non-disjunctive a, in a trivial way: E_a is then the conjunction of all $C \in \mathtt{sub}^{-}(\mathcal{O})$ such that $\mathtt{ch}^{\mathrm{sp}}_{\mathcal{O}}(\mathcal{A}_{C_0}) \vdash C(a)$.

We start at the root a_0 (if it is disjunctive). Recall our assumption that $\mathtt{ch}^{\mathrm{sp}}_{\mathcal{O}}(\mathcal{A}_{C_0}) \not\vdash \mathtt{Dis}^{\mathcal{EL}}_{\mathcal{O}}(C_0)(a_0)$. There must be a disjunct E_{a_0} of D_{a_0} such that $\emptyset \not\models E_{a_0} \sqsubseteq \mathtt{Dis}^{\mathcal{EL}}_{\mathcal{O}}(C_0)$ as otherwise $\emptyset \models D_{a_0} \sqsubseteq \mathtt{Dis}^{\mathcal{EL}}_{\mathcal{O}}(C_0)$ and thus (P2) yields $\mathtt{Dis}^{\mathcal{EL}}_{\mathcal{O}}(C_0)(a_0) \in \mathtt{ch}^{\mathrm{sp}}_{\mathcal{O}}(\mathcal{A}_{C_0})$. Together with Rules 2 and 3, this yields $\mathtt{ch}^{\mathrm{sp}}_{\mathcal{O}}(\mathcal{A}_{C_0}) \vdash \mathtt{Dis}^{\mathcal{EL}}_{\mathcal{O}}(C_0)(a_0)$, a contradiction. Let $\mathcal{A}^{+}$ denote the result of extending $\mathtt{ch}^{\mathrm{sp}}_{\mathcal{O}}(\mathcal{A}_{C_0})$ for E_{a_0} as described above. We observe the following counterpart of (P1) for $\mathcal{A}^{+}$.

(P3) if $\mathcal{A}^{+} \vdash D(a_0)$ with $D \in \mathtt{Dis}(\mathcal{O})$, then $\emptyset \models E_{a_0} \sqsubseteq D$.
If $\mathtt{ch}^{\mathrm{sp}}_{\mathcal{O}}(\mathcal{A}_{C_0}) \vdash D(a_0)$, then this follows from (P1). Otherwise, by definition of $\vdash$ and construction of $\mathcal{A}^{+}$, we must have $\emptyset \models \bigsqcap S \sqsubseteq D$ where S contains

1. all concepts $D' \in \mathtt{Dis}(\mathcal{O})$ such that $\mathtt{ch}^{\mathrm{sp}}_{\mathcal{O}}(\mathcal{A}_{C_0}) \vdash D'(a_0)$ and
2. all concepts D' with $D'(a_0)$ fresh in $\mathcal{A}^{+}$.

(P1) implies $\emptyset \models E_{a_0} \sqsubseteq D'$ for all concepts D' from Point 1 and the construction of $\mathcal{A}^{+}$ yields $\emptyset \models E_{a_0} \sqsubseteq D'$ for all concepts D' from Point 2. Thus $\emptyset \models E_{a_0} \sqsubseteq D$.

We show that no new rule applications are possible, except for applications of Rule 3 to original disjunctive individuals:

- Rule 1. Assume that $\mathcal{A}^{+} \vdash C(a_0)$ and that $C \sqsubseteq D \in \mathcal{O}$. Then (P3) yields $\emptyset \models E_{a_0} \sqsubseteq C$. By definition of D_{a_0}, this implies that C is a conjunct of E_{a_0}. Let D be of the form $C_1 \sqcap \cdots \sqcap C_n \sqcap D_1 \sqcap \cdots \sqcap D_m$ where $C_1, \dots, C_n$ are concept names or existential restrictions and $D_1, \dots, D_m$ are disjunctions. Then $C_1, \dots, C_n$ must also be conjuncts in E_{a_0}. Moreover, for $1 \leq i \leq m$, $\mathtt{DNF}(D_i)$ must contain a disjunct G such that all conjuncts of G are in E_{a_0}. This implies $\emptyset \models E_{a_0} \sqsubseteq \mathtt{DNF}(D)$. Consequently, $\mathtt{DNF}(D)(a_0)$ is in $\mathcal{A}^{+}$.
- Rule 2. If $C_1 \sqcap C_2(a_0)$ is fresh in $\mathcal{A}^{+}$, then $C_1 \sqcap C_2 \in \mathtt{Dis}(\mathcal{O})$ is such that $\emptyset \models E_{a_0} \sqsubseteq C_1 \sqcap C_2$. Thus $\emptyset \models E_{a_0} \sqsubseteq C_i$ for $i \in \{1, 2\}$ and as a consequence, $C_1(a_0), C_2(a_0)$ are also in $\mathcal{A}^{+}$.
- Rule 3. New applications of Rule 3 are possible only to a_0, which is original and disjunctive.
- Rule 4. Assume that $D_1(a_0)$ is in $\mathcal{A}^{+}$, $D_1 \in \mathtt{Dis}^{-}(\mathcal{O})$, that $\mathcal{A}^{+} \vdash D_2(a_0)$, and that $\mathcal{O} \models D_1 \sqcap D_2 \sqsubseteq D_3$. We have $\emptyset \models E_{a_0} \sqsubseteq D_1$: if $D_1(a_0)$ is fresh in $\mathcal{A}^{+}$, then this is clear; otherwise, $D_1(a_0) \in \mathtt{ch}^{\mathrm{sp}}_{\mathcal{O}}(\mathcal{A}_{C_0})$, which implies $\mathtt{ch}^{\mathrm{sp}}_{\mathcal{O}}(\mathcal{A}_{C_0}) \vdash D_1(a_0)$ since $D_1 \in \mathtt{Dis}^{-}(\mathcal{O})$, and thus (P1) yields $\emptyset \models E_{a_0} \sqsubseteq D_1$. Moreover, $\emptyset \models E_{a_0} \sqsubseteq D_2$ by (P3). By definition of D_{a_0}, $\emptyset \models E_{a_0} \sqsubseteq D_3$ and thus $D_3(a_0)$ is in $\mathcal{A}^{+}$.
- Rule 5. Trivially not applicable since a_0 has no predecessors.
- Rule 6 and 7. Only apply to anonymous individuals, but a_0 is original.

This finishes the extension of $\mathtt{ch}^{\mathrm{sp}}_{\mathcal{O}}(\mathcal{A}_{C_0})$ at a_0. From now on, we assume that this extension has been incorporated into $\mathtt{ch}^{\mathrm{sp}}_{\mathcal{O}}(\mathcal{A}_{C_0})$, that is, we write $\mathtt{ch}^{\mathrm{sp}}_{\mathcal{O}}(\mathcal{A}_{C_0})$ in place of $\mathcal{A}^{+}$. Trivially, property (P1) still holds for all $a \neq a_0$ and property (P2) is preserved.

We then apply the following extension as long as possible. Choose some $r(b, a) \in \mathtt{ch}^{\mathrm{sp}}_{\mathcal{O}}(\mathcal{A}_{C_0})$ with a original and disjunctive and assume that E_b was already determined and $\mathtt{ch}^{\mathrm{sp}}_{\mathcal{O}}(\mathcal{A}_{C_0})$ extended accordingly (the latter only if b is disjunctive). We argue that there must be a disjunct E_a of D_a such that the following properties are satisfied:

(a) $\emptyset \models E_a \sqsubseteq C$ and $\exists r.C \in \mathtt{sub}(\mathcal{O})$ implies $\mathtt{ch}^{\mathrm{sp}}_{\mathcal{O}}(\mathcal{A}_{C_0}) \vdash \exists r.C(b)$;

(b) $\emptyset \models E_a \sqsubseteq D_1 \in \mathtt{Dis}^-(\mathcal{O})$ and $\mathcal{O} \models \exists r.D_1 \sqsubseteq D_2$ with $D_2 \in \mathtt{Dis}(\mathcal{O})$ implies that $D_2(b) \in \mathtt{ch}^{\mathrm{sp}}_{\mathcal{O}}(\mathcal{A}_{C_0})$.

Assume that this is not the case. Let $D_a = E_1 \sqcup \cdots \sqcup E_k$. For $1 \leq i \leq k$, we then find one of the following:

(i) $D'_i = \exists r.D_i \in \mathtt{sub}(\mathcal{O})$ with $\emptyset \models E_i \sqsubseteq D_i$ and $\mathtt{ch}^{\mathrm{sp}}_{\mathcal{O}}(\mathcal{A}_{C_0}) \not\vdash D'_i(b)$.

(ii) $D_i \in \mathtt{Dis}^-(\mathcal{O})$ and $D'_i \in \mathtt{Dis}(\mathcal{O})$ such that $\emptyset \models E_i \sqsubseteq D_i$, $\mathcal{O} \models \exists r.D_i \sqsubseteq D'_i$, and $D'_i(b) \notin \mathtt{ch}^{\mathrm{sp}}_{\mathcal{O}}(\mathcal{A}_{C_0})$.

Let D_L denote the result of removing identical disjuncts from $\bigsqcup_{1\leq i\leq k} D_i$ and D_R the result of removing identical disjuncts from $\bigsqcup_{1\leq i\leq k} D'_i$. We have $D_L, D_R \in \mathtt{Dis}(\mathcal{O})$ while this need not be true for the disjunctions that they have been obtained from. Clearly, $\mathcal{O} \models \exists r.D_L \sqsubseteq D_R$. Since each D_i is from $\mathtt{Dis}^-(\mathcal{O})$, $D_L \in \mathtt{Dis}^-(\mathcal{O})$ while this is not guaranteed for D_R even when $k > 1$. Since $\mathcal{O} \models \exists r.D_L \sqsubseteq D_R$ and since D_L is from $\mathtt{Dis}^-(\mathcal{O})$, Rule 5 yields $D_R(b) \in \mathtt{ch}^{\mathrm{sp}}_{\mathcal{O}}(\mathcal{A}_{C_0})$. due to Rules 2 and 3, this implies $\mathtt{ch}^{\mathrm{sp}}_{\mathcal{O}}(\mathcal{A}_{C_0}) \vdash D_R(b)$. We distinguish two cases.

First assume that b is disjunctive. Then D_b is defined and $\mathtt{ch}^{\mathrm{sp}}_{\mathcal{O}}(\mathcal{A}_{C_0}) \vdash D_R(b)$ and (P2) yield $\emptyset \models D_b \sqsubseteq D_R$. It follows that there is a disjunct K of D_R with $\emptyset \models E_b \sqsubseteq K$. Consequently, $\emptyset \models E_b \sqsubseteq D'_i$ for some i. It follows that when $\mathtt{ch}^{\mathrm{sp}}_{\mathcal{O}}(\mathcal{A}_{C_0})$ was extended for b, then $D'_i(b)$ has been added to $\mathtt{ch}^{\mathrm{sp}}_{\mathcal{O}}(\mathcal{A}_{C_0})$. If D_i, D'_i come from Case (ii), then this is an immediate contradiction. Otherwise, non-applicability of Rules 2 and 3 yields $\mathtt{ch}^{\mathrm{sp}}_{\mathcal{O}}(\mathcal{A}_{C_0}) \vdash D'_i(b)$, a contradiction to Case (i).

Now assume that b is not disjunctive. Since $D_R(b) \in \mathtt{ch}^{\mathrm{sp}}_{\mathcal{O}}(\mathcal{A}_{C_0})$, this implies that D_R has only a single disjunct K. This implies that $D'_1 = \cdots = D'_k = K$. From $\mathtt{ch}^{\mathrm{sp}}_{\mathcal{O}}(\mathcal{A}_{C_0}) \vdash D_R(b)$ and $D_R(b) \in \mathtt{ch}^{\mathrm{sp}}_{\mathcal{O}}(\mathcal{A}_{C_0})$, we thus obtain $\mathtt{ch}^{\mathrm{sp}}_{\mathcal{O}}(\mathcal{A}_{C_0}) \vdash D'_1(b)$ and $D'_1(b) \in \mathtt{ch}^{\mathrm{sp}}_{\mathcal{O}}(\mathcal{A}_{C_0})$, a contradiction in Case (i) and (ii), respectively.

We now extend $\mathtt{ch}^{\mathrm{sp}}_{\mathcal{O}}(\mathcal{A}_{C_0})$ for E_a as described above. We observe the same property (P3) as above, the proof is identical:

(P3) if $\mathcal{A}^+ \vdash D(a)$ with $D \in \mathtt{Dis}(\mathcal{O})$, then $\emptyset \models E_a \sqsubseteq D$.

We again show that no new rule applications are possible except applications of Rule 3 to original disjunctive individuals. We only consider those cases explicitly for which the arguments are not the same as above:

- Rule 1. Assume that $\mathcal{A}^+ \vdash C(c)$ and that $C \sqsubseteq D \in \mathcal{O}$. We can use Property (a) to show that the former implies $\mathtt{ch}^{\mathrm{sp}}_{\mathcal{O}}(\mathcal{A}_{C_0}) \vdash C(c)$ whenever $c \neq a$, and thus non-applicability of Rule 1 before the extension ensures $\mathtt{DNF}(D)(c) \in \mathtt{ch}^{\mathrm{sp}}_{\mathcal{O}}(\mathcal{A}_{C_0}) \subseteq \mathcal{A}^+$. Now assume that $c = a$. From $\mathcal{A}^+ \vdash C(a)$ and (P3), we obtain $\emptyset \models E_a \sqsubseteq C$. By definition of D_a, this implies $\emptyset \models E_a \sqsubseteq \mathtt{DNF}(D)$, and consequently $\mathtt{DNF}(a)$ is in $\mathcal{A}^+$.
- Rule 4. Assume that $D_1(c)$ is in $\mathcal{A}^+$, $D_1 \in \mathtt{Dis}^-(\mathcal{O})$, that $\mathcal{A}^+ \vdash D_2(c)$, and that $\mathcal{O} \models D_1 \sqcap D_2 \sqsubseteq D_3$. First assume that $c \neq a$. Then $D_1(c) \in \mathtt{ch}^{\mathrm{sp}}_{\mathcal{O}}(\mathcal{A}_{C_0})$. Moreover, we can use Property (a) to show that $\mathcal{A}^+ \vdash D_2(c)$ implies $\mathtt{ch}^{\mathrm{sp}}_{\mathcal{O}}(\mathcal{A}_{C_0}) \vdash D_2(c)$. Thus, Rule 4 yields $D_2(c) \in \mathtt{ch}^{\mathrm{sp}}_{\mathcal{O}}(\mathcal{A}_{C_0}) \subseteq \mathcal{A}^+$. Now assume that $c = a$. Then $\emptyset \models E_a \sqsubseteq D_1$: if $D_1(a)$ is fresh in $\mathcal{A}^+$, then this is clear; otherwise, otherwise, $D_1(a) \in \mathtt{ch}^{\mathrm{sp}}_{\mathcal{O}}(\mathcal{A}_{C_0})$, which implies $\mathtt{ch}^{\mathrm{sp}}_{\mathcal{O}}(\mathcal{A}_{C_0}) \vdash D_1(a)$ since $D_1 \in \mathtt{Dis}^-(\mathcal{O})$, and thus (P1) yields $\emptyset \models E_a \sqsubseteq D_1$. Moreover, $\emptyset \models E_a \sqsubseteq D_2$ by (P3). By definition of D_a, $\emptyset \models E_a \sqsubseteq D_2$ and thus $D_2(a)$ is in $\mathcal{A}^+$.
- Rule 5. Assume that $r(b,c), D_1(c) \in \mathcal{A}^+$ with $D_1(c)$ fresh and $D_1 \in \mathtt{Dis}^-(\mathcal{O})$. Assume further that $\mathcal{O} \models \exists r.D_1 \sqsubseteq D_2$. Clearly, we must have $c = a$. Since $D_1(a) \in \mathcal{A}^+$ and $D_1 \in \mathtt{Dis}^-(\mathcal{O})$, $\mathcal{A}^+ \vdash D_1(a)$. Thus (P2) yields $\emptyset \models E_a \sqsubseteq D_1$ and from Property (b) we obtain $D_2(b) \in \mathtt{ch}^{\mathrm{sp}}_{\mathcal{O}}(\mathcal{A}_{C_0}) \subseteq \mathcal{A}^+$.

This finishes the extension of $\mathtt{ch}^{\mathrm{sp}}_{\mathcal{O}}(\mathcal{A}_{C_0})$ at a. It is not hard to verify that (P1) holds for all original disjunctive a for which the extension has not yet been carried out, (P2) is preserved, and (P3) holds for all original disjunctive a for which the extension has already been carried out. In particular, the extension of $\mathtt{ch}^{\mathrm{sp}}_{\mathcal{O}}(\mathcal{A}_{C_0})$ at a does not invalidate (P3) for original disjunctive a' that have been treated earlier due to Property (a). We continue until the extension has taken place for all original disjunctive a and use $\mathcal{E}$ to denote the ABox that is obtained in the limit. No new rule applications are possible with the exception of applications of Rule 3 to original disjunctive individuals.

Recall that we aim to construct a model $\mathcal{I}$ of $\mathcal{O}$ with an element d such that $d \in C_0^{\mathcal{I}} \setminus (\mathtt{Dis}^{\mathcal{EL}}_{\mathcal{O}}(C_0))^{\mathcal{I}}$. We are going to start from $\mathcal{E}^-$, that is, $\mathcal{E}$ restricted to role assertions and atomic concept assertions, viewed as an interpretation. The resulting interpretation $\mathcal{I}$, however, need not be a model of $\mathcal{O}$, for two reasons. First, new applications of Rule 3 to original disjunctive individuals a are possible which means that there might be assertions $\exists r.C(a) \in \mathcal{E}$ such that $a \notin (\exists r.C)^{\mathcal{I}}$, and this in turn means that some CIs in $\mathcal{O}$ might not be satisfied. And second, we have chosen disjuncts E_a of the disjunctions D_a at original disjunctive individuals to ensure that all disjunctions are satisfied at original individuals, but we have not ensured the same at anonymous individuals. We thus modify the initial $\mathcal{I}$ in two ways, which both involve grafting additional tree-shaped interpretations that we select in what follows. We first observe that

($*$) If $C(a) \in \mathcal{E}$ with a original and disjunctive, then $\emptyset \models E_a \sqsubseteq C$.

To see this, first assume that $\exists r.C(a) \in \mathtt{ch}^{\mathrm{sp}}_{\mathcal{O}}(\mathcal{A}_{C_0})$ already before the extension to $\mathcal{E}$. Then non-applicability of Rules 2 and 3 implies $\mathtt{ch}^{\mathrm{sp}}_{\mathcal{O}}(\mathcal{A}_{C_0}) \vdash C(a)$ and (P3) yields $\emptyset \models E_a \sqsubseteq C$. Otherwise, $\emptyset \models E_a \sqsubseteq C$ by construction of $\mathcal{E}$.

By ($*$) and the (semantic!) definition of D_a of which E_a is a disjunct, we find for each $\exists r.C(a) \in \mathcal{E}$ with a original and disjunctive, a tree model $\mathcal{I}_{\exists r.C(a)}$ of $\mathcal{O}$ with root d such that $d \in C^{\mathcal{I}_{\exists r.C(a)}}$ and $d \in F^{\mathcal{I}_{\exists r.C(a)}}$ implies $\emptyset \models E_a \sqsubseteq \exists r.F$ for all $\exists r.F \in \mathtt{sub}(\mathcal{O})$.

Let Γ denote the set of individuals a in $\mathtt{ch}^{\mathrm{sp}}_{\mathcal{O}}(\mathcal{A}_{C_0})$ that are anonymous and marked and whose predecessor is anonymous and unmarked.[3] Let $a \in \Gamma$ have been introduced for C_a and let $r(b,a)$ be the unique assertion of this form in $\mathtt{ch}^{\mathrm{sp}}_{\mathcal{O}}(\mathcal{A}_{C_0})$.

[3] We work here with the anonymous part, which is identical in $\mathtt{ch}^{\mathrm{sp}}_{\mathcal{O}}(\mathcal{A}_{C_0})$ and in $\mathcal{E}$.

Since Rule 5 is not applicable, $\mathtt{Dis}_{\mathcal{O}}(\exists r.C_a)(b) \in \mathtt{ch}^{\mathrm{sp}}_{\mathcal{O}}(\mathcal{A}_{C_0})$ and since b is anonymous and not marked, $\mathtt{Dis}_{\mathcal{O}}(\exists r.C_a) \notin \mathtt{Dis}^-(\mathcal{O})$. Furthermore, $\mathtt{Dis}_{\mathcal{O}}(\exists r.C_a)$ is not empty since we assume that $\mathtt{ch}^{\mathrm{sp}}_{\mathcal{O}}(\mathcal{A}_{C_0})$ contains no assertion of the form $\bot(b)$. Consequently, we find a tree model $\mathcal{I}_a$ of $\mathcal{O}$ with root $a \in C_a^{\mathcal{I}_a}$ such that, in the extension $\mathcal{J}$ of $\mathcal{I}_a$ obtained by adding an r-predecessor b to the root a of $\mathcal{I}_a$, we have $b \in (\exists r.C)^{\mathcal{J}}$ iff $\mathcal{O} \models C_a \sqsubseteq C$ for all $\exists r.C \in \mathtt{sub}(\mathcal{O})$.

Construct an interpretation $\mathcal{I}$ as follows:

- start with $\mathcal{E}^-$ viewed as an interpretation;
- for each $\exists r.C(a) \in \mathcal{E}$ with a original and disjunctive, disjointly add the interpretation $\mathcal{I}_{\exists r.C(a)}$ with root d and extend $r^{\mathcal{I}}$ with (a, d);
- for each $a \in \Gamma$, replace the subtree rooted at a with $\mathcal{I}_a$.

We next observe the following.

Claim 1. Let $a \in \Delta^{\mathcal{I}}$ be an individual of $\mathtt{ch}^{\mathrm{sp}}_{\mathcal{O}}(\mathcal{A}_{C_0})$ and let $C \in \mathtt{sub}(\mathcal{O})$ be an $\mathcal{EL}$ concept. Then $a \in C^{\mathcal{I}}$ implies

1. $\emptyset \models E_a \sqsubseteq C$ if a is original and disjunctive, and
2. $\mathtt{ch}^{\mathrm{sp}}_{\mathcal{O}}(\mathcal{A}_{C_0}) \vdash C(a)$ otherwise.

The proof is by induction on the structure of C. In the induction start, $C = A$ is a concept name. Since $a \in A^{\mathcal{I}}$, we must have $A(a) \in \mathcal{E}$ and thus $\mathcal{E} \vdash A(a)$. If a is original and disjunctive, then (P3) yields $\emptyset \models E_a \sqsubseteq C$ as required. If this is not the case, then $A(a) \in \mathcal{E}$ implies $A(a) \in \mathtt{ch}^{\mathrm{sp}}_{\mathcal{O}}(\mathcal{A}_{C_0})$, thus $\mathtt{ch}^{\mathrm{sp}}_{\mathcal{O}}(\mathcal{A}_{C_0}) \vdash A(a)$ as required.

The case $C = C_1 \sqcap C_2$ is straightforward using the semantics and induction hypothesis. Details are left to the reader.

It thus remains to deal with the case $C = \exists r.C_1$. Then $a \in C^{\mathcal{I}}$ implies that there is a $d \in C_1^{\mathcal{I}}$ with $(a, d) \in r^{\mathcal{I}}$. We distinguish several cases. First assume that d is an individual from $\mathtt{ch}^{\mathrm{sp}}_{\mathcal{O}}(\mathcal{A}_{C_0})$ that is not in Γ. We have the following subcase:

1. a and d are original and disjunctive.
   The induction hypothesis yields $\emptyset \models E_d \sqsubseteq C_1$. Thus Condition (a) from the extension step ensures that $\mathtt{ch}^{\mathrm{sp}}_{\mathcal{O}}(\mathcal{A}_{C_0}) \vdash \exists r.C_1(a)$. Property (P1) yields $\emptyset \models E_a \sqsubseteq \exists r.C_1$, as required.
2. a is original and disjunctive and d is not.
   The induction hypothesis yields $\mathtt{ch}^{\mathrm{sp}}_{\mathcal{O}}(\mathcal{A}_{C_0}) \vdash C_1(d)$ and the construction of $\mathcal{I}$ yields $r(a, d) \in \mathtt{ch}^{\mathrm{sp}}_{\mathcal{O}}(\mathcal{A}_{C_0})$, thus $\mathtt{ch}^{\mathrm{sp}}_{\mathcal{O}}(\mathcal{A}_{C_0}) \vdash \exists r.C_1(a)$. Property (P1) yields $\emptyset \models E_a \sqsubseteq \exists r.C_1$, as required.
3. d is original and disjunctive and a is not.
   The induction hypothesis yields $\emptyset \models E_d \sqsubseteq C_1$. Thus Condition (a) from the extension step ensures that $\mathtt{ch}^{\mathrm{sp}}_{\mathcal{O}}(\mathcal{A}_{C_0}) \vdash \exists r.C_1(a)$, as required.
4. neither a nor d are original and disjunctive.
   The induction hypothesis yields $\mathtt{ch}^{\mathrm{sp}}_{\mathcal{O}}(\mathcal{A}_{C_0}) \vdash C_1(d)$ and the construction of $\mathcal{I}$ yields $r(a, d) \in \mathtt{ch}^{\mathrm{sp}}_{\mathcal{O}}(\mathcal{A}_{C_0})$, thus $\mathtt{ch}^{\mathrm{sp}}_{\mathcal{O}}(\mathcal{A}_{C_0}) \vdash \exists r.C_1(a)$, as required.

The following cases remain:

5. a is original and disjunctive and there is an $\exists r.E(a) \in \mathtt{ch}^{\mathrm{sp}}_{\mathcal{O}}(\mathcal{A}_{C_0})$ such that d is the root of $\mathcal{I}_{\exists r.E(a)}$.
   By choice of $\mathcal{I}_{\exists r.E(a)}$, this implies $\emptyset \models E_a \sqsubseteq \exists r.C_1$, as required.
6. $d \in \Gamma$ and thus the root of $\mathcal{I}_d$.
   Then $d \in C_1^{\mathcal{I}_d}$. By choice of $\mathcal{I}_d$ and since $\exists r.C_1 \in \mathtt{sub}(\mathcal{O})$, we have $\mathcal{O} \models C_d \sqsubseteq C_1$ where C_d is the concept that d was introduced for. We moreover have $C_d(d) \in \mathtt{ch}^{\mathrm{sp}}_{\mathcal{O}}(\mathcal{A}_{C_0})$ and by non-applicability of Rules 2 and 3 also $\mathtt{ch}^{\mathrm{sp}}_{\mathcal{O}}(\mathcal{A}_{C_0}) \vdash C_d(d)$. Since d is in Γ, it is marked. Thus $D(d) \in \mathtt{ch}^{\mathrm{sp}}_{\mathcal{O}}(\mathcal{A}_{C_0})$ for some $D \in \mathtt{Dis}^-(\mathcal{O})$. Clearly, $\mathcal{O} \models D \sqcap C_d \sqsubseteq C_1$. We can thus invoke Rule 4 with $D_1 = D$, $D_2 = C_d$, and $D_3 = C_1$ to yield $C_1(d) \in \mathtt{ch}^{\mathrm{sp}}_{\mathcal{O}}(\mathcal{A}_{C_0})$. Rules 2 and 3 thus ensure that $\mathtt{ch}^{\mathrm{sp}}_{\mathcal{O}}(\mathcal{A}_{C_0}) \vdash C_1(d)$. From $(a, d) \in r^{\mathcal{I}}$, we obtain $r(a, d) \in \mathtt{ch}^{\mathrm{sp}}_{\mathcal{O}}(\mathcal{A}_{C_0})$, and thus $\mathtt{ch}^{\mathrm{sp}}_{\mathcal{O}}(\mathcal{A}_{C_0}) \vdash \exists r.C_1(a)$ as required (since a cannot be original).

This finishes the proof of the claim.

Claim 2. Let $a \in \Delta^{\mathcal{I}}$ be an individual of $\mathtt{ch}^{\mathrm{sp}}_{\mathcal{O}}(\mathcal{A}_{C_0})$ and let $C \in \mathtt{sub}(\mathcal{O})$. Then $\mathtt{ch}^{\mathrm{sp}}_{\mathcal{O}}(\mathcal{A}_{C_0}) \vdash C(a)$ implies $a \in C^{\mathcal{I}}$.

The proof is by induction on the structure of C. The case that C is a concept name is clear by construction of $\mathcal{I}$. The case that $C = C_1 \sqcap C_2$ and $C = \exists r.C_1$ are straightforward using the fact that Rules 2 and 3 are not applicable and the induction hypothesis. It remains to deal with the case $C = C_1 \sqcup C_2$. Then $C_1 \sqcup C_2(a) \in \mathtt{ch}^{\mathrm{sp}}_{\mathcal{O}}(\mathcal{A}_{C_0})$ and thus a is disjunctive and $C_1 \sqcup C_2$ is a conjunct of every disjunct of D_a, including the disjunct E_a chosen for a during the extension of $\mathtt{ch}^{\mathrm{sp}}_{\mathcal{O}}(\mathcal{A}_{C_0})$. By definition of D_a, it follows that some C_i, $i \in \{1, 2\}$, is also a conjunct of E_a. Thus $C_i(a)$ has been added in the extension of $\mathtt{ch}^{\mathrm{sp}}_{\mathcal{O}}(\mathcal{A}_{C_0})$ and it remains to apply the induction hypothesis. This finishes the proof of Claim 4

Note that $\mathcal{I}$ is a tree interpretation with root a_0. By construction of $\mathcal{I}$, it is clear that $a_0 \in C_0^{\mathcal{I}}$. We next show that $\mathcal{I}$ is a model of $\mathcal{O}$.

Let $C \sqsubseteq D \in \mathcal{O}$ and let $d \in C^{\mathcal{I}}$. If d is in the domain of some interpretation $\mathcal{I}_{\exists r.C(a)}$ or $\mathcal{I}_a$, then it follows from the construction of $\mathcal{I}$ and the fact that all interpretations $\mathcal{I}_{\exists r.C(a)}$ and $\mathcal{I}_a$ are models of $\mathcal{O}$ that $d \in D^{\mathcal{I}}$. Thus let a be an individual of $\mathtt{ch}^{\mathrm{sp}}_{\mathcal{O}}(\mathcal{A}_{C_0})$. We distinguish two cases.

First assume that a is original and disjunctive. Then Point 1 of Claim 1 yields $\emptyset \models E_a \sqsubseteq C$ and as a consequence we have $\emptyset \models E_a \sqsubseteq D$ which implies $F(a) \in \mathcal{E}$ for all top-level conjuncts F of D. If F is a concept name, then this yields $a \in F^{\mathcal{I}}$ by construction of $\mathcal{I}$. If F takes the form $\exists r.G$, then the addition of $\mathcal{I}_{\exists r.G}$ ensures that $a \in (\exists r.G)^{\mathcal{I}}$. As a consequence, $a \in D^{\mathcal{I}}$.

Now assume that a is not original or not disjunctive. Then Point 2 of Claim 1 yields $\mathtt{ch}^{\mathrm{sp}}_{\mathcal{O}}(\mathcal{A}_{C_0}) \vdash C(a)$. Non-applicability of Rule 1 of the chase yields $\mathtt{DNF}(D)(a) \in \mathtt{ch}^{\mathrm{sp}}_{\mathcal{O}}(\mathcal{A}_{C_0})$ and non-applicability of Rules 2 and 3 yields $\mathtt{ch}^{\mathrm{sp}}_{\mathcal{O}}(\mathcal{A}_{C_0}) \vdash \mathtt{DNF}(D)(a)$. By Claim 2, $a \in D^{\mathcal{I}}$.

It remains to show that $a_0 \notin \mathtt{Dis}^{\mathcal{EL}}_{\mathcal{O}}(C_0)^{\mathcal{I}}$. Assume to the contrary that $a_0 \in \mathtt{Dis}^{\mathcal{EL}}_{\mathcal{O}}(C_0)^{\mathcal{I}}$. Then there is a disjunct K

of $\mathtt{Dis}^{\mathcal{EL}}_{\mathcal{O}}(C_0)^{\mathcal{I}}$ such that $a_0 \in C^{\mathcal{I}}$ for every conjunct C of K. We distinguish two cases.

First assume that a_0 is disjunctive. By Point 1 of Claim 1, $\emptyset \models E_{a_0} \sqsubseteq C$ for all conjuncts C of K. Thus $\emptyset \models E_{a_0} \sqsubseteq \mathtt{Dis}^{\mathcal{EL}}_{\mathcal{O}}(C_0)$, in contradiction to our choice of E_{a_0}.

Now assume that a_0 is not disjunctive. By Point 2 of Claim 1, $\mathtt{ch}^{\mathrm{sp}}_{\mathcal{O}}(\mathcal{A}_{C_0}) \vdash C(a_0)$ for all disjuncts C of K. By Lemma 10, $\mathcal{O} \models C_0 \sqsubseteq C$ for all such C. Since $\mathcal{I}$ is a model of $\mathcal{O}$, this implies that $\mathtt{Dis}^{\mathcal{EL}}_{\mathcal{O}}(C_0)$ has only the disjunct K. We have thus shown that $\mathtt{ch}^{\mathrm{sp}}_{\mathcal{O}}(\mathcal{A}_{C_0}) \vdash \mathtt{Dis}^{\mathcal{EL}}_{\mathcal{O}}(C_0)$, a contraction. ❑

D Proof of Theorem 3

Let $\mathtt{1Dis}(\mathcal{O}_S)$ denote the set of disjunctions of concepts from $\mathtt{sub}^-(\mathcal{O}_S)$, without repetition, of which there are clearly only single exponentially many. For each $D \in \mathtt{Dis}(\mathcal{O}_S)$, let K_D be the set of disjunctions $D' \in \mathtt{1Dis}(\mathcal{O}_S)$ such that D' is a conjunct of the result of converting D viewed as a DNF formula (in which all concept names and concepts $\exists r.E$ serve as propositional variables) into KNF. We then have $\emptyset \models D \equiv \bigsqcap K_D$. For $\ell \in \mathbb{N} \cup \{\omega\}$, let the $\mathcal{EL}$ ontology $\widehat{\mathcal{O}}^{\ell}_T$ be defined as $\mathcal{O}^{\ell}_T$ in Figure 2, but with every occurrence of a concept name X_D replaced by $\bigsqcap_{C \in K_D} Y_C$.

Theorem 3 is a consequence of the following.

Lemma 11. *Let $\ell \in \mathbb{N} \cup \{\omega\}$. For all $\mathcal{EL}$ concepts C_0, D_0 over $\mathtt{sig}(\mathcal{O}_S)$, $\mathcal{O}^{\ell}_T \models C \sqsubseteq D$ iff $\widehat{\mathcal{O}}^{\ell}_T \models C \sqsubseteq D$.*

Proof. It suffices to show that for every model $\mathcal{I}$ of $\mathcal{O}^{\ell}_T$ there is a model $\widehat{\mathcal{I}}$ of $\widehat{\mathcal{O}}^{\ell}_T$ such that the restrictions of $\mathcal{I}$ and $\widehat{\mathcal{I}}$ to the symbols of $\mathtt{sig}(\mathcal{O}_S)$ are identical and vice versa.

We start with the easier direction. Thus let $\widehat{\mathcal{I}}$ be a model of $\widehat{\mathcal{O}}^{\ell}_T$. Let $\mathcal{I}$ be defined like $\widehat{\mathcal{I}}$ except that $X^{\mathcal{I}}_D = \bigsqcap_{D' \in K_D} Y^{\widehat{\mathcal{I}}}_{D'}$. Observe that the concept names X_D are not in $\mathtt{sig}(\mathcal{O}_S)$ and thus the restrictions of $\mathcal{I}$ and $\widehat{\mathcal{I}}$ to the symbols of $\mathtt{sig}(\mathcal{O}_S)$ are identical, as required. It is not straightforward to verify that $\mathcal{I}$ satisfies every CI $C_1 \sqsubseteq C_2 \in \mathcal{O}^{\ell}_T$ given that $\widehat{\mathcal{O}}^{\ell}_T$ contains a CI $C'_1 \sqsubseteq C'_2$ such that C'_i can be obtained from C' by replacing each X_D with $\bigsqcap_{D' \in K_D} Y_{D'}$ and that $\widehat{\mathcal{I}}$ satisfies $C'_1 \sqsubseteq C'_2$.

For the converse direction, let $\mathcal{I}$ be a model of $\mathcal{O}^{\ell}_T$. We cannot define a corresponding $\widehat{\mathcal{I}}$ of $\widehat{\mathcal{O}}^{\ell}_T$ by setting $\bigsqcap_{D' \in K_D} Y^{\widehat{\mathcal{I}}}_{D'} = X^{\mathcal{I}}_D$ because we need to interpret individual concept names $Y_{D'}$ rather than conjunctions thereof. To achieve this, we resort to semantic disjunctions $\mathtt{Dis}_{\mathcal{O}_S}(D)$. In fact, we define $\widehat{\mathcal{I}}$ to be like $\mathcal{I}$ except that

$$Y^{\widehat{\mathcal{I}}}_D = \big(\mathtt{Dis}_{\mathcal{O}_S}(D)^{\uparrow} \sqcap \big(\bigsqcup_{D' \in \mathtt{Dis}^-(\mathcal{O}_S) \mid D \in K_{D'}} X_{D'}\big)\big)^{\mathcal{I}}$$

for every $D \in \mathtt{1Dis}(\mathcal{O}_S)$. It remains to show that $\widehat{\mathcal{I}}$ is a model of of $\widehat{\mathcal{O}}^{\ell}_T$. We can argue exactly as in the converse direction if we know that $\bigsqcap_{D' \in K_D} Y^{\widehat{\mathcal{I}}}_{D'} = X^{\mathcal{I}}_D$ for all $D \in \mathtt{Dis}^-(\mathcal{O}_S)$. By definition of $\widehat{\mathcal{I}}$, this amounts to showing that

$$\big(\bigsqcap_{D' \in K_D} \mathtt{Dis}_{\mathcal{O}_S}(D')^{\uparrow} \sqcap \big(\bigsqcup_{D'' \in \mathtt{Dis}^-(\mathcal{O}_S) \mid D' \in K_{D''}} X_{D''}\big)\big)^{\mathcal{I}} = X^{\mathcal{I}}_D.$$

For the '$\supseteq$' direction, we note that $\mathcal{O}_S \models D \sqsubseteq \mathtt{Dis}_{\mathcal{O}_S}(D')^{\uparrow}$ for every $D' \in K_D$ since $\emptyset \models D \sqsubseteq D'$ and due to the definition of $\mathtt{Dis}_{\mathcal{O}_S}(D')$. Thus $\mathcal{O}^{\ell}_T$ contains the CI $X_D \sqcap X_D \sqsubseteq \mathtt{Dis}_{\mathcal{O}_S}(D')^{\uparrow}$ and consequently $\mathcal{O}^{\ell}_T \models X_D \sqsubseteq \mathtt{Dis}_{\mathcal{O}_S}(D') \sqcap X_D$ for every $D' \in K_D$. It suffices to recall that $\mathcal{I}$ is a model of $\mathcal{O}^{\ell}_T$.

For the '$\subseteq$' direction, assume that

$$d \in \big(\bigsqcap_{D' \in K_D} \mathtt{Dis}_{\mathcal{O}_S}(D')^{\uparrow} \sqcap X_{D''_0}\big)^{\mathcal{I}}$$

for some $D''_0 \in \mathtt{Dis}^-(\mathcal{O}_S)$. Let $K_D = \{D_1, \dots, D_k\}$ and let $D'_1 = \mathtt{Dis}_{\mathcal{O}_S}(D''_0 \sqcap D_1)$.

We first argue that $d \in (D'^{\uparrow}_1)^{\mathcal{I}}$. This follows from $\mathcal{O}_S \models D''_0 \sqcap D_1 \sqsubseteq D'_1$ and the fact that, thus, $\mathcal{O}^{\ell}_T$ contains the CI $X_{D''_0} \sqcap D^{\uparrow}_1 \sqsubseteq D'^{\uparrow}_1$ and since $\mathcal{I}$ is a model of $\mathcal{O}^{\ell}_T$.

In the very special case that D'_1 consists of a single disjunct that contains all concepts from $\mathtt{sub}^-(\mathcal{O}_S)$ as conjuncts, we actually have $\mathcal{O}_S \models D''_0 \sqcap D_1 \sqsubseteq D$, and thus we can argue as above that $d \in (D^{\uparrow})^{\mathcal{I}}$ and are done since $D^{\uparrow} = X_D$ given that $D \in \mathtt{Dis}^-(\mathcal{O}_S)$.

Otherwise, we can find a $D''_1 \in \mathtt{Dis}(\mathcal{O}_S)$ with at least two disjuncts such that $\emptyset \models D'_1 \equiv D''_1$. Let $D'_2 = \mathtt{Dis}_{\mathcal{O}_S}(D''_1 \sqcap D_2)$. As in the case of D''_1, we can show that $d \in (D'^{\uparrow}_2)^{\mathcal{I}}$. We can repeat this until we have shown that $d \in (D'^{\uparrow}_k)^{\mathcal{I}}$, $D'_k = \mathtt{Dis}_{\mathcal{O}_S}(D''_{k-1} \sqcap D_k)$. Again, we are done if D'_k consists of a single disjunct that contains all concepts from $\mathtt{sub}^-(\mathcal{O}_S)$ as conjuncts. Otherwise, we can find a a $D''_k \in \mathtt{Dis}(\mathcal{O}_S)$ with at least two disjuncts such that $\emptyset \models D'_k \equiv D''_1$.

By construction of D'_k and choice of D''_k, we must have $\mathcal{O}_S \models D''_k \sqsubseteq D$. Thus $\mathcal{O}^{\ell}_T$ contains the CI $X_{D''_k} \sqcap D''^{\uparrow}_k \sqsubseteq D^{\uparrow}$. It follows that $d \in X^{\mathcal{I}}_D$. ❑

E Proof of Theorem 5

Theorem 2 follows from the following lemma.

Lemma 12. *Let $\mathcal{O}$ be an acyclic $\mathcal{ELU}$ ontology and let*

$$C = A_1 \sqcap \cdots \sqcap A_n \sqcap \exists r_1.E_1 \sqcap \cdots \sqcap \exists r_m.E_m, \quad D = \exists r.E$$

be $\mathcal{EL}$ concepts such that $\mathcal{O} \models C \sqsubseteq D$ and there does not exist any $i \le m$ with $r_i = r$ and $\mathcal{O} \models E_i \sqsubseteq E$. Then there exists $i \le n$ with $\mathcal{O} \models A_i \sqsubseteq D$.

Proof. Assume the lemma does not hold. Take tree shaped interpretations $\mathcal{I}_i$, $1 \le i \le n$, with root a_i and $\mathcal{J}_i$, $1 \le i \le m$, with root b_i such that all $\mathcal{I}_i, \mathcal{J}_i$ are models of $\mathcal{O}$ and

- $a_i \in A^{\mathcal{I}_i}_i$ and $a_i \notin D^{\mathcal{I}_i}$, for all $1 \le i \le n$;
- $b_i \in E^{\mathcal{J}_i}_i$ and $b_i \notin E^{\mathcal{I}_i}$, for all $1 \le i \le m$.

Construct a model $\mathcal{I}$ by taking the disjoint union of all $\mathcal{I}_i, \mathcal{J}_i$ and then identifying all a_i, $1 \le i \le m$, to a single node a and adding (a, b_i) to the interpretation of r_i. Next define $\mathcal{I}'$ by adding in $\mathcal{I}$, recursively, a to the interpretation of a concept name A if there exists C such that $A \equiv C \in \mathcal{O}$ and $a \in C^{\mathcal{I}}$. We claim that $\mathcal{I}'$ is a model of $\mathcal{O}$ and $a \notin D^{\mathcal{I}}$. The latter holds by definition. For the former, consider some $A' \equiv C' \in \mathcal{O}'$ for which a has not been added to the interpretation of A'

in the step above (the remaining CIs are trivially true in $\mathcal{I}'$). Then it only remains to check that $a \in A'^{\mathcal{I}}$ implies $a \in C'^{\mathcal{I}}$, but this follows by construction again. ❑

F Proof of Theorems 6 and 7

The $\mathcal{EL}$ chase introduced in Appendix C.1 can be extended to $\mathcal{EL}_\bot$ in a straightforward way. Recall that we assume $\bot$ to occur only in CIs of the form $C \sqsubseteq \bot$. The $\mathcal{EL}_\bot$ chase is defined exactly as the $\mathcal{EL}$ chase. In particular, it also treats CIs of the form $C \sqsubseteq \bot$, adding $\bot(a)$ to an ABox $\mathcal{A}$ when $\mathcal{A} \models C(a)$, and thus producing $\mathcal{EL}_\bot$ extended ABoxes. We write $\mathtt{ch}_\mathcal{O}(\mathcal{A}_C) \models \bot$ if there is some a with $\bot(a) \in \mathtt{ch}_\mathcal{O}(\mathcal{A}_C)$. The correctness of the chase now reads as follows.

Lemma 13. *Let $\mathcal{O}$ be an $\mathcal{EL}_\bot$ ontology and let C, D be $\mathcal{EL}_\bot$ concepts. Then $\mathcal{O} \models C \sqsubseteq D$ iff $\mathtt{ch}_\mathcal{O}(\mathcal{A}_C) \models D(a_0)$ or $\mathtt{ch}_\mathcal{O}(\mathcal{A}_C) \models \bot$.*

Based on Lemma 13, we can prove the soundness of the approximation. The proof is essentially identical to that of Lemma 5, that is, to the correctness of the approximation in the $\mathcal{ELU}$-to-$\mathcal{EL}$ case. We omit details.

Lemma 14. *$\mathcal{O}_T^\omega \models C_0 \sqsubseteq D_0$ implies $\mathcal{O}_S \models C_0 \sqsubseteq D_0$ for all $\mathcal{EL}$ concepts C_0, D_0 over $\mathtt{sig}(\mathcal{O}_S)$.*

Now for completeness. Recall that we have established the central Lemma 6 already for the case where $\mathcal{O}_S$ is an $\mathcal{ELU}_\bot$ ontology. The same is true for Lemma 9.

Lemma 15. *Let $\ell \in \mathbb{N} \cup \{\omega\}$. Then $\mathcal{O}_S \models C_0 \sqsubseteq D_0$ implies $\mathcal{O}_T^\ell \models C_0 \sqsubseteq D_0$ for all $\mathcal{EL}$ concepts C_0, D_0 over $\mathtt{sig}(\mathcal{O}_S)$ such that the role depth of D_0 is bounded by ℓ.*

Proof. Assume that $\mathcal{O}_S \models C_0 \sqsubseteq D_0$ with C_0, D_0 $\mathcal{EL}$ concepts over $\mathtt{sig}(\mathcal{O}_S)$ such that the role depth of D_0 is bounded by ℓ. If C_0 contains $\bot$, then clearly $\mathcal{O}_T^\ell \models C_0 \sqsubseteq D_0$. If D_0 contains $\bot$, then it is equivalent to $\bot$. We can thus assume that C_0 is an $\mathcal{EL}$ concept and it suffices to consider the cases where D_0 is $\bot$, a concept name, or of the form $\exists r.E_0$.

We start with the case $D_0 = \bot$. Then $\mathcal{O}_S \models C_0 \sqsubseteq D_0$ implies that $\mathtt{Dis}^{\mathcal{EL}}_{\mathcal{O}_S}(C_0)$ is $\bot$ (that is, it is the empty disjunction), and consequently Lemma 6 yields $\mathcal{O}_T^\ell \models C_0 \sqsubseteq \bot$ as required.

Now let $D_0 = A$. Clearly, $\mathcal{O}_S \models C_0 \sqsubseteq A$ implies $\mathcal{O}_S \models \mathtt{Dis}^{\mathcal{EL}}_{\mathcal{O}_S}(C_0) \sqsubseteq A$. It thus follows from Lemma 6 that $O_T^\ell \models C_0 \sqsubseteq A$ (see proof of Lemma 15 for details).

The case where $D_0 = \exists r.E_0$ is a consequence of the following claim. For each $a \in \mathtt{Ind}(\mathcal{A}_{C_0})$, we write C_0^a as an abbreviation for $C^a_{\mathcal{A}_{C_0}}$.

Claim. For all $a \in \mathtt{Ind}(\mathcal{A}_{C_0})$ and $\mathcal{EL}$ concepts $\exists r.E$ of depth $\ell - \mathtt{depth}(a)$, $\mathcal{O}_S \models C_0^a \sqsubseteq \exists r.E$ implies $\mathcal{O}_T^\ell \models C_0^a \sqsubseteq \exists r.E$.

Proof of claim. The proof is by induction on the codepth of a.

Induction start. Then a is a leaf in $\mathcal{A}_{C_0}$ and thus C_0^a does not have any top-level conjuncts of the form $\exists r.E'$. Thus Condition 1 from Figure 3 is satisfied for $F = C_0^a$. Consequently, $\mathcal{O}_T^\ell$ contains the CI $C_0^a \sqsubseteq \exists r.E$ and we are done.

Induction step. Then a is a non-leaf in $\mathcal{A}_{C_0}$. We distinguish two cases.

$$
\begin{array}{rcll}
C & \sqsubseteq & E^\uparrow & \text{if } C \sqsubseteq E \in \mathcal{O}_S \\
X_D \sqcap D_1^\uparrow & \sqsubseteq & D_2^\uparrow & \text{if } \mathcal{O}_S \models D \sqcap D_1 \sqsubseteq D_2 \\
\exists r.X_D & \sqsubseteq & D_1^\uparrow & \text{if } \mathcal{O}_S \models \exists r.D \sqsubseteq D_1 \\
X_D & \sqsubseteq & \exists r.D_1^\uparrow & \text{if } \mathcal{O}_S \models D \sqsubseteq \exists r.D_1 \\
F^\uparrow & \sqsubseteq & \exists r.G & \text{if } \mathcal{O}_S \models F \sqsubseteq \exists r.G
\end{array}
$$

where in the last line F is an $\mathcal{EL}$ concept over $\mathtt{sig}(\mathcal{O}_S)$ decorated with disjunctions from $\mathtt{Dis}(\mathcal{O}_S)$ at leaves and G is an $\mathcal{EL}$ concept over $\mathtt{sig}(\mathcal{O}_S)$ such that

1. F has no top-level conjunct $\exists r.F'$ s.t. $\mathcal{O}_S \models F' \sqsubseteq G$;
2. it is not the case that $\mathtt{Dis}_{\mathcal{O}_S}(F)$ has at least two disjuncts and $\mathcal{O}_S \models \mathtt{Dis}_{\mathcal{O}_S}(F) \sqsubseteq \exists r.G$;
3. $\mathtt{depth}(F) \le \mathtt{depth}(G) < \ell$.

Figure 4: Optimized ℓ-bounded $\mathcal{EL}_\bot$ approximation $\mathcal{O}_T^\ell$.

Case 1. There is a top-level conjunct $\exists r.E'$ in C_0^a such that $\mathcal{O}_\mathcal{S} \models E' \sqsubseteq E$. Then a has an r-successor b in $\mathcal{A}_{C_0}$ such that $C_0^b = E'$. Let

$$E = A_1 \sqcap \cdots \sqcap A_n \sqcap \exists r_1.E_1 \sqcap \cdots \sqcap \exists r_m.E_m.$$

Since we have already shown Lemma 15 for the case where D_0 is a concept name, we obtain $\mathcal{O}_T^\ell \models C_0^b \sqsubseteq A_i$ for $1 \le i \le n$. From the induction hypothesis, we further obtain $\mathcal{O}_T^\ell \models C_0^b \sqsubseteq \exists r_i.E_i$ for $1 \le i \le m$. Thus $\mathcal{O}_T^\ell \models C_0^b \sqsubseteq E$ and consequently $\mathcal{O}_T^\ell \models C_0^a \sqsubseteq \exists r.E$ as required.

Case 2. There is no top-level conjunct $\exists r.E'$ in C_0^a such that $\mathcal{O}_\mathcal{S} \models E' \sqsubseteq E$. Then Condition 1 from Figure 3 is satisfied for $F = C_0^a$. Let $\mathcal{A}$ be the ditree-shaped subABox of $\mathcal{A}_{C_0}$ rooted at a and let $\mathcal{A}^\pm$ be the extended ABox obtained from $\mathcal{A}|_k$, with k the depth of $\exists r.E$, by adding $\mathtt{Dis}^{\mathcal{EL}}_{\mathcal{O}_S}(C^c_\mathcal{A})^\uparrow(c)$ whenever c is a leaf in $\mathcal{A}|_k$. Applying Lemma 9 to $\mathcal{A}$ and $\mathcal{A}^\pm$ and with $\exists r.E$ in place of $\exists r.C$, we obtain $\mathcal{O}_S, \mathcal{A}^\pm \models \exists r.E(a)$. Let $C^\pm$ be $\mathcal{A}^\pm$ viewed as an $\mathcal{EL}$ concept decorated with disjunctions from $\mathtt{Dis}(\mathcal{O}_S)$ at leaves. Since there is no top-level conjunct $\exists r.E'$ in C_0^a such that $\mathcal{O}_\mathcal{S} \models E' \sqsubseteq E$ there is no top-level conjunct $\exists r.E'$ in $C^\pm$ such that $\mathcal{O}_\mathcal{S} \models E' \sqsubseteq E$ either: if this was the case with $\exists r.E'$ corresponding to the successor $r(a,b)$ of a in $\mathcal{A}^\pm$, then we can apply Lemma 9 to the subABox of $\mathcal{A}$ rooted at b and the subABox of $\mathcal{A}^\pm$ rooted at b to obtain $\mathcal{O}_\mathcal{S}, \mathcal{A} \models \exists r.E(b)$ and thus b in $\mathcal{A}$ corresponds to a top-level conjunct $\exists r.E'$ in C_0^a with $\mathcal{O}_\mathcal{S} \models E' \sqsubseteq E$. ❑

G More Optimization for Figure 3

A further optimization of the approximation from Figure 3 is shown in Figure 4 where $\mathtt{Dis}_{\mathcal{O}_S}(C_0)$ is defined just like $\mathtt{Dis}^{\mathcal{EL}}_{\mathcal{O}_S}(C_0)$ except that the disjunctions and conjunctions are based on all concepts from $\mathtt{sub}^-(\mathcal{O}_S)$ rather than only those formulated in $\mathcal{EL}$. Compared to Figure 3, the second last concept inclusion and Condition 2 have been added, with the aim of invoking the expensive bottommost concept inclusion less often.

Example 6. *Consider the following variation of the $\mathcal{ELU}$ ontology in Proposition 1:*

$$\begin{array}{rrcll} \mathcal{O}_S = \{ & A & \sqsubseteq & B_1 \sqcup B_2, & \\ & \exists r.B_i & \sqsubseteq & B_i, & \text{for } i \in \{1,2\} \\ & B_i \sqcap A' & \sqsubseteq & \exists r.M & \text{for } i \in \{1,2\} \,\}. \end{array}$$

The approximation $\mathcal{O}_T^\omega$ in Figure 3 would contain the CI

$$A' \sqcap \exists r^n.A \sqsubseteq \exists r.M \qquad (\dagger)$$

for all $n \geq 1$. However, $\mathtt{Dis}_{\mathcal{O}_S}(A' \sqcap \exists r^n.A)$ is

$$(A' \sqcap B_1 \sqcap \exists r.B_1) \sqcup (A' \sqcap B_2 \sqcap \exists r.B_2)$$

and we have $\mathcal{O}_S \models \mathtt{Dis}_{\mathcal{O}_S}(A' \sqcap \exists r^n.A) \sqsubseteq \exists r.M$. Consequently, the CIs (†) are not contained in the approximation $\mathcal{O}_T^\omega$ according to Figure 4. It is compensated by the CIs

$$\begin{array}{rcl} A & \sqsubseteq & X_{B_1 \sqcup B_2} \\ \exists r.X_{B_1 \sqcup B_2} & \sqsubseteq & X_{(B_1 \sqcap \exists r.B_1) \sqcup (B_2 \sqcap \exists r.B_2)} \\ A' \sqcap X_{(B_1 \sqcap \exists r.B_1) \sqcup (B_2 \sqcap \exists r.B_2)} & \sqsubseteq & X_{\mathtt{Dis}_{\mathcal{O}_S}(A' \sqcap \exists r^n.A)} \\ X_{\mathtt{Dis}_{\mathcal{O}_S}(A' \sqcap \exists r^n.A)} & \sqsubseteq & \exists r.M \end{array}$$

with the last line being an instantiation of the new second last CI schema in Figure 4.

Proposition 2 and Example 3 provide cases where the last line of Figure 4 is still needed. Arguably, the cases illustrated by these examples are not too likely to occur in practice.

It should be clear that the new CIs in the second last line are sound and thus soundness of the approximation is not compromised. In what follows, we proof completeness. For our proof to go through, we need to assume that $\top$ is always contained in $\mathtt{sub}^-(\mathcal{O}_S)$. We start with observing two technical lemmas, the first one being a variant of Lemma 6.

Lemma 16. *Let C_0 be an $\mathcal{EL}$ concept over $\mathtt{sig}(\mathcal{O}_S)$ decorated with disjunctions from $\mathtt{Dis}(\mathcal{O}_S)$ at leaves such that $\mathtt{Dis}_{\mathcal{O}_S}(C_0)$ has at least two disjuncts. Then $\mathcal{O}_T^- \models C_0^\uparrow \sqsubseteq \mathtt{Dis}_{\mathcal{O}_S}(C_0)$.*

Proof. (sketch) The proof is almost identical to that of Lemma 6, we only sketch the differences. The fact that C_0 is no longer an $\mathcal{EL}$ concept but is decorated with disjunctions from $\mathtt{Dis}(\mathcal{O}_S)$ at leaves is no problem at all. It is simply carried through the entire proof and does not prompt any further modifications. The fact that we work with $\mathtt{Dis}_{\mathcal{O}_S}(C_0)$ instead of $\mathtt{Dis}^{\mathcal{EL}}_{\mathcal{O}_S}(C_0)$, however, does require some changes. In the main proof of Lemma 6, we need a very slight modification of the special chase plus an adapted formulation of Lemma 7.

We define a variant $\vdash'$ of $\vdash$ that only differs in the clause for disjunction:

- $\mathcal{A} \vdash' C_1 \sqcup C_2(a)$ if (a) $\mathcal{A} \models C_1(a)$ or (b) $\mathcal{A} \models C_2(a)$ or (c) $C_1 \sqcup C_2(a) \in \mathcal{A}$.

Now, the only modification of the special chase is that, in Rule 4, we replace $\mathcal{A} \vdash D_2(a)$ with $\mathcal{A} \vdash' D_2(a)$. The adapted formulation of Lemma 7 then reads as follows.

Claim 1. Let $\mathcal{O}$ be an $\mathcal{ELU}_\bot$ ontology and C_0 be an $\mathcal{EL}$ concept over $\mathtt{sig}(\mathcal{O})$ decorated with disjunctions from $\mathtt{Dis}(\mathcal{O}_S)$ at leaves such that $\mathtt{Dis}_{\mathcal{O}_S}(C_0)$ has at least two disjuncts. Then $X_{\mathtt{Dis}_\mathcal{O}(C_0)}(a_0) \in \mathtt{ch}^{\mathrm{sp}}_\mathcal{O}(\mathcal{A}_{C_0})$.

In the proof of Lemma 6, in the claim stating that $\mathcal{O}_T^- \models C_i^\uparrow \sqsubseteq C_{i+1}^\uparrow$ for all $i \geq 0$, we need to adapt the Case of Rule 4, as follows.

Then there are $D_1(a) \in \mathcal{A}_i$ with $D_1 \in \mathtt{Dis}^-(\mathcal{O}_S)$ and $D_2, D_3 \in \mathtt{Dis}(\mathcal{O}_S)$ such that $\mathcal{A}_i \vdash' D_2(a)$, $\mathcal{O}_S \models D_1 \sqcap D_2 \sqsubseteq D_3$, and $\mathcal{A}_{i+1} = \mathcal{A}_i \cup \{D_3(a)\}$. Let E_a be the subconcept of C_i that corresponds to the subtree rooted at a in $\mathcal{A}_i$ and let F_a be the subconcept of C_{i+1} that corresponds to the subtree rooted at a in $\mathcal{A}_{i+1}$. Then $F_a = E_a \sqcap D_3$. From $D_1(a) \in \mathcal{A}$ and $D_1 \in \mathtt{Dis}^-(\mathcal{O}_S)$, we obtain that X_{D_1} is a top-level conjunct of $E_a^\uparrow$. From $\mathcal{A}_i \vdash' D_2(a)$, we obtain an $\mathcal{EL}$ concept D_2' with $\emptyset \models E_a^\uparrow \sqsubseteq {D_2'}^\uparrow$ and $\mathcal{O}_S \models D_2' \sqsubseteq D_2$; we in fact obtain D_2 by 'following' $\mathcal{A}_i \vdash' D_2(a)$ using the definition of $\vdash'$ and whenever we arrive at $\mathcal{A}_i \vdash' F_1 \sqcup F_2(b)$ and this holds because of Case (a) from the definition of $\vdash'$ (resp. Case (b)), replacing the occurrence of $F_1 \sqcup F_2$ in D_2 that gave rise to this with F_1 (resp. F_2). From $\mathcal{O}_S \models D_1 \sqcap D_2 \sqsubseteq D_3$ and $\mathcal{O}_S \models D_2' \sqsubseteq D_2$, we obtain $\mathcal{O}_S \models D_1 \sqcap D_2' \sqsubseteq D_3$ and thus $\mathcal{O}_T^-$ contains the CI $X_{D_1} \sqcap {D_2'}^\uparrow \sqsubseteq D_3^\uparrow$. Consequently, $\mathcal{O}_T^- \models C_i^\uparrow \sqsubseteq C_{i+1}^\uparrow$ as required.

It remains to prove Claim 1. The proof is, in turn, a slight modification of the proof of Lemma 7. Again, we concentrate on sketching the differences. Of course, we replace $\mathtt{Dis}^{\mathcal{EL}}_\mathcal{O}(C_0)$ with $\mathtt{Dis}_\mathcal{O}(C_0)$ throughout the proof. Further, we replace $\vdash$ with $\vdash'$ in property (P1) and in (the two incarnations of) property (P3). We then go on to construct the interpretation $\mathcal{I}$ as before and show, also as before, that it is a model of $\mathcal{O}$. It remains to show that $a_0 \notin \mathtt{Dis}_\mathcal{O}(C_0)$. For this, we first need to observe the following version of Claim 1 in the proof of Lemma 7.

Claim 2. Let $a \in \Delta^\mathcal{I}$ be an individual of $\mathtt{ch}^{\mathrm{sp}}_\mathcal{O}(\mathcal{A}_{C_0})$ and let $C \in \mathtt{sub}(\mathcal{O})$ (not necessarily be an $\mathcal{EL}$ concept). Then $a \in C^\mathcal{I}$ implies

1. $\emptyset \models E_a \sqsubseteq C$ if a is original and disjunctive, and
2. $\mathtt{ch}^{\mathrm{sp}}_\mathcal{O}(\mathcal{A}_{C_0}) \vdash' C(a)$ otherwise.

The proof is by induction on the structure of C. All cases except $C = C_1 \sqcup C_2$ are as in the proof of Claim 1 in the proof of Lemma 7. Due to the use of $\vdash'$ in place of $\vdash$, however, the additional case is straightforward using the semantics and induction hypothesis.

We next argue that a_0 is disjunctive. Assume to the contrary that it is not. It can be verified that Lemma 10 (soundness of the special chase) still holds when the precondition $\mathtt{ch}^{\mathrm{sp}}_\mathcal{O}(\mathcal{A}_{C_0}) \vdash D(a_0)$ is replaced with $\mathtt{ch}^{\mathrm{sp}}_\mathcal{O}(\mathcal{A}_{C_0}) \vdash' D(a_0)$. Let K be the conjunction of all $C \in \mathtt{sub}^-(\mathcal{O})$ such that $a_0 \in C^\mathcal{I}$. By Claim 2, $\mathtt{ch}^{\mathrm{sp}}_\mathcal{O}(\mathcal{A}_{C_0}) \vdash' C(a_0)$ for all such C. Thus the modified Lemma 10 yields $\mathcal{O} \models C_0 \sqsubseteq K$. Since $\mathcal{I}$ is a model of $\mathcal{O}$, this implies that $\mathtt{Dis}_\mathcal{O}(C_0)$ has only the disjunct K, a contradiction to $\mathtt{Dis}_\mathcal{O}(C_0)$ having two disjuncts.

Now back to our proof that $a_0 \notin \mathtt{Dis}_\mathcal{O}(C_0)$. It remains to show that $a_0 \notin \mathtt{Dis}_\mathcal{O}(C_0)^\mathcal{I}$. Then there is a disjunct K of $\mathtt{Dis}_\mathcal{O}(C_0)^\mathcal{I}$ such that $a_0 \in C^\mathcal{I}$ for every conjunct C of K. Since a_0 is disjunctive, Point 1 of Claim 1, yields $\emptyset \models E_{a_0} \sqsubseteq$

C for all conjuncts C of K. Thus $\emptyset \models E_{a_0} \sqsubseteq \mathtt{Dis}_{\mathcal{O}}(C_0)$, in contradiction to our choice of E_{a_0}. ❑

Lemma 17. *Let $D \in \mathtt{Dis}(\mathcal{O}_S)$ be satisfiable w.r.t. $\mathcal{O}_S$ and let C be an $\mathcal{EL}$ concept. Then $\mathcal{O}_S \models D \sqsubseteq \exists r.C$, implies that there is a $D' \in \mathtt{Dis}(\mathcal{O}_S)$ with $\mathcal{O}_S \models D \sqsubseteq \exists r.D'$ and $\mathcal{O}_S \models D' \sqsubseteq C$.*

Proof. For an interpretation $\mathcal{I}$ and $d \in \Delta^{\mathcal{I}}$, let $\mathtt{Con}(d)$ denote the conjunction $K \in \mathtt{Con}(\mathcal{O}_S)$ such that for all $C \in \mathtt{sub}^-(\mathcal{O}_S)$, $d \in C^{\mathcal{I}}$ iff C is a conjunct of K. Now consider all models $\mathcal{I}$ of $\mathcal{O}_S$ and all $d \in \Delta^{\mathcal{I}}$ with $d \in D^{\mathcal{I}}$. We use $\mathcal{K}_{\mathcal{I},d}$ to denote the set of all $K \in \mathtt{Con}(\mathcal{O}_S)$ such that $K = \mathtt{Con}(e)$ for some r-successor e of d in $\mathcal{I}$. Further, we use $\mathfrak{K}$ to denote the set of all $\mathcal{K}_{\mathcal{I},d}$.

Claim. For every $\mathcal{K}_{\mathcal{I},d} \in \mathfrak{K}$, there is a $K \in \mathcal{K}_{\mathcal{I},d}$ with $\mathcal{O}_S \models K \sqsubseteq C$.

Assume that this is not the case. Then for each $K \in \mathcal{K}_{\mathcal{I},d}$ take a tree model $\mathcal{J}_K$ of $\mathcal{O}_S$ with root e_K such that $e_k \in K^{\mathcal{J}_K} \setminus C^{\mathcal{J}_K}$. Then let the interpretation $\mathcal{J}$ be obtained from the unraveling of $\mathcal{I}$ at d by dropping all subtrees rooted at r-successors of the root d, taking the disjoint union with all $\mathcal{J}_K$ and making each e_K an r-successor of d. It can be verified that the resulting $\mathcal{J}$ is a model of $\mathcal{O}_S$ and that $d \in D^{\mathcal{J}} \setminus (\exists r.C)^{\mathcal{J}}$, in contradiction to $\mathcal{O}_S \models D \sqsubseteq \exists r.C$. This finishes the proof of the claim.

Now let D' be the disjunction of all $K \in \mathcal{K}_{\mathcal{I},d}$ with $\mathcal{O}_S \models K \sqsubseteq C$, over all $\mathcal{K}_{\mathcal{I},d} \in \mathfrak{K}$. By the claim, $\mathcal{O}_S \models D' \sqsubseteq C$. Moreover, by definition of $\mathfrak{K}$, we have $\mathcal{O}_S \models D \sqsubseteq \exists r.D'$ and are done. ❑

Now back to the completeness proof of the modified approximation shown in Figure 4. Due to Lemma 15, it suffices to show that for all CIs $F^{\uparrow} \sqsubseteq \exists r.G$ with F and G of the form required for the last line of Figure 4 and Property 2 from Figure 4 not satisfied, then the restriction $\mathcal{O}_T^*$ of $\mathcal{O}_T^{\omega}$ to the first four lines is such that $\mathcal{O}_T^* \models F^{\uparrow} \sqsubseteq \exists r.G$.

Thus take a CI $F^{\uparrow} \sqsubseteq \exists r.G$ as described. Then $D = \mathtt{Dis}_{\mathcal{O}_S}(F)$ has more than one disjunct and $\mathcal{O}_S \models \mathtt{Dis}_{\mathcal{O}_S}(F) \sqsubseteq \exists r.G$. By Lemma 16, $\mathcal{O}_T^- \models F^{\uparrow} \sqsubseteq X_{\mathtt{Dis}_{\mathcal{O}_S}(F)}$. Moreover, $\mathtt{Dis}_{\mathcal{O}_S}(F)$ is satisfiable w.r.t. $\mathcal{O}_S$ since it contains at least two disjuncts. To show that $\mathcal{O}_T^* \models F^{\uparrow} \sqsubseteq \exists r.G$, it thus suffices to establish the following.

Claim. If $\mathcal{O}_S \models D \sqsubseteq C$ with $D \in \mathtt{Dis}^-(\mathcal{O}_S)$ satisfiable w.r.t. $\mathcal{O}_S$ and C an $\mathcal{EL}$ concept, then $\mathcal{O}_T^* \models X_D \sqsubseteq C$.

We prove the claim by induction on C. If $C = A$ is a concept name, then it follows from $\mathcal{O}_S \models D \sqsubseteq C$ that $\mathcal{O}_T^-$ contains a CI $X_D \sqcap X_D \sqsubseteq A$, and thus we are done. The case that $C = C_1 \sqcap C_2$ is straightforward using the semantics and induction hypothesis. Thus assume that $C = \exists r.C_1$. By Lemma 17, there is a $D' \in \mathtt{Dis}(\mathcal{O}_S)$ with $\mathcal{O}_S \models D \sqsubseteq \exists r.D'$ and $\mathcal{O}_S \models D' \sqsubseteq C_1$. We can find a disjunction D'' with at least two disjuncts such that $\mathcal{O}_S \models D' \equiv D''$: if D' has only a single disjunct that does not contain $\top$ as a conjunct, we can choose $D'' = D' \sqcup (D' \sqcap \top)$ and if D' has only a single disjunct that does contain $\top$ as a conjunct, we can choose $D'' = D' \sqcup D^-$ where D^- is D' with conjunct $\top$ removed. We can apply the induction hypothesis to D'' and C_1 to obtain $\mathcal{O}_T^* \models X_{D''} \sqsubseteq C_1$. Moreover, by the second last line in Figure 4, $\mathcal{O}_T^*$ contains $X_D \sqsubseteq \exists r.D''$ and thus we have $\mathcal{O}_T^* \models X_D \sqsubseteq \exists r.C_1$, as required.

Theorem 7. *Let C_0, D_0 be $\mathcal{EL}_{\bot}$ concepts with $D_0 \in \mathtt{sub}(\mathcal{O}_S)$. Then $\mathcal{O}_S \models C_0 \sqsubseteq D_0$ iff $\mathcal{O}_T^- \models C_0 \sqsubseteq D_0$.*

Proof. The 'if' direction follows from Lemma 13. For 'only if', assume that $\mathcal{O}_S \models C_0 \sqsubseteq D_0$. By Lemma 6, $\mathcal{O}_T^- \models C_0 \sqsubseteq \mathtt{Dis}^{\mathcal{EL}}_{\mathcal{O}_S}(C_0)^{\uparrow}$. By definition of $\mathtt{Dis}^{\mathcal{EL}}_{\mathcal{O}_S}(C_0)$, $\mathcal{O}_S \models C_0 \sqsubseteq D_0$ and $D_0 \in \mathtt{sub}(\mathcal{O}_S)$ implies that every top-level conjunct of D_0 is a conjunct in every disjunct of $\mathtt{Dis}^{\mathcal{EL}}_{\mathcal{O}_S}(C_0)$. First assume that there is only a single such disjunct. Then $\mathtt{Dis}^{\mathcal{EL}}_{\mathcal{O}_S}(C_0)$ with conjunct D_0, and since D_0 is an $\mathcal{EL}$ concept it is also a conjunct of $\mathtt{Dis}^{\mathcal{EL}}_{\mathcal{O}_S}(C_0)^{\uparrow}$. Thus $\mathcal{O}_T^- \models C_0 \sqsubseteq \mathtt{Dis}^{\mathcal{EL}}_{\mathcal{O}_S}(C_0)^{\uparrow}$ implies $\mathcal{O}_T^- \models C_0 \sqsubseteq D_0$ as required. Now assume that $\mathtt{Dis}^{\mathcal{EL}}_{\mathcal{O}_S}(C_0)$ has more than one disjunct. Then $\mathtt{Dis}^{\mathcal{EL}}_{\mathcal{O}_S}(C_0)^{\uparrow} = X_{\mathtt{Dis}^{\mathcal{EL}}_{\mathcal{O}_S}(C_0)}$ and $\mathcal{O}_T^-$ contains the CI $X_{\mathtt{Dis}^{\mathcal{EL}}_{\mathcal{O}_S}(C_0)} \sqcap X_{\mathtt{Dis}^{\mathcal{EL}}_{\mathcal{O}_S}(C_0)} \sqsubseteq D_0$. Thus again $\mathcal{O}_T^- \models C_0 \sqsubseteq D_0$. ❑

H Proof of Theorem 8

Theorem 8. *Let $\mathcal{O}_S$ be an $\mathcal{ALC}$ ontology, $\mathtt{sig}(\mathcal{O}_S) = \Sigma$. Then*

1. *the ontology $\mathcal{O}_T^{\omega}$ from Section 4 is an $\mathcal{EL}_{\bot}$ approximation of $\mathcal{O}_S$ w.r.t. ELQ(Σ);*
2. *the ontology $\mathcal{O}_T^-$ from Section 4 is a (finite) $\mathcal{EL}_{\bot}$ approximation of $\mathcal{O}_S$ w.r.t. AQ;*
3. *if $\mathcal{O}_S$ falls within $\mathcal{ELU}$, then the ontology $\mathcal{O}_T^{\omega}$ from Section 3 is an $\mathcal{EL}$ approximation of $\mathcal{O}_S$ w.r.t. ELQ(Σ).*

Proof. Define the unfolding $\mathcal{A}_a^*$ of an ABox $\mathcal{A}$ at an individual names a as the (possibly infinite) ABox whose individuals are words w of the form $a_0 r_1 a_1 \dots a_n$ with $a_0 = a$ and $r_{i+1}(a_i, a_{i+1}) \in \mathcal{A}$ for all $i < n$, and containing the assertions $A(a_0 r_1 a_1 \dots a_n)$ if $A(a_n) \in \mathcal{A}$ and $r(a_0 r_1 a_1 \dots a_n, a_0 r_1 \dots a_n r_{n+1} a_{n+1})$ if $r_{n+1}(a_n, a_{n+1}) \in \mathcal{A}$. The following has been proved in [Lutz and Wolter, 2010].

Fact 1. The following conditions are equivalent for any $\mathcal{EL}_{\bot}$ ontology $\mathcal{O}$ and $\mathcal{EL}$ concept C:

1. $\mathcal{O}, \mathcal{A} \models C(a)$;
2. $\mathcal{O}, \mathcal{A}_a^* \models C(a)$.

We now show the first claim of Theorem 8. The proofs of the remaining two claims are similar and omitted. Let $\mathcal{O}_S$ be an $\mathcal{ALC}$ ontology with $\mathtt{sig}(\mathcal{O}_S) = \Sigma$ and let $\mathcal{O}_T^{\omega}$ be the ontology from Section 4. To show that $\mathcal{O}_T^{\omega}$ is an $\mathcal{EL}_{\bot}$ approximation of $\mathcal{O}_S$ w.r.t. ELQ(Σ), we have to check the conditions of Definition 3. For Condition 1, assume that $C(x)$ is in ELQ(Σ) and that $\mathcal{A}$ is an ABox using no symbols from $\mathtt{sig}(\mathcal{O}_T^{\omega}) \setminus \mathtt{sig}(\mathcal{O}_S)$ such that $\mathcal{O}_T^{\omega}, \mathcal{A} \models C(a)$. By Fact 1, $\mathcal{O}_T^{\omega}, \mathcal{A}_a^* \models C(a)$. Denote by $(\mathcal{A}_a^*)_{|\Sigma}$ the ABox obtained from $\mathcal{A}_a^*$ by removing all assertions using symbols not in Σ. Then still $\mathcal{O}_T^{\omega}, (\mathcal{A}_a^*)_{|\Sigma} \models C(a)$ as $\mathcal{O}_T^{\omega}$ and C do not use any of the symbols used in the assertions we removed. By compactness there exists an $\mathcal{EL}$ concept D corresponding to a finite sub-ABox $\mathcal{A}_1$ of $(\mathcal{A}_a^*)_{|\Sigma}$ with root a such that $\mathcal{O}_T^\omega \models D \sqsubseteq C$. Then $\mathcal{O}_S \models D \sqsubseteq C$ since $\mathcal{O}_T^\omega$ is an $\mathcal{EL}_\bot$ approximation of $\mathcal{O}_S$ and C, D use symbols in Σ only. Then $\mathcal{O}_S, \mathcal{A} \models C(a)$ since there is a homomorphism from $\mathcal{A}_1$ to $\mathcal{A}$ mapping a to a.

For Condition 2, let $C(x)$ be in ELQ(Σ) and $Q = (\mathcal{O}, \Sigma', C(x))$ such that $(\mathcal{O}_S, \Sigma', C(x)) \supseteq Q$, where Σ' is a signature with $\Sigma' \cap \mathtt{sig}(\mathcal{O}_T^\omega) \subseteq \mathtt{sig}(\mathcal{O}_S)$ and $\mathcal{O}$ is an $\mathcal{EL}_\bot$ ontology. To show that $(\mathcal{O}_T^\omega, \Sigma', C(x)) \supseteq Q$, consider a Σ' ABox $\mathcal{A}$ such that $\mathcal{O}, \mathcal{A} \models C(a)$. Then by Fact 1, $\mathcal{O}, \mathcal{A}_a^* \models C(a)$. Hence $\mathcal{O}_S, \mathcal{A}_a^* \models C(a)$ since $(\mathcal{O}_S, \Sigma', C(x)) \supseteq Q$. Then one can argue as above that $\mathcal{O}_T^\omega, \mathcal{A}_a^* \models C(a)$. Hence, by Fact 1, $\mathcal{O}_T^\omega, \mathcal{A} \models C(a)$, as required. ❑