Technische Universität Dresden
Institute for Theoretical Computer Science
Chair for Automata Theory

LTCS–Report

Most Specific Generalizations w.r.t. General $\mathcal{EL}$-TBoxes

Benjamin Zarrieß   Anni-Yasmin Turhan

LTCS-Report 13-06

Postal Address:
Lehrstuhl für Automatentheorie
Institut für Theoretische Informatik
TU Dresden
01062 Dresden

http://lat.inf.tu-dresden.de

Visiting Address:
Nöthnitzer Str. 46
Dresden

Most Specific Generalizations w.r.t. General $\mathcal{EL}$-TBoxes

Benjamin Zarrieß and Anni-Yasmin Turhan*

Institute for Theoretical Computer Science
Technische Universität Dresden, Germany
{zarriess,turhan}@tcs.inf.tu-dresden.de

Abstract

In the area of Description Logics the least common subsumer (lcs) and the most specific concept (msc) are inferences that generalize a set of concepts or an individual, respectively, into a single concept. If computed w.r.t. a general $\mathcal{EL}$-TBox neither the lcs nor the msc need to exist. So far in this setting no exact conditions for the existence of lcs- or msc-concepts are known. This report provides necessary and suffcient conditions for the existence of these two kinds of concepts. For the lcs of a fixed number of concepts and the msc we show decidability of the existence in PTime and polynomial bounds on the maximal role-depth of the lcs- and msc-concepts. The latter allows to compute the lcs and the msc, respectively.

1 Introduction

Description Logics (DL) allow to model application domains in a structured and well-understood way. Due to their formal semantics, DLs can offer powerful reasoning services. In recent years the lightweight DL $\mathcal{EL}$ became popular as an ontology language for large-scale ontologies. $\mathcal{EL}$ provides the logical underpinning of the OWL 2 EL profile of the W3C web ontology language OWL [W3C09], which is used in important life science ontologies, as for instance, SNOMED CT [Spa00] and the thesaurus of the US national cancer institute (NCI) [SdH$^+$07], which contain ten thousands of concepts. The reason for the success of $\mathcal{EL}$ is that it offers limited, but sufficient expressive power, while reasoning can still be done in polynomial time [BBL05].

In DLs basic categories from an application domain can be captured by *concepts* and binary relations by *roles*. Implications between concepts can be specified in the so-called *TBox*. A *general TBox* allows complex concepts on both sides of implications. Facts from the application domain can be captured by *individuals* and their relations in the *ABox*.

Classical inferences for DLs are *subsumption*, which computes the sub- and super-concept relationships of named concepts and *instance checking*, which determines for a given individual whether it belongs to a given concept. Reasoning support for the design and maintenance of large ontologies can be provided by the *bottom-up approach*, which allows to derive a new concept from a set of example individuals, see [BKM99]. For this kind of task the generalization inferences *least common subsumer* (lcs) and *most specific concept* (msc) are investigated for lightweight DLs like $\mathcal{EL}$. The lcs of a collection of concepts is a complex concept that captures

*Partially supported by the German Research Foundation (DFG) in the Collaborative Research Center 912 [Pleaseinsert\PrerenderUnicode{âĂIJ}intopreamble]Highly Adaptive Energy-Efficient Computing[Pleaseinsert\PrerenderUnicode{âĂİ}intopreamble].

all commonalities of these concepts. The msc generalizes an individual into a complex concept, that is the most specific one of which the individual is an instance of.

Unfortunately, neither the lcs nor the msc need to exist, if computed w.r.t. general $\mathcal{EL}$-TBoxes [Baa03] or cyclic ABoxes written in $\mathcal{EL}$ [KM02]. Let's consider the TBox statements:

$$\begin{aligned}
\mathsf{Penicillin} &\sqsubseteq \mathsf{Antibiotic} \sqcap \exists \mathsf{kills}.\mathsf{S\text{-}aureus},\\
\mathsf{Carbapenem} &\sqsubseteq \mathsf{Antibiotic} \sqcap \exists \mathsf{kills}.\mathsf{E\text{-}coli},\\
\mathsf{S\text{-}aureus} &\sqsubseteq \mathsf{Bacterium} \sqcap \exists \mathsf{resistantMutant}.\mathsf{Penicillin},\\
\mathsf{E\text{-}coli} &\sqsubseteq \mathsf{Bacterium} \sqcap \exists \mathsf{resistantMutant}.\mathsf{Carbapenem}
\end{aligned}$$

We want to compute the lcs of $\mathsf{Penicillin}$ and $\mathsf{Carbapenem}$. Now, both concepts are defined by the type of bacterium they kill. These, in turn, are defined by the substance a mutant of theirs is resistant to. This leads to a cyclic definition and thus the common subsumer cannot be captured by a finite $\mathcal{EL}$-concept, since this would need to express the cycle. If computed w.r.t. a TBox that in addition to the above ones also contains the axioms:

$$\begin{aligned}
\mathsf{Antibiotic} &\sqsubseteq \exists \mathsf{kills}.\mathsf{Bacterium},\\
\mathsf{Bacterium} &\sqsubseteq \exists \mathsf{resistantMutant}.\mathsf{Antibiotic},
\end{aligned}$$

then the lcs exists. With the additional statements the lcs of $\mathsf{Penicillin}$ and $\mathsf{Carbapenem}$ is just $\mathsf{Antibiotic}$. We can observe that the existence of the lcs does not merely depend on whether the TBox is cyclic. In fact, for cyclic $\mathcal{EL}$-TBoxes exact conditions for the existence of the lcs have been devised [Baa04]. However, for the case of general $\mathcal{EL}$-TBoxes such conditions are unknown.

There are several approaches to compute generalizations even in this setting. In [LPW10] an extension of $\mathcal{EL}$ with greatest fixpoints was introduced, where the generalization concepts always exist. Computation algorithms for approximative solutions for the lcs were devised in [BST07, PT11a] and for the msc in [KM02]. The last two methods simply compute the generalization concept up to a given k, a bound on the maximal nestings of quantifiers. If the lcs or msc exists and a large enough k was given, then these methods yield the exact solutions. However, to obtain the *least* common subsumer and the *most* specific concept by these methods in practice, a decision procedure for the existence of the lcs or msc, resp., and a method for computing a sufficiently large k are still needed. This paper provides these methods for the lcs and the msc.

In this paper we first introduce basic notions for the DL $\mathcal{EL}$ and its canonical models, which serve as a basis for the characterization of the lcs introduced in the subsequent section. There we show that the characterization can be used to verify whether a given generalization is the most specific one and that the size of the lcs, if it exists, is polynomially bounded by the size of the input, which yields a decision procedure for the existence problem. In Section 4 we show the corresponding results for the msc. We end with some conclusions.

2 Preliminaries

2.1 The Description Logic $\mathcal{EL}$

Let N_C, N_R and N_I be disjoint sets of *concept*, *role* and *individual names*. Let $A \in N_C$ and $r \in N_R$. *$\mathcal{EL}$-concepts* are built according to the syntax rule

$$C ::= \top \mid A \mid C \sqcap D \mid \exists r.C$$

An *interpretation* $\mathcal{I} = (\Delta^\mathcal{I}, \cdot^\mathcal{I})$ consists of a non-empty domain $\Delta^\mathcal{I}$ and a function $\cdot^\mathcal{I}$ that assigns subsets of $\Delta^\mathcal{I}$ to concept names, binary relations on $\Delta^\mathcal{I}$ to role names and elements of $\Delta^\mathcal{I}$ to individual names. The function is extended to complex concepts in the usual way. For a detailed description of the semantic of DLs see [BCM$^+$03].

Let C, D denote $\mathcal{EL}$-concepts. A *general concept inclusions* (GCIs) is an expression of the form $C \sqsubseteq D$. A *(general) TBox* $\mathcal{T}$ is a finite set of GCIs. A GCI $C \sqsubseteq D$ is satisfied in an interpretation $\mathcal{I}$ if $C^\mathcal{I} \subseteq D^\mathcal{I}$. An interpretation $\mathcal{I}$ is a *model* of a TBox $\mathcal{T}$ if it satisfies all GCIs in $\mathcal{T}$.

Let $a, b \in N_I$, $r \in N_R$ and C a concept, then $C(a)$ is a *concept assertion* and $r(a, b)$ a *role assertion*. An interpretation $\mathcal{I}$ satisfies an assertion $C(a)$ if $a^\mathcal{I} \in C^\mathcal{I}$ and $r(a, b)$ if $(a^\mathcal{I}, b^\mathcal{I}) \in r^\mathcal{I}$ holds. An *ABox* $\mathcal{A}$ is a finite set of assertions. An interpretation $\mathcal{I}$ is a *model* of an ABox $\mathcal{A}$ if it satisfies all assertions in $\mathcal{A}$. A *knowledge base* (KB) $\mathcal{K}$ consists of a TBox and an ABox ($\mathcal{K} = (\mathcal{T}, \mathcal{A})$). An interpretation is a model of $\mathcal{K} = (\mathcal{T}, \mathcal{A})$ if it is a model of $\mathcal{T}$ and $\mathcal{A}$.[1]

Important reasoning tasks considered for DLs are *subsumption* and *instance checking*. A concept C *is subsumed by* a concept D w.r.t. a TBox $\mathcal{T}$ (denoted $C \sqsubseteq_\mathcal{T} D$) if $C^\mathcal{I} \subseteq D^\mathcal{I}$ holds in all models $\mathcal{I}$ of $\mathcal{T}$. A concept C is *equivalent* to a concept D w.r.t. a TBox $\mathcal{T}$ (denoted $C \equiv_\mathcal{T} D$) if $C \sqsubseteq_\mathcal{T} D$ and $D \sqsubseteq_\mathcal{T} C$ hold. A reasoning service dealing with a KB is instance checking. An individual *a is instance of* the concept C w.r.t. $\mathcal{K}$ (denoted $\mathcal{K} \models C(a)$) if $a^\mathcal{I} \in C^\mathcal{I}$ holds in all models $\mathcal{I}$ of $\mathcal{K}$. These two reasoning problems can be decided for $\mathcal{EL}$ in polynomial time [BBL05].

Based on subsumption and instance checking our two inferences of interest *least common subsumer* (lcs) and *most specific concept* (msc) are defined.

Definition 1. Let C, D be concepts and $\mathcal{T}$ a TBox. The concept E is the *lcs of C, D w.r.t. $\mathcal{T}$* ($\text{lcs}_\mathcal{T}(C, D)$) if the properties

1. $C \sqsubseteq_\mathcal{T} E$ and $D \sqsubseteq_\mathcal{T} E$, and
2. $C \sqsubseteq_\mathcal{T} F$ and $D \sqsubseteq_\mathcal{T} F$ implies $E \sqsubseteq_\mathcal{T} F$.

are satisfied. If a concept E satisfies Property 1 it is a *common subsumer of C and D w.r.t. $\mathcal{T}$*.

Thus the lcs is unique up to equivalence, while common subsumers are not unique, thus we write $F \in \text{cs}_\mathcal{T}(C, D)$.

The *role depth* ($rd(C)$) of a concept C denotes the maximal nesting depth of $\exists$ in C. If, in Definition 1 the concepts E and F are of role-depth up to k, then E is the *role-depth bounded lcs* (k-$\text{lcs}_\mathcal{T}(C, D)$) of C and D w.r.t. $\mathcal{T}$.

$N_{I,\mathcal{A}}$ is the set of individual names used in an ABox $\mathcal{A}$.

Definition 2. Let $a \in N_{I,\mathcal{A}}$ and $\mathcal{K} = (\mathcal{T}, \mathcal{A})$ a KB. A concept C is the *most specific concept of a w.r.t. $\mathcal{K}$* ($\text{msc}_\mathcal{K}(a)$) if it satisfies:

1. $\mathcal{K} \models C(a)$, and
2. $\mathcal{K} \models D(a)$ implies $C \sqsubseteq_\mathcal{T} D$.

If in the last definition the concepts C and D have a role-depth limited to k, then C is the *role depth bounded msc* of a w.r.t. $\mathcal{K}$ (k-$\text{msc}_\mathcal{K}(a)$). The msc and the k-msc are unique up to equivalence in $\mathcal{EL}$.

[1] Since we only use the DL $\mathcal{EL}$, we write 'concept' instead of '$\mathcal{EL}$-concept' and assume all TBoxes, ABoxes and KBs to be written in $\mathcal{EL}$ in the following.

2.2 Canonical Models and Simulation Relations

The correctness proof of the computation algorithms for the lcs and msc depends on the characterization of subsumption and instance checking. In case of an empty TBox, homomorphisms between syntax trees of concepts [BKM99] were used. A characterization w.r.t. general TBoxes using *canonical models* and *simulations* was given in [LW10a], which we want to use in the following.

Let X be a concept, TBox, ABox or KB, then $\mathsf{sub}(X)$ denotes the set of subconcepts occurring in X.

Definition 3 (canonical model)**.** Let C be a concept and $\mathcal{T}$ a TBox. The *canonical model* $\mathcal{I}_{C,\mathcal{T}}$ of C and $\mathcal{T}$ is defined as follows:

- $\Delta^{\mathcal{I}_{C,\mathcal{T}}} := \{d_C\} \cup \{d_{C'} \mid \exists r.C' \in \mathsf{sub}(C) \cup \mathsf{sub}(\mathcal{T})\}$;
- $A^{\mathcal{I}_{C,\mathcal{T}}} := \{d_D \mid D \sqsubseteq_{\mathcal{T}} A\}$, for all $A \in N_C$;
- $r^{\mathcal{I}_{C,\mathcal{T}}} := \{(d_D, d_{D'}) \mid D \sqsubseteq_{\mathcal{T}} \exists r.D'$ for $\exists r.D' \in \mathsf{sub}(\mathcal{T})$ or $\exists r.D'$ is a conjunct in $D\}$ for all $r \in N_R$.

The notion of a canonical model can be extended to a KB.

Definition 4 (canonical model of a knowledge base)**.** Let $\mathcal{K} = (\mathcal{T}, \mathcal{A})$ be a knowledge base. The canonical model $\mathcal{I}_\mathcal{K}$ of $\mathcal{K}$ is defined as follows:

- $\Delta^{\mathcal{I}_\mathcal{K}} := \{d_a \mid a \in N_{I,\mathcal{A}}\} \cup \{d_C \mid \exists r.C \in \mathsf{sub}(\mathcal{K})\}$;
- $A^{\mathcal{I}_\mathcal{K}} := \{d_a \mid \mathcal{K} \models A(a)\} \cup \{d_C \mid C \sqsubseteq_{\mathcal{T}} A\}$, for all $A \in N_C$;
- $r^{\mathcal{I}_\mathcal{K}} := \{(d_C, d_D) \mid C \sqsubseteq_{\mathcal{T}} \exists r.D, \exists r.D \in \mathsf{sub}(\mathcal{K})\} \cup$
  $\{(d_a, d_b) \mid r(a,b) \in \mathcal{A}\} \cup$
  $\{(d_a, d_C) \mid \mathcal{K} \models \exists r.C(a), \exists r.C \in \mathsf{sub}(\mathcal{K})\}$ for all $r \in N_R$;
- $a^{\mathcal{I}_\mathcal{K}} := d_a$, for all $a \in N_{I,\mathcal{A}}$.

To identify some properties of canonical models we use *simulation relations* between interpretations.

Definition 5 (simulation)**.** Let $\mathcal{I}_1$ and $\mathcal{I}_2$ be interpretations. $\mathcal{S} \subseteq \Delta^{\mathcal{I}_1} \times \Delta^{\mathcal{I}_2}$ is called *simulation* from $\mathcal{I}_1$ to $\mathcal{I}_2$ if all of the following conditions are satisfied:

(S1) For all concept names $A \in N_C$ and all $(e_1, e_2) \in \mathcal{S}$ it holds: $e_1 \in A^{\mathcal{I}_1}$ implies $e_2 \in A^{\mathcal{I}_2}$.

(S2) For all role names $r \in N_R$ and all $(e_1, e_2) \in \mathcal{S}$ and all $f_1 \in \Delta^{\mathcal{I}_1}$ with $(e_1, f_1) \in r^{\mathcal{I}_1}$ there exists $f_2 \in \Delta^{\mathcal{I}_2}$ such that $(e_2, f_2) \in r^{\mathcal{I}_2}$ and $(f_1, f_2) \in \mathcal{S}$.

To denote an interpretation $\mathcal{I}$ with $d \in \Delta^\mathcal{I}$ we write $(\mathcal{I}, d)$. It holds that $(\mathcal{I}, d)$ *is simulated by* $(\mathcal{J}, e)$ (written as $(\mathcal{I}, d) \lesssim (\mathcal{J}, e)$) if there exists a simulation $\mathcal{S} \subseteq \Delta^\mathcal{I} \times \Delta^\mathcal{J}$ with $(d, e) \in \mathcal{S}$. The relation $\lesssim$ is a preorder, i.e. it is reflexive and transitive. $(\mathcal{I}, d)$ is *simulation-equivalent* to $(\mathcal{J}, e)$ (written as $(\mathcal{I}, d) \simeq (\mathcal{J}, e)$) if $(\mathcal{I}, d) \lesssim (\mathcal{J}, e)$ and $(\mathcal{J}, e) \lesssim (\mathcal{I}, d)$ holds.

Now we summarize some important properties of canonical models that were shown in [LW10a].

Lemma 6. *Let C be a concept and $\mathcal{T}$ a TBox.*

1. $d_E \in E^{\mathcal{I}_{C,\mathcal{T}}}$ *for all* $d_E \in \Delta^{\mathcal{I}_{C,\mathcal{T}}}$.
2. $\mathcal{I}_{C,\mathcal{T}}$ *is a model of* $\mathcal{T}$.
3. $(\mathcal{I}_{C,\mathcal{T}}, d_D) \simeq (\mathcal{I}_{C',\mathcal{T}}, d_D)$, *for all concepts* C' *and all* $d_D \in \Delta^{\mathcal{I}_{C,\mathcal{T}}} \cap \Delta^{\mathcal{I}_{C',\mathcal{T}}}$.
4. *For all models* $\mathcal{I}$ *of* $\mathcal{T}$ *and all* $d \in \Delta^{\mathcal{I}}$, *the following conditions are equivalent:*
   (a) $d \in C^{\mathcal{I}}$;
   (b) $(\mathcal{I}_{C,\mathcal{T}}, d_C) \lesssim (\mathcal{I}, d)$.
5. *The following conditions are equivalent:*
   (a) $C \sqsubseteq_{\mathcal{T}} D$;
   (b) $d_C \in D^{\mathcal{I}_{C,\mathcal{T}}}$;
   (c) $(\mathcal{I}_{D,\mathcal{T}}, d_D) \lesssim (\mathcal{I}_{C,\mathcal{T}}, d_C)$.

This lemma gives us a characterization of subsumption. A similar lemma was shown in [LW10b] for the instance relationship.

Lemma 7. *Let* $\mathcal{K}$ *be a knowledge base.* $\mathcal{I}_{\mathcal{K}}$ *satisfies the following properties:*

1. $\mathcal{I}_{\mathcal{K}}$ *is a model of* $\mathcal{K}$.
2. *The following conditions are equivalent:*
   (a) $\mathcal{K} \models C(a)$;
   (b) $d_a \in C^{\mathcal{I}_{\mathcal{K}}}$.

Next we recall some known operations on interpretations.

Taking an element of the domain of an interpretation as the root, the interpretation can be unraveled into a possibly infinite tree. The nodes of the tree are words that correspond to paths starting in d. Now, $\pi = dr_1d_1r_2d_2r_3\cdots$ is a path in an interpretation $\mathcal{I}$ if the domain elements d_i and d_{i+1} are connected via $r_{i+1}^{\mathcal{I}}$ for all i.

Definition 8 (tree unraveling of an interpretation)**.** Let $\mathcal{I}$ be an interpretation w.r.t. the names N_C and N_R with $d \in \Delta^{\mathcal{I}}$. The *tree unraveling* $\mathcal{I}_d$ of $\mathcal{I}$ in d is defined as follows:

$$\begin{aligned}
\Delta^{\mathcal{I}_d} &:= \{dr_1d_1r_2\cdots r_nd_n \mid (d_i, d_{i+1}) \in r_{i+1}^{\mathcal{I}} \land 0 \le i < n \land d_0 = d\};\\
A^{\mathcal{I}_d} &:= \{\sigma d' \mid \sigma d' \in \Delta^{\mathcal{I}_d} \land d' \in A^{\mathcal{I}}\}, \text{ for all } A \in N_C;\\
r^{\mathcal{I}_d} &:= \{(\sigma, \sigma r d') \mid (\sigma, \sigma r d') \in \Delta^{\mathcal{I}_d} \times \Delta^{\mathcal{I}_d}\}, \text{ for all } r \in N_R.
\end{aligned}$$

The *length* of an element $\sigma \in \Delta^{\mathcal{I}_d}$, denoted by $|\sigma|$, is the number of role names occurring in σ. If σ is of the form $dr_1d_1r_2\cdots r_md_m$, then d_m is the *tail* of σ denoted by $\mathsf{tail}(\sigma) = d_m$. The interpretation $\mathcal{I}_d^{\ell}$ denotes the finite subtree rooted in d of the tree unraveling $\mathcal{I}_d$ containing all elements up to depth ℓ. Such a finite tree can be translated into a complex concept which is called *characteristic concept*.

Definition 9 (characteristic concept)**.** Let $(\mathcal{I}, d)$ be an interpretation. The ℓ-*characteristic concept* $X^{\ell}(\mathcal{I}, d)$ is defined as follows: [2]

$$\begin{aligned}
X^0(\mathcal{I}, d) &:= \bigsqcap\{A \in N_C \mid d \in A^{\mathcal{I}}\}\\
X^{\ell}(\mathcal{I}, d) &:= X^0(\mathcal{I}, d) \sqcap \bigsqcap_{r \in N_R} \bigsqcap\{\exists r.X^{\ell-1}(\mathcal{I}, d') \mid (d, d') \in r^{\mathcal{I}}\}
\end{aligned}$$

[2] For a set M of concepts we write $\bigsqcap M$ as shorthand for $\bigsqcap_{F \in M} F$. If M is empty, then $\bigsqcap M$ is equal to $\top$.

Later we will need the following basic property of characteristic concepts that was shown in [LPW10].

Lemma 10. *Let $(\mathcal{I}, d)$ and $(\mathcal{J}, e)$ be interpretations. Then $e \in (X^\ell(\mathcal{I}, d))^\mathcal{J}$ if and only if $(\mathcal{I}^\ell_d, d) \lesssim (\mathcal{J}, e)$.*

Another operation that we will use later is the product of two interpretations that is defined as follows.

Definition 11 (product interpretation)**.** Let $\mathcal{I}$ and $\mathcal{J}$ be interpretations. The *product interpretation* $\mathcal{I} \times \mathcal{J}$ is defined by

$$\begin{aligned}
\Delta^{\mathcal{I}\times\mathcal{J}} &:= \Delta^\mathcal{I} \times \Delta^\mathcal{J};\\
A^{\mathcal{I}\times\mathcal{J}} &:= \{(d,e) \mid (d,e) \in \Delta^{\mathcal{I}\times\mathcal{J}} \wedge d \in A^\mathcal{I} \wedge e \in A^\mathcal{J}\}, \text{ for all } A \in N_C;\\
r^{\mathcal{I}\times\mathcal{J}} &:= \{((d,e),(f,g)) \mid ((d,e),(f,g)) \in \Delta^{\mathcal{I}\times\mathcal{J}} \times \Delta^{\mathcal{I}\times\mathcal{J}}\\
&\qquad\qquad \wedge (d,f) \in r^\mathcal{I} \wedge (e,g) \in r^\mathcal{J}\}, \text{ for all } r \in N_R.
\end{aligned}$$

3 Existence of the Least Common Subsumer

In this section we develop a decision procedure for the problem whether for two given concepts and a given TBox the least common subsumer of these two concepts exists w.r.t. the given TBox. If not stated otherwise, the two input concepts are denoted by C and D and the TBox by $\mathcal{T}$.

Similar to the approach used in [Baa04] we proceed by the following steps:

1. Devise a method to identify lcs-candidates. The set of lcs-candidates is a possibly infinite set of common subsumers of C and D w.r.t. $\mathcal{T}$, such that if the lcs exists then one of these lcs-candidates actually is the lcs.

2. Characterize the existence of the lcs. Find a condition such that the problem whether a given common subsumer of C and D w.r.t. $\mathcal{T}$ is least (w.r.t. $\sqsubseteq_\mathcal{T}$), can be decided by testing this condition.

3. Establish an upper bound on the role-depth of the lcs. We give a bound ℓ such that *if the lcs exists*, then it has a role-depth less or equal ℓ. By such an upper bound one needs to check only for finitely many of the lcs-candidates if they are least (w.r.t. $\sqsubseteq_\mathcal{T}$).

The next subsection addresses the first two problems, afterwards we show that such a desired upper bound exists.

3.1 Characterizing the existence of the lcs

In this section canonical models and simulation relations are used to obtain in a first step a set of possible candidates for the lcs and then to characterize whether a common subsumer is least or not.

In [PT11a] so called role-depth bounded least common subsumers were introduced as approximations of the lcs, denoted by k-lcs$_\mathcal{T}(C, D)$. For a fixed natural number k the k-lcs$_\mathcal{T}(C, D)$ is a common subsumer that is the least one of all common subsumers with a role-depth $\leq k$. To obtain the k-lcs$_\mathcal{T}(C, D)$ we build the product of the canonical models $(\mathcal{I}_{C,\mathcal{T}}, d_C)$ and $(\mathcal{I}_{D,\mathcal{T}}, d_D)$ and then take the k-characteristic concept of this product model. This product construction is adopted from [Baa03, LPW10], where a similar construction was used to define the lcs in $\mathcal{EL}$ with gfp-semantics and in the DL $\mathcal{EL}^\nu$ respectively.

In order to prove that the k-lcs can be computed as described above, we first show some properties of product models and their characteristic concepts.

Lemma 12. *Let $\mathcal{I}_{C,\mathcal{T}} \times \mathcal{I}_{D,\mathcal{T}}$ and $\mathcal{I}_{E,\mathcal{T}} \times \mathcal{I}_{F,\mathcal{T}}$ be products of canonical models with $(d_G, d_H) \in \Delta^{\mathcal{I}_{C,\mathcal{T}} \times \mathcal{I}_{D,\mathcal{T}}} \cap \Delta^{\mathcal{I}_{E,\mathcal{T}} \times \mathcal{I}_{F,\mathcal{T}}}$.*

1. *For any $k \in \mathbb{N}$ it holds that $X^k(\mathcal{I}_{C,\mathcal{T}} \times \mathcal{I}_{D,\mathcal{T}}, (d_G, d_H)) = X^k(\mathcal{I}_{E,\mathcal{T}} \times \mathcal{I}_{F,\mathcal{T}}, (d_G, d_H))$*
2. *Let N be a concept. $(d_G, d_H) \in N^{\mathcal{I}_{C,\mathcal{T}} \times \mathcal{I}_{D,\mathcal{T}}}$ iff $G \sqsubseteq_\mathcal{T} N$ and $H \sqsubseteq_\mathcal{T} N$.*

Proof.
1. By Claim 3 of Lemma 6 it is implied that for any k $X^k(\mathcal{I}_{C,\mathcal{T}}, d_G) = X^k(\mathcal{I}_{E,\mathcal{T}}, d_G)$ and $X^k(\mathcal{I}_{D,\mathcal{T}}, d_H) = X^k(\mathcal{I}_{F,\mathcal{T}}, d_H)$, respectively. Obviously, this implies the claim.
2. This claim follows directly from the definition of products of interpretations and Claim 5 of Lemma 6.

□

Now we show that the k-characteristic concept of $(\mathcal{I}_{C,\mathcal{T}} \times \mathcal{I}_{D,\mathcal{T}}, (d_C, d_D))$ yields the k-lcs$_\mathcal{T}(C, D)$.

Lemma 13. *Let k be a natural number.*

1. *$X^k(\mathcal{I}_{C,\mathcal{T}} \times \mathcal{I}_{D,\mathcal{T}}, (d_C, d_D)) \in cs_\mathcal{T}(C, D)$.*
2. *Let E be a concept with $rd(E) \leq k$ and $C \sqsubseteq_\mathcal{T} E$ and $D \sqsubseteq_\mathcal{T} E$.*
   It holds that $X^k(\mathcal{I}_{C,\mathcal{T}} \times \mathcal{I}_{D,\mathcal{T}}, (d_C, d_D)) \sqsubseteq_\mathcal{T} E$.

Proof. 1. We show the claim by induction on k.

$k = 0$: By Definition 9 it holds that

$$X^0(\mathcal{I}_{C,\mathcal{T}} \times \mathcal{I}_{D,\mathcal{T}}, (d_C, d_D)) = \bigsqcap\{A \in N_C \mid (d_C, d_D) \in A^{\mathcal{I}_{C,\mathcal{T}} \times \mathcal{I}_{D,\mathcal{T}}}\}. \tag{1}$$

For any concept name A in this conjunction it holds that $(d_C, d_D) \in A^{\mathcal{I}_{C,\mathcal{T}} \times \mathcal{I}_{D,\mathcal{T}}}$ and therefore $d_C \in A^{\mathcal{I}_{C,\mathcal{T}}}$ and $d_D \in A^{\mathcal{I}_{D,\mathcal{T}}}$. From point 5 of Lemma 6 it follows that $C \sqsubseteq_\mathcal{T} A$ and $D \sqsubseteq_\mathcal{T} A$ and therefore $C \sqsubseteq_\mathcal{T} X^0(\mathcal{I}_{C,\mathcal{T}} \times \mathcal{I}_{D,\mathcal{T}}, (d_C, d_D))$ and $D \sqsubseteq_\mathcal{T} X^0(\mathcal{I}_{C,\mathcal{T}} \times \mathcal{I}_{D,\mathcal{T}}, (d_C, d_D))$.

$k > 0$: By applying the definition of X^k we get

$$\begin{aligned} X^k(\mathcal{I}_{C,\mathcal{T}} \times \mathcal{I}_{D,\mathcal{T}}, (d_C, d_D)) =& X^0(\mathcal{I}_{C,\mathcal{T}} \times \mathcal{I}_{D,\mathcal{T}}, (d_C, d_D)) \sqcap \\ & \bigsqcap_{r \in N_R} \bigsqcap\{\exists r.X^{k-1}(\mathcal{I}_{C,\mathcal{T}} \times \mathcal{I}_{D,\mathcal{T}}, (d_E, d_F)) \\ & \mid ((d_C, d_D), (d_E, d_F)) \in r^{\mathcal{I}_{C,\mathcal{T}} \times \mathcal{I}_{D,\mathcal{T}}}\}. \end{aligned} \tag{2}$$

From Lemma 12.1 it follows that $X^{k-1}(\mathcal{I}_{C,\mathcal{T}} \times \mathcal{I}_{D,\mathcal{T}}, (d_E, d_F)) = X^{k-1}(\mathcal{I}_{E,\mathcal{T}} \times \mathcal{I}_{F,\mathcal{T}}, (d_E, d_F))$. Now the induction hypothesis can be applied as follows:

$$\begin{aligned} E &\sqsubseteq_\mathcal{T} X^{k-1}(\mathcal{I}_{E,\mathcal{T}} \times \mathcal{I}_{F,\mathcal{T}}, (d_E, d_F)) \\ F &\sqsubseteq_\mathcal{T} X^{k-1}(\mathcal{I}_{E,\mathcal{T}} \times \mathcal{I}_{F,\mathcal{T}}, (d_E, d_F)). \end{aligned}$$

By Lemma 12.1 it is implied that

$$E \sqsubseteq_{\mathcal{T}} X^{k-1}(\mathcal{I}_{C,\mathcal{T}} \times \mathcal{I}_{D,\mathcal{T}}, (d_E, d_F))$$
$$F \sqsubseteq_{\mathcal{T}} X^{k-1}(\mathcal{I}_{C,\mathcal{T}} \times \mathcal{I}_{D,\mathcal{T}}, (d_E, d_F))$$

and by Lemma 6.5

$$d_E \in (X^{k-1}(\mathcal{I}_{C,\mathcal{T}} \times \mathcal{I}_{D,\mathcal{T}}, (d_E, d_F)))^{\mathcal{I}_{E,\mathcal{T}}}$$
$$d_F \in (X^{k-1}(\mathcal{I}_{C,\mathcal{T}} \times \mathcal{I}_{D,\mathcal{T}}, (d_E, d_F)))^{\mathcal{I}_{F,\mathcal{T}}}.$$

From Lemma 6.3 it follows $(\mathcal{I}_{E,\mathcal{T}}, d_E) \simeq (\mathcal{I}_{C,\mathcal{T}}, d_E)$ and $(\mathcal{I}_{F,\mathcal{T}}, d_F) \simeq (\mathcal{I}_{D,\mathcal{T}}, d_F)$ consequently

$$d_E \in (X^{k-1}(\mathcal{I}_{C,\mathcal{T}} \times \mathcal{I}_{D,\mathcal{T}}, (d_E, d_F)))^{\mathcal{I}_{C,\mathcal{T}}}$$
$$d_F \in (X^{k-1}(\mathcal{I}_{C,\mathcal{T}} \times \mathcal{I}_{D,\mathcal{T}}, (d_E, d_F)))^{\mathcal{I}_{D,\mathcal{T}}}.$$

and by definition of the product of interpretation it holds that

$$(d_E, d_F) \in (X^{k-1}(\mathcal{I}_{C,\mathcal{T}} \times \mathcal{I}_{D,\mathcal{T}}, (d_E, d_F)))^{\mathcal{I}_{C,\mathcal{T}} \times \mathcal{I}_{D,\mathcal{T}}}.$$

Since (d_E, d_F) is an r-successor of (d_C, d_D) in $\mathcal{I}_{C,\mathcal{T}} \times \mathcal{I}_{D,\mathcal{T}}$ it is implied that

$$(d_C, d_D) \in (\exists r.X^{k-1}(\mathcal{I}_{C,\mathcal{T}} \times \mathcal{I}_{D,\mathcal{T}}, (d_E, d_F)))^{\mathcal{I}_{C,\mathcal{T}} \times \mathcal{I}_{D,\mathcal{T}}}$$

and with Lemma 12.2 we obtain

$$C \sqsubseteq_{\mathcal{T}} \exists r.X^{k-1}(\mathcal{I}_{C,\mathcal{T}} \times \mathcal{I}_{D,\mathcal{T}}, (d_E, d_F))$$
$$D \sqsubseteq_{\mathcal{T}} \exists r.X^{k-1}(\mathcal{I}_{C,\mathcal{T}} \times \mathcal{I}_{D,\mathcal{T}}, (d_E, d_F)).$$

As shown in the base case $X^0(\mathcal{I}_{C,\mathcal{T}} \times \mathcal{I}_{D,\mathcal{T}}, (d_C, d_D))$ is also a common subsumer of C and D w.r.t. $\mathcal{T}$. It is now implied that $X^k(\mathcal{I}_{C,\mathcal{T}} \times \mathcal{I}_{D,\mathcal{T}}, (d_C, d_D))$ is a common subsumer of C and D w.r.t. $\mathcal{T}$.

2. The claim is proven by induction on the role-depth of an arbitrary common subsumer E of C and D w.r.t. $\mathcal{T}$ with $rd(E) \leq k$.

$rd(E) = 0:$ E is a conjunction of concept names of the form $\bigsqcap_i A_i$. We show that the concept names A_i occur in the conjunction $X^0(\mathcal{I}_{C,\mathcal{T}} \times \mathcal{I}_{D,\mathcal{T}}, (d_C, d_D))$. Since $C \sqsubseteq_{\mathcal{T}} E$ and $D \sqsubseteq_{\mathcal{T}} E$ holds, it follows by Lemma 6.5 that $d_C \in E^{\mathcal{I}_{C,\mathcal{T}}}$ and $d_D \in E^{\mathcal{I}_{D,\mathcal{T}}}$. So we have that $d_C \in A_i^{\mathcal{I}_{C,\mathcal{T}}}$ and $d_D \in A_i^{\mathcal{I}_{D,\mathcal{T}}}$ for all i and $(d_C, d_D) \in A_i^{\mathcal{I}_{C,\mathcal{T}} \times \mathcal{I}_{D,\mathcal{T}}}$ for all i. By definition of $X^k(\mathcal{I}_{C,\mathcal{T}} \times \mathcal{I}_{D,\mathcal{T}}, (d_C, d_D))$ and (1) it is implied that $X^0(\mathcal{I}_{C,\mathcal{T}} \times \mathcal{I}_{D,\mathcal{T}}, (d_C, d_D)) \sqsubseteq_{\mathcal{T}} E$.

$rd(E) = n > 0:$ Let
$$E = A_1 \sqcap \cdots \sqcap A_\ell \sqcap \exists r_1.E'_1 \sqcap \cdots \sqcap \exists r_m.E'_m$$

It can be shown like in the base case that the conjunction $A_1 \sqcap ... \sqcap A_\ell$ subsumes $X^k(\mathcal{I}_{C,\mathcal{T}} \times \mathcal{I}_{D,\mathcal{T}}, (d_C, d_D))$. Let $\exists r_j.E'_j$ with $1 \leq j \leq m$ be an existential restriction in E. Since it holds that $C \sqsubseteq_{\mathcal{T}} \exists r_j.E'_j$ and $D \sqsubseteq_{\mathcal{T}} \exists r_j.E'_j$, we get $d_C \in (\exists r_j.E'_j)^{\mathcal{I}_{C,\mathcal{T}}}$ and $d_D \in (\exists r_j.E'_j)^{\mathcal{I}_{D,\mathcal{T}}}$ by Lemma 6.5. There are r_j-successors d_G and d_H of d_C and d_D in $\mathcal{I}_{C,\mathcal{T}}$ and $\mathcal{I}_{D,\mathcal{T}}$, respectively, with $d_G \in (E'_j)^{\mathcal{I}_{C,\mathcal{T}}}$ and $d_H \in (E'_j)^{\mathcal{I}_{D,\mathcal{T}}}$. It holds that

$$d_G \in (E'_j)^{\mathcal{I}_{C,\mathcal{T}}}$$
$$\Rightarrow (\mathcal{I}_{E'_j,\mathcal{T}}, d_{E'_j}) \lesssim (\mathcal{I}_{C,\mathcal{T}}, d_G) \simeq (\mathcal{I}_{G,\mathcal{T}}, d_G) \text{ (by Lemma 6.4 and 6.3)}$$
$$\Rightarrow G \sqsubseteq_{\mathcal{T}} E'_j \text{ (by Lemma 6.5).}$$

The same argument holds for d_H. By induction hypothesis and $rd(E'_j) = n - 1$ we now have that $X^{n-1}(\mathcal{I}_{G,\mathcal{T}} \times \mathcal{I}_{H,\mathcal{T}}, (d_G, d_H)) \sqsubseteq_{\mathcal{T}} E'_j$. From Lemma 12.1 it follows that

$$X^{n-1}(\mathcal{I}_{G,\mathcal{T}} \times \mathcal{I}_{H,\mathcal{T}}, (d_G, d_H)) = X^{n-1}(\mathcal{I}_{C,\mathcal{T}} \times \mathcal{I}_{D,\mathcal{T}}, (d_G, d_H))$$

and therefore $X^{n-1}(\mathcal{I}_{C,\mathcal{T}} \times \mathcal{I}_{D,\mathcal{T}}, (d_G, d_H)) \sqsubseteq_{\mathcal{T}} E'_j$ and

$$\exists r_j.X^{n-1}(\mathcal{I}_{C,\mathcal{T}} \times \mathcal{I}_{D,\mathcal{T}}, (d_G, d_H)) \sqsubseteq_{\mathcal{T}} \exists r_j.E'_j.$$

Since $\exists r_j.X^{n-1}(\mathcal{I}_{C,\mathcal{T}} \times \mathcal{I}_{D,\mathcal{T}}, (d_G, d_H))$ is a conjunct in $X^n(\mathcal{I}_{C,\mathcal{T}} \times \mathcal{I}_{D,\mathcal{T}}, (d_C, d_D))$, it is implied that $X^n(\mathcal{I}_{C,\mathcal{T}} \times \mathcal{I}_{D,\mathcal{T}}, (d_C, d_D)) \sqsubseteq_{\mathcal{T}} \exists r_j.E'_j$.

□

In the following we take $X^k(\mathcal{I}_{C,\mathcal{T}} \times \mathcal{I}_{D,\mathcal{T}}, (d_C, d_D))$ as representation of k-lcs$_{\mathcal{T}}(C, D)$. It is implied by Lemma 13 that the set of k-characteristic concepts of the product model $(\mathcal{I}_{C,\mathcal{T}} \times \mathcal{I}_{D,\mathcal{T}}, (d_C, d_D))$ for all k is the set of possible candidates for the lcs$_{\mathcal{T}}(C, D)$. This can be stated as follows.

Corollary 14. *The lcs$_{\mathcal{T}}(C, D)$ exists if and only if there exists a $k \in \mathbb{N}$ such that for all $\ell \in \mathbb{N}$: k-lcs$_{\mathcal{T}}(C, D) \sqsubseteq_{\mathcal{T}} \ell$-lcs$_{\mathcal{T}}(C, D)$.*

Obviously, this doesn't yield a decision procedure for the problem whether the k-lcs$_{\mathcal{T}}(C, D)$ is the lcs, since subsumption cannot be checked for infinitely many ℓ in finite time.

Next, we address step 2 and show a condition on the common subsumers that decides whether a common subsumer is the least one or not. The main idea is that the product model captures all commonalities of the input concepts by means of canonical models. Thus we compare the canonical models of the common subsumers and the product model using $\lesssim$ and simulation equivalence $\simeq$.

First it can be stated that the canonical model of the k-lcs simulates the tree unraveling of the product model limited to depth k.

Lemma 15. *Let $\mathcal{J}_{(d_C,d_D)}$ be the tree unraveling of $(\mathcal{I}_{C,\mathcal{T}} \times \mathcal{I}_{D,\mathcal{T}}, (d_C, d_D))$ in (d_C, d_D) and K the k-lcs$_{\mathcal{T}}(C, D)$ w.r.t. $\mathcal{T}$. It holds that $\mathcal{J}^k_{(d_C,d_D)} \lesssim (\mathcal{I}_{K,\mathcal{T}}, d_K)$.*

Proof. The concept $X^k(\mathcal{I}_{C,\mathcal{T}} \times \mathcal{I}_{D,\mathcal{T}}, (d_C, d_D))$ is by Lemma 13 a common subsumer of C, D w.r.t. $\mathcal{T}$. Since $X^k(\mathcal{I}_{C,\mathcal{T}} \times \mathcal{I}_{D,\mathcal{T}}, (d_C, d_D))$ has role-depth $\leq k$, it is implied that $K \sqsubseteq_{\mathcal{T}} X^k(\mathcal{I}_{C,\mathcal{T}} \times \mathcal{I}_{D,\mathcal{T}}, (d_C, d_D))$ and therefore $d_K \in (X^k(\mathcal{I}_{C,\mathcal{T}} \times \mathcal{I}_{D,\mathcal{T}}, (d_C, d_D)))^{\mathcal{I}_{K,\mathcal{T}}}$ by Lemma 6.5. From Lemma 10 it now follows $\mathcal{J}^k_{(d_C,d_D)} \lesssim (\mathcal{I}_{K,\mathcal{T}}, d_K)$. □

The following lemma recalls a simple property about products of interpretations.

Lemma 16 ([LPW10])**.** *Let $(\mathcal{J}, e), (\mathcal{I}_1, d_1)$ and $(\mathcal{I}_2, d_2)$ be interpretations. If $(\mathcal{J}, e) \lesssim (\mathcal{I}_1, d_1)$ and $(\mathcal{J}, e) \lesssim (\mathcal{I}_2, d_2)$, then $(\mathcal{J}, e) \lesssim (\mathcal{I}_1 \times \mathcal{I}_2, (d_1, d_2))$.*

Now we show that a common subsumer is the lcs if and only if its canonical model is simulation-equivalent to the product of the canonical models of the input concepts.

Lemma 17. *Let E be a concept.*
E is the lcs of C and D w.r.t. $\mathcal{T}$ iff $(\mathcal{I}_{C,\mathcal{T}} \times \mathcal{I}_{D,\mathcal{T}}, (d_C, d_D)) \simeq (\mathcal{I}_{E,\mathcal{T}}, d_E)$.

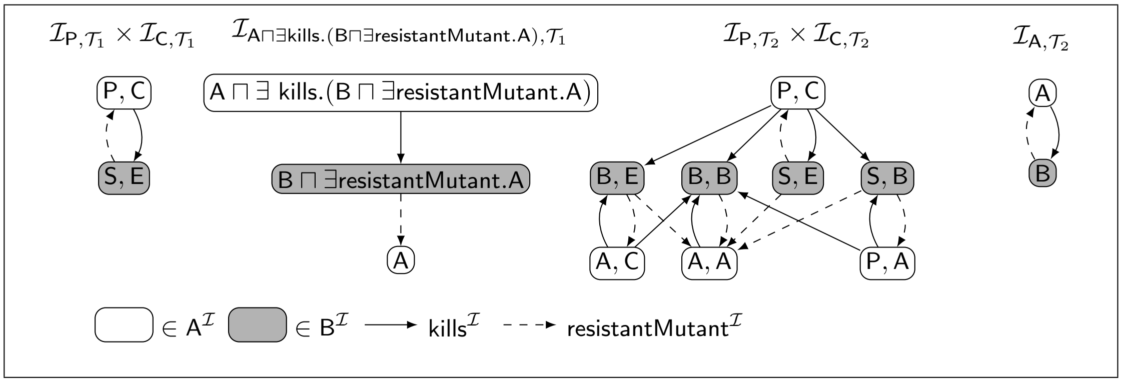

Figure 1: Product of canonical models of $\mathcal{T}_1$ and $\mathcal{T}_2$

The proof idea of this claim can be outlined as follows:

Assume $(\mathcal{I}_{E,\mathcal{T}}, d_E)$ is simulation-equivalent to the product model. We need to show that $E \equiv_{\mathcal{T}} \text{lcs}_{\mathcal{T}}(C, D)$.

For any $F \in \text{cs}_{\mathcal{T}}(C, D)$ it holds by Lemma 6.5 that $(\mathcal{I}_{F,\mathcal{T}}, d_F)$ is simulated by $(\mathcal{I}_{C,\mathcal{T}}, d_C)$ and by $(\mathcal{I}_{D,\mathcal{T}}, d_D)$ and therefore also by $(\mathcal{I}_{C,\mathcal{T}} \times \mathcal{I}_{D,\mathcal{T}}, (d_C, d_D))$. By transitivity of $\lesssim$ it is implied that $(\mathcal{I}_{F,\mathcal{T}}, d_F) \lesssim (\mathcal{I}_{E,\mathcal{T}}, d_E)$ and $E \sqsubseteq_{\mathcal{T}} F$ by Lemma 6. Therefore $E \equiv_{\mathcal{T}} \text{lcs}_{\mathcal{T}}(C, D)$.

For the other direction assume $E \equiv_{\mathcal{T}} \text{lcs}_{\mathcal{T}}(C, D)$. It has to be shown that $(\mathcal{I}_{E,\mathcal{T}}, d_E)$ simulates the product model. Let $\mathcal{J}_{(d_C,d_D)}$ be the tree unraveling of the product model. Since E is more specific than the k-characteristic concepts of the product model for all k (by Corollary 14), $(\mathcal{I}_{E,\mathcal{T}}, d_E)$ simulates the subtree $\mathcal{J}^k_{(d_C,d_D)}$ of $\mathcal{J}_{(d_C,d_D)}$ limited to elements up to depth k, for all k. For each k we consider the maximal simulation from $\mathcal{J}^k_{(d_C,d_D)}$ to $(\mathcal{I}_{E,\mathcal{T}}, d_E)$. Note that $((d_C, d_D), d_E)$ is contained in any of these simulations. Let σ be an element of $\Delta^{\mathcal{J}_{(d_C,d_D)}}$ at an arbitrary depth ℓ. We show how to determine the elements of $\Delta^{\mathcal{I}_{E,\mathcal{T}}}$, that simulate this fixed element σ. Let $\mathcal{S}_n(\sigma)$ be the maximal set of elements from $\Delta^{\mathcal{I}_{E,\mathcal{T}}}$ that simulate σ in each of the trees $\mathcal{J}^n_{(d_C,d_D)}$ with $n \geq \ell$. We can observe that the infinite sequence $(\mathcal{S}_{\ell+i}(\sigma))_{i=0,1,2,\ldots}$ is decreasing (w.r.t. $\supseteq$). Therefore at a certain depth we reach a fixpoint set. This fixpoint set exists for any σ. It can be shown that the union of all these fixpoint sets yields a simulation from the product model to $(\mathcal{I}_{E,\mathcal{T}}, d_E)$.

Proof of Lemma 17. "⇒":
Assume that E is the lcs of C, D w.r.t. $\mathcal{T}$, thus $C \sqsubseteq_{\mathcal{T}} E$ and $D \sqsubseteq_{\mathcal{T}} E$ and by Lemma 6.5 $(\mathcal{I}_{E,\mathcal{T}}, d_E) \lesssim (\mathcal{I}_{C,\mathcal{T}}, d_C)$ and $(\mathcal{I}_{E,\mathcal{T}}, d_E) \lesssim (\mathcal{I}_{D,\mathcal{T}}, d_D)$ holds. It is now implied by Lemma 16 that

$$(\mathcal{I}_{E,\mathcal{T}}, d_E) \lesssim (\mathcal{I}_{C,\mathcal{T}} \times \mathcal{I}_{D,\mathcal{T}}, (d_C, d_D)). \tag{3}$$

We now show $(\mathcal{I}_{C,\mathcal{T}} \times \mathcal{I}_{D,\mathcal{T}}, (d_C, d_D)) \lesssim (\mathcal{I}_{E,\mathcal{T}}, d_E)$ by constructing a simulation from the tree unraveling $\mathcal{J}_{(d_C,d_D)}$ of $(\mathcal{I}_{C,\mathcal{T}} \times \mathcal{I}_{D,\mathcal{T}}, (d_C, d_D))$ to $(\mathcal{I}_{E,\mathcal{T}}, d_E)$. We first write $\mathcal{J}_{(d_C,d_D)}$ as an infinite union of the subtrees $\mathcal{J}^k_{(d_C,d_D)}$.

$$\Delta^{\mathcal{J}_{(d_C,d_D)}} = \bigcup_{k=0} \Delta^{\mathcal{J}^k_{(d_C,d_D)}}, \tag{4}$$

$$A^{\mathcal{J}_{(d_C,d_D)}} = \bigcup_{k=0} A^{\mathcal{J}^k_{(d_C,d_D)}}, \text{ for all } A \in N_C \tag{5}$$

$$r^{\mathcal{J}_{(d_C,d_D)}} = \bigcup_{k=0} r^{\mathcal{J}^k_{(d_C,d_D)}}, \text{ for all } r \in N_R \tag{6}$$

Let K be the k-lcs$_{\mathcal{T}}(C,D)$ for an arbitrary k. By Lemma 15 we have:

$$\mathcal{J}^k_{(d_C,d_D)} \lesssim (\mathcal{I}_{K,\mathcal{T}}, d_K). \tag{7}$$

Since E is the lcs, E is subsumed by K w.r.t. $\mathcal{T}$ and therefore it holds (by Lemma 6.5) that $(\mathcal{I}_{K,\mathcal{T}}, d_K) \lesssim (\mathcal{I}_{E,\mathcal{T}}, d_E)$. With (7) and transitivity of $\lesssim$ we have

$$\mathcal{J}^n_{(d_C,d_D)} \lesssim (\mathcal{I}_{E,\mathcal{T}}, d_E)$$

for all $n \in \mathbb{N}$. If $\mathcal{J}_{(d_C,d_D)}$ is finite, then there exists an $m \in \mathbb{N}$ such that $\mathcal{J}^m_{(d_C,d_D)} = \mathcal{J}_{(d_C,d_D)}$. In this case we are done. It remains to be shown that $\mathcal{J}_{(d_C,d_D)} \lesssim (\mathcal{I}_{E,\mathcal{T}}, d_E)$ also holds if $\mathcal{J}_{(d_C,d_D)}$ is an infinite tree. Consequently, there exists for each n a maximal simulation $\mathcal{S}_n \subseteq \Delta^{\mathcal{J}^n_{(d_C,d_D)}} \times \Delta^{\mathcal{I}_{E,\mathcal{T}}}$ with $((d_C,d_D),d_E) \in \mathcal{S}_n$. For the infinite sequence of subtrees

$$\mathcal{J}^0_{(d_C,d_D)}, \mathcal{J}^1_{(d_C,d_D)}, \mathcal{J}^2_{(d_C,d_D)}, ...$$

of $\mathcal{J}_{(d_C,d_D)}$ there exists an infinite sequence $\mathcal{S}_0, \mathcal{S}_1, \mathcal{S}_2, ...$ of maximal simulations. Using this sequence we construct now a simulation that shows $\mathcal{J}_{(d_C,d_D)} \lesssim (\mathcal{I}_{E,\mathcal{T}}, d_E)$. To do this we select an $\ell \in \mathbb{N}$ and an arbitrary element $\sigma \in \Delta^{\mathcal{J}_{(d_C,d_D)}}$ with $|\sigma| = \ell$.

The element σ occurs in all subtrees $\mathcal{J}^m_{(d_C,d_D)}$ with $m \geq \ell$. So there are pairs in the corresponding maximal simulations $\mathcal{S}_m$ that consist of σ and an element $d \in \Delta^{\mathcal{I}_{E,\mathcal{T}}}$. For this σ and all $m \geq \ell$ we now collect exactly those pairs that occur in the maximal simulation $\mathcal{S}_m$ and denote it by:

$$\mathcal{S}_m(\sigma) := (\{\sigma\} \times \Delta^{\mathcal{I}_{E,\mathcal{T}}}) \cap \mathcal{S}_m.$$

For all m the corresponding sets $\mathcal{S}_m(\sigma) \subseteq \mathcal{S}_m$ are non-empty.

We can also observe, that if an element σ is simulated by d in $\mathcal{S}_{i+2}$ (i.e. $(\sigma,d) \in \mathcal{S}_{i+2}(\sigma)$) it is also simulated by the same d in $\mathcal{S}_{i+1}$ since these simulations are maximal. Therefore the sets $\mathcal{S}_m(\sigma)$ don't increase with increasing m. This is shown in the following claim.

Claim. *Let $\sigma \in \Delta^{\mathcal{J}_{(d_C,d_D)}}$ with $\ell = |\sigma|$. It holds that:*

$$\mathcal{S}_\ell(\sigma) \supseteq \mathcal{S}_{\ell+1}(\sigma) \supseteq \mathcal{S}_{\ell+2}(\sigma) ...$$

Proof of the claim. We show by induction on $n \geq \ell$ that

$$\mathcal{S}_n(\sigma) \subseteq \mathcal{S}_{n-1}(\sigma) \subseteq ... \subseteq \mathcal{S}_{\ell+1}(\sigma) \subseteq \mathcal{S}_\ell(\sigma).$$

This obviously holds for the base case $n = \ell$.
Let $n > \ell$ and $(\sigma,d) \in \mathcal{S}_n(\sigma)$. It has to be shown that $(\sigma,d) \in \mathcal{S}_{n-1}(\sigma)$ and therefore $\mathcal{S}_n(\sigma) \subseteq \mathcal{S}_{n-1}(\sigma)$. Let $\mathcal{S}_n \subseteq \Delta^{\mathcal{J}^n_{(d_C,d_D)}} \times \Delta^{\mathcal{I}_{E,\mathcal{T}}}$ be the maximal simulation from $\mathcal{J}^n_{(d_C,d_D)}$ to $(\mathcal{I}_{E,\mathcal{T}}, d_E)$. Let $\mathcal{S}_{n\restriction_{n-1}}$ defined as

$$\mathcal{S}_{n\restriction_{n-1}} := \mathcal{S}_n \cap (\Delta^{\mathcal{J}^{n-1}_{(d_C,d_D)}} \times \Delta^{\mathcal{I}_{E,\mathcal{T}}})$$

be the restriction of $\mathcal{S}_n$ to pairs, whose first components are elements of the tree unraveling with depth less or equal $n-1$. Since $\mathcal{S}_{n\restriction_{n-1}}$ is a simulation from $\mathcal{J}^{n-1}_{(d_C,d_D)}$ to $(\mathcal{I}_{E,\mathcal{T}}, d_E)$, it holds that $\mathcal{S}_{n\restriction_{n-1}}$ is contained in the maximal simulation $\mathcal{S}_{n-1}$. We have now $(\sigma,d) \in \mathcal{S}_n(\sigma) \subseteq \mathcal{S}_{n\restriction_{n-1}} \subseteq \mathcal{S}_{n-1}$, because $|\sigma| < n$. Then it is implied that $(\sigma,d) \in \mathcal{S}_{n-1}(\sigma)$ and therefore $\mathcal{S}_n(\sigma) \subseteq \mathcal{S}_{n-1}(\sigma)$. By applying the induction hypothesis to $\mathcal{S}_{n-1}(\sigma)$ we get

$$\mathcal{S}_n(\sigma) \subseteq \mathcal{S}_{n-1}(\sigma) \overset{\text{I.H.}}{\subseteq} ... \overset{\text{I.H.}}{\subseteq} \mathcal{S}_{\ell+1}(\sigma) \overset{\text{I.H.}}{\subseteq} \mathcal{S}_\ell(\sigma)$$

which finishes the proof of the claim. $\square$

From this claim it follows that there exists an $f \in \mathbb{N}$ such that

$$\mathcal{S}_f(\sigma) = \bigcap_{\ell \geq |\sigma|}^{\infty} \mathcal{S}_\ell(\sigma). \tag{8}$$

We construct a relation $\mathcal{S} \subseteq \Delta^{\mathcal{J}_{(d_C,d_D)}} \times \Delta^{\mathcal{I}_{E,\mathcal{T}}}$ as follows:

$$\mathcal{S} := \bigcup_{\sigma \in \Delta^{\mathcal{J}_{(d_C,d_D)}}} \left(\bigcap_{\ell \geq |\sigma|}^{\infty} \mathcal{S}_\ell(\sigma) \right)$$

To show $\mathcal{J}_{(d_C,d_D)} \lesssim (\mathcal{I}_{E,\mathcal{T}}, d_E)$ it has to be shown that $\mathcal{S}$ is a simulation with $((d_C, d_D), d_E) \in \mathcal{S}$.

For all $n \in \mathbb{N}$ we have $((d_C, d_D), d_E) \in \mathcal{S}_n((d_C, d_D))$ and therefore $((d_C, d_D), d_E) \in \mathcal{S}$. Next we show that $\mathcal{S}$ fulfills the conditions $(S1)$ and $(S2)$ of Definition 5.

$(S1)$: Let $(\sigma, d) \in \mathcal{S}$ with $\sigma \in A^{\mathcal{J}_{(d_C,d_D)}}$ for a concept name A. It has to be shown that $d \in A^{\mathcal{I}_{E,\mathcal{T}}}$.

There exists an $x \in \mathbb{N}$ with $(\sigma, d) \in \mathcal{S}_x$. From $\sigma \in A^{\mathcal{J}_{(d_C,d_D)}}$ and (5) it follows that $\sigma \in A^{\mathcal{J}^x_{(d_C,d_D)}}$. $\mathcal{S}_x$ is a simulation from $\mathcal{J}^x_{(d_C,d_D)}$ to $(\mathcal{I}_{E,\mathcal{T}}, d_E)$ and satisfies $(S1)$. It follows that $d \in A^{\mathcal{I}_{E,\mathcal{T}}}$.

$(S2)$: Let $(\sigma, d) \in \mathcal{S}$ and $(\sigma, \sigma re) \in r^{\mathcal{J}_{(d_C,d_D)}}$. It has to be shown that there is a g with $(d, g) \in r^{\mathcal{I}_{E,\mathcal{T}}}$ and $(\sigma re, g) \in \mathcal{S}$.

By (8) there are numbers n, m with $\mathcal{S}_n(\sigma) = \bigcap_{i \geq |\sigma|}^{\infty} \mathcal{S}_i(\sigma)$ and $\mathcal{S}_m(\sigma re) = \bigcap_{j \geq |\sigma re|}^{\infty} \mathcal{S}_j(\sigma re)$. Let $m > n$ w.l.o.g. It is implied that $\mathcal{S}_m(\sigma) = \mathcal{S}_n(\sigma)$. Since $(\sigma, d) \in \mathcal{S}_m$ and $(\sigma, \sigma re) \in r^{\mathcal{J}^m_{(d_C,d_D)}}$ (by (6)), there is a g with $(d, g) \in r^{\mathcal{I}_{E,\mathcal{T}}}$ and $(\sigma re, g) \in \mathcal{S}_m(\sigma re) \subseteq \mathcal{S}_m$, because $\mathcal{S}_m$ is a simulation and satisfies $(S2)$. The number m was chosen such that $\mathcal{S}_m(\sigma re) \subseteq \mathcal{S}$ holds and therefore it is implied that $(\sigma re, g) \in \mathcal{S}$.

It is implied that $\mathcal{J}_{(d_C,d_D)} \lesssim (\mathcal{I}_{E,\mathcal{T}}, d_E)$ and therefore also $(\mathcal{I}_{C,\mathcal{T}} \times \mathcal{I}_{D,\mathcal{T}}, (d_C, d_D)) \lesssim (\mathcal{I}_{E,\mathcal{T}}, d_E)$. Together with (3) we have $(\mathcal{I}_{C,\mathcal{T}} \times \mathcal{I}_{D,\mathcal{T}}, (d_C, d_D)) \simeq (\mathcal{I}_{E,\mathcal{T}}, d_E)$.
"$\Leftarrow$":
Assume E is a common subsumer of C and D and $(\mathcal{I}_{C,\mathcal{T}} \times \mathcal{I}_{D,\mathcal{T}}, (d_C, d_D)) \simeq (\mathcal{I}_{E,\mathcal{T}}, d_E)$. It has to be shown that E is the *least* common subsumer. Let F be an arbitrary concept with $C \sqsubseteq_{\mathcal{T}} F$ and $D \sqsubseteq_{\mathcal{T}} F$. From Lemma 6.5 it follows that

$$(\mathcal{I}_{F,\mathcal{T}}, d_F) \lesssim (\mathcal{I}_{C,\mathcal{T}}, d_C)$$
$$(\mathcal{I}_{F,\mathcal{T}}, d_F) \lesssim (\mathcal{I}_{D,\mathcal{T}}, d_D)$$

From Lemma 16 it follows that

$$(\mathcal{I}_{F,\mathcal{T}}, d_F) \lesssim (\mathcal{I}_{C,\mathcal{T}} \times \mathcal{I}_{D,\mathcal{T}}, (d_C, d_D))$$

and by assumption

$$(\mathcal{I}_{F,\mathcal{T}}, d_F) \lesssim (\mathcal{I}_{C,\mathcal{T}} \times \mathcal{I}_{D,\mathcal{T}}, (d_C, d_D)) \lesssim (\mathcal{I}_{E,\mathcal{T}}, d_E).$$

We now have $(\mathcal{I}_{F,\mathcal{T}}, d_F) \lesssim (\mathcal{I}_{E,\mathcal{T}}, d_E)$ and $E \sqsubseteq_{\mathcal{T}} F$ by Lemma 6.5. So E is the least common subsumer of C and D w.r.t. $\mathcal{T}$. $\square$

By the use of this Lemma it can be verified whether a given common subsumer is the least one or not, which we illustrate by an example.

Example 18. Consider again the TBox from the introduction (now displayed with abbreviated concept names)

$$\begin{aligned}\mathcal{T}_1 = \{&\mathsf{P} \sqsubseteq \mathsf{A} \sqcap \exists \mathsf{kills}.\mathsf{S}, \quad \mathsf{S} \sqsubseteq \mathsf{B} \sqcap \exists \mathsf{resistantMutant}.\mathsf{P},\\ &\mathsf{C} \sqsubseteq \mathsf{A} \sqcap \exists \mathsf{kills}.\mathsf{E}, \quad \mathsf{E} \sqsubseteq \mathsf{B} \sqcap \exists \mathsf{resistantMutant}.\mathsf{C}\}\end{aligned}$$

and the following extended TBox

$$\mathcal{T}_2 = \mathcal{T}_1 \cup \{\mathsf{A} \sqsubseteq \exists \mathsf{kills}.\mathsf{B}, \ \ \mathsf{B} \sqsubseteq \exists \mathsf{resistantMutant}.\mathsf{A}\}.$$

In Figure 1 we can see that

$$\mathsf{A} \sqcap \exists \mathsf{kills}.(\mathsf{B} \sqcap \exists \mathsf{resistantMutant}.\mathsf{A}) \in \mathrm{cs}_{\mathcal{T}_1}(\mathsf{P}, \mathsf{C}),$$

but it is not the lcs, because its canonical model cannot simulate the product model $(\mathcal{I}_{\mathsf{P},\mathcal{T}_1} \times \mathcal{I}_{\mathsf{C},\mathcal{T}_1}, (d_{\mathsf{P}}, d_{\mathsf{C}}))$. The concept A, however, is the lcs of P and C w.r.t. $\mathcal{T}_2$. We have $(\mathcal{I}_{\mathsf{P},\mathcal{T}_2} \times \mathcal{I}_{\mathsf{C},\mathcal{T}_2}, (d_{\mathsf{P}}, d_{\mathsf{C}})) \lesssim (\mathcal{I}_{\mathsf{A},\mathcal{T}_2}, d_{\mathsf{A}})$ since any element from $\Delta^{\mathcal{I}_{\mathsf{P},\mathcal{T}_2} \times \mathcal{I}_{\mathsf{C},\mathcal{T}_2}}$ in $\mathsf{A}^{\mathcal{I}_{\mathsf{P},\mathcal{T}_2} \times \mathcal{I}_{\mathsf{C},\mathcal{T}_2}}$ or $\mathsf{B}^{\mathcal{I}_{\mathsf{P},\mathcal{T}_2} \times \mathcal{I}_{\mathsf{C},\mathcal{T}_2}}$ is simulated by A or B, respectively.

The characterization of the existence of the lcs given in Corollary 14 can be reformulated using Lemma 17.

Corollary 19. *The* $\mathrm{lcs}_{\mathcal{T}}(C, D)$ *exists iff there exists a* k *such that the canonical model of* $X^k(\mathcal{I}_{C,\mathcal{T}} \times \mathcal{I}_{D,\mathcal{T}}, (d_C, d_D))$ *w.r.t.* $\mathcal{T}$ *simulates* $(\mathcal{I}_{C,\mathcal{T}} \times \mathcal{I}_{D,\mathcal{T}}, (d_C, d_D))$.

This corollary still doesn't yield a decision procedure for the existence problem of the lcs, since the depth k is still unrestricted. Such a restriction will be developed in the next section.

3.2 A Polynomial Upper Bound on the Role-depth of the LCS

In this section we show that, if the lcs exists, its role-depth is bounded by the size of the product model. First, consider again the TBox $\mathcal{T}_2$ from Example 18, where $\mathsf{A} \sqsubseteq_{\mathcal{T}_2} \exists \mathsf{kills}.(\mathsf{B} \sqcap \exists \mathsf{resistantMutant}.\mathsf{A})$ holds, which results in a loop in the product model through the elements A, A and B, B. Furthermore, the cycles in the product model involving the roles kills and $\mathsf{resistantMutant}$ are captured by the canonical model $\mathcal{I}_{\mathsf{A},\mathcal{T}_2}$. Therefore $\mathsf{A} \equiv_{\mathcal{T}_2} \mathrm{lcs}_{\mathcal{T}_2}(\mathsf{P}, \mathsf{C})$. On this observation we build our general method.

We call elements $(d_F, d_{F'}) \in \Delta^{\mathcal{I}_{C,\mathcal{T}} \times \mathcal{I}_{D,\mathcal{T}}}$ *synchronous* if $F = F'$ and *asynchronous* otherwise. The structure of $(\mathcal{I}_{C,\mathcal{T}} \times \mathcal{I}_{D,\mathcal{T}}, (d_C, d_D))$ can now be simplified by considering only synchronous successors of synchronous elements.

Lemma 20. *Let* $(d_E, d_E) \in \Delta^{\mathcal{I}_{C,\mathcal{T}} \times \mathcal{I}_{D,\mathcal{T}}}$. $(\mathcal{I}_{C,\mathcal{T}} \times \mathcal{I}_{D,\mathcal{T}}, (d_E, d_E)) \simeq (\mathcal{I}_{E,\mathcal{T}}, d_E)$.

Proof. We define relations $\mathcal{S} \subseteq \Delta^{\mathcal{I}_{C,\mathcal{T}} \times \mathcal{I}_{D,\mathcal{T}}} \times \Delta^{\mathcal{I}_{E,\mathcal{T}}}$ and $\mathcal{Z} \subseteq \Delta^{\mathcal{I}_{E,\mathcal{T}}} \times \Delta^{\mathcal{I}_{C,\mathcal{T}} \times \mathcal{I}_{D,\mathcal{T}}}$ with $((d_E, d_E), d_E) \in \mathcal{S}$ and $(d_E, (d_E, d_E)) \in \mathcal{Z}$ as follows.

$$\begin{aligned}\mathcal{S} &:= \{((d_F, d_G), d_F) \mid (d_F, d_G) \in \Delta^{\mathcal{I}_{C,\mathcal{T}} \times \mathcal{I}_{D,\mathcal{T}}}, d_F \in \Delta^{\mathcal{I}_{E,\mathcal{T}}}\}\\ \mathcal{Z} &:= \{(d_F, (d_F, d_F)) \mid d_F \in \Delta^{\mathcal{I}_{E,\mathcal{T}}}, (d_F, d_F) \in \Delta^{\mathcal{I}_{C,\mathcal{T}} \times \mathcal{I}_{D,\mathcal{T}}}\}\end{aligned}$$

Obviously $\mathcal{S}$ and $\mathcal{Z}$ satisfy $(S1)$ and $(S2)$ of Definition 5. Since $((d_E, d_E), d_E) \in \mathcal{S}$ and $\mathcal{S}$ is a simulation, $(\mathcal{I}_{C,\mathcal{T}} \times \mathcal{I}_{D,\mathcal{T}}, (d_E, d_E)) \lesssim (\mathcal{I}_{E,\mathcal{T}}, d_E)$. And analogous we have $(d_E, (d_E, d_E)) \in \mathcal{Z}$, $\mathcal{Z}$ is a simulation and therefore $(\mathcal{I}_{E,\mathcal{T}}, d_E) \lesssim (\mathcal{I}_{C,\mathcal{T}} \times \mathcal{I}_{D,\mathcal{T}}, (d_E, d_E))$. The composition $\mathcal{S} \circ \mathcal{Z} \subseteq \Delta^{\mathcal{I}_{C,\mathcal{T}} \times \mathcal{I}_{D,\mathcal{T}}} \times \Delta^{\mathcal{I}_{C,\mathcal{T}} \times \mathcal{I}_{D,\mathcal{T}}}$ is also a simulation with $((d_E, d_E), (d_E, d_E)) \in \mathcal{S} \circ \mathcal{Z}$. The second component of the pairs in $\mathcal{S} \circ \mathcal{Z}$ are synchronous by definition of $\mathcal{Z}$. Therefore any asynchronous successor of (d_E, d_E) is simulated by its synchronous counterparts in $\mathcal{S} \circ \mathcal{Z}$. $\square$

In order to find a number k, such that the product model is simulated by the canonical model of $K = X^k(\mathcal{I}_{C,\mathcal{T}} \times \mathcal{I}_{D,\mathcal{T}}, (d_C, d_D))$, we first represent the model $(\mathcal{I}_{K,\mathcal{T}}, d_K)$ as a subtree of the tree unraveling of the product model $\mathcal{J}_{(d_C,d_D)}$ with root (d_C, d_D). We construct this representation by extending the subtree $\mathcal{J}^k_{(d_C,d_D)}$ by new tree models at depth k. We need to ensure that the resulting interpretation, denoted by $\widehat{\mathcal{J}}^k_{(d_C,d_D)}$, is a model of $\mathcal{T}$, that is simulation-equivalent to $(\mathcal{I}_{K,\mathcal{T}}, d_K)$. The elements $\sigma \in \Delta^{\mathcal{J}^k_{(d_C,d_D)}}$ with $|\sigma| = k$ we extend and the corresponding trees we append to them are selected as follows:

First we consider elements that have a tail that is a synchronous element. If $\mathsf{tail}(\sigma) = (d_F, d_F)$, then F is called *tail concept* of σ. To select the elements with a synchronous tail, that we extend by the canonical model of their tail concept, we use embeddings of $\mathcal{J}^k_{(d_C,d_D)}$ into $(\mathcal{I}_{K,\mathcal{T}}, d_K)$. We show that such an embedding exists.

Lemma 21. *Let $\mathcal{I}_{K,\mathcal{T}}$ be the canonical model of $X^k(\mathcal{I}_{C,\mathcal{T}} \times \mathcal{I}_{D,\mathcal{T}}, (d_C, d_D))$ w.r.t. $\mathcal{T}$. For any k there exists a simulation $Z \subseteq \Delta^{\mathcal{J}^k_{(d_C,d_D)}} \times \Delta^{\mathcal{I}_{K,\mathcal{T}}}$ that is functional and $Z((d_C, d_D)) = d_K$.*

Proof. It holds by Definition 9 and by definition of the tree unraveling that:

$$X^k((\mathcal{I}_{C,\mathcal{T}} \times \mathcal{I}_{D,\mathcal{T}}, (d_C, d_D))) = X^k(\mathcal{J}_{(d_C,d_D)}, (d_C, d_D)) = X^k(\mathcal{J}^k_{(d_C,d_D)}, (d_C, d_D)).$$

By Definition 3 $(\mathcal{I}_{K,\emptyset}, d_K)$ is a subinterpretation of $(\mathcal{I}_{K,\mathcal{T}}, d_K)$, which means $\Delta^{\mathcal{I}_{K,\emptyset}} \subseteq \Delta^{\mathcal{I}_{K,\mathcal{T}}}$, $A^{\mathcal{I}_{K,\emptyset}} \subseteq A^{\mathcal{I}_{K,\mathcal{T}}}$ for all concept names A and $r^{\mathcal{I}_{K,\emptyset}} \subseteq r^{\mathcal{I}_{K,\mathcal{T}}}$ for all role names r. From Definition 3 and 9 it follows that there even exists a bijective total function Z between $\Delta^{\mathcal{I}_{K,\emptyset}}$ and $\Delta^{\mathcal{J}^k_{(d_C,d_D)}}$ such that $\sigma \in A^{\mathcal{J}^k_{(d_C,d_D)}}$ iff $Z(\sigma) \in A^{\mathcal{I}_{K,\emptyset}}$ for all A and $(\sigma, \sigma') \in r^{\mathcal{J}^k_{(d_C,d_D)}}$ iff $(Z(\sigma), Z(\sigma')) \in r^{\mathcal{I}_{K,\emptyset}}$ for all r. Z is a functional simulation from $(\mathcal{J}^k_{(d_C,d_D)}, (d_C, d_D))$ to $(\mathcal{I}_{K,\mathcal{T}}, d_K)$. □

Let $\mathcal{H} = \{Z_1, ..., Z_n\}$ be the set of all functional simulations Z_i from $\mathcal{J}^k_{(d_C,d_D)}$ to $(\mathcal{I}_{K,\mathcal{T}}, d_K)$ with $Z_i((d_C, d_D)) = d_K$. We say that σ with tail concept F is *matched* by Z_i if $Z_i(\sigma) \in F^{\mathcal{I}_{K,\mathcal{T}}}$. The set of elements $\sigma \in \Delta^{\mathcal{J}^k_{(d_C,d_D)}}$ with $|\sigma| = k$, that are matched by a functional simulation Z_i is called *matching set* denoted by $\mathcal{M}(Z_i)$.

The elements from $\Delta^{\mathcal{J}^k_{(d_C,d_D)}}$, we extend, are called *stubs*.

Definition 22. Let $\sigma \in \Delta^{\mathcal{J}^k_{(d_C,d_D)}}$ with $|\sigma| = k$. σ is contained in the *set of stubs* of $\mathcal{J}^k_{(d_C,d_D)}$, denoted by $\mathsf{stubs}(\mathcal{J}^k_{(d_C,d_D)})$, if it satisfies one of the following properties:

1. Let M be a conjunction of concept names and $\exists r.F \in \mathsf{sub}(\mathcal{T})$. It holds that $\sigma \in M^{\mathcal{J}^k_{(d_C,d_D)}}$ and $M \sqsubseteq_{\mathcal{T}} \exists r.F$
2. Let $\mathcal{M}(\mathcal{H}) := \{\mathcal{M}(Z) \mid Z \in \mathcal{H}\}$ be the set of all matching sets. It holds that σ is contained in all maximal sets in $\mathcal{M}(\mathcal{H})$.

Now we define the set of trees $\Upsilon(\sigma)$ that are appended to a stub σ. Consider $\sigma \in \mathsf{stubs}(\mathcal{J}^k_{(d_C,d_D)})$ that satisfies the first condition for $\exists r.F$. Let $(\mathcal{I}_{\exists r.F,\mathcal{T}}, d_{\exists r.F})$ be the canonical model. By definition of $\mathcal{J}_{(d_C,d_D)}$ it holds that $\sigma r(d_F, d_F) \in \Delta^{\mathcal{J}_{(d_C,d_D)}}$ and the subtree $\mathcal{J}_{\sigma r(d_F,d_F)}$ of $\mathcal{J}_{(d_C,d_D)}$ is simulation-equivalent to $(\mathcal{I}_{\exists r.F,\mathcal{T}}, d_{\exists r.F})$ (by Lemma 20). Thus $\Upsilon(\sigma)$ contains $\mathcal{J}_{\sigma r(d_F,d_F)}$.

Assume $\sigma \in \mathsf{stubs}(\mathcal{J}^k_{(d_C,d_D)})$ satisfies the second property for the tail concept F. In this case the subtree $\mathcal{J}_\sigma$ of $\mathcal{J}_{(d_C,d_D)}$ is simulation-equivalent to $(\mathcal{I}_{F,\mathcal{T}}, d_F)$ as shown in Lemma 20. Thus $\Upsilon(\sigma)$ contains $\mathcal{J}_\sigma$.

We define the extended interpretation $\widehat{\mathcal{J}}^k_{(d_C,d_D)}$ as follows:

$$\begin{aligned}
\Delta^{\widehat{\mathcal{J}}^k_{(d_C,d_D)}} &:= \Delta^{\mathcal{J}^k_{(d_C,d_D)}} \cup \bigcup_{\sigma \in \mathsf{stubs}(\Delta^{\mathcal{J}^k_{(d_C,d_D)}})} \bigcup_{\mathcal{J} \in \Upsilon(\sigma)} \Delta^{\mathcal{J}} \\
A^{\widehat{\mathcal{J}}^k_{(d_C,d_D)}} &:= A^{\mathcal{J}^k_{(d_C,d_D)}} \cup \bigcup_{\sigma \in \mathsf{stubs}(\Delta^{\mathcal{J}^k_{(d_C,d_D)}})} \bigcup_{\mathcal{J} \in \Upsilon(\sigma)} A^{\mathcal{J}}, \text{ for all } A \in N_C \\
r^{\widehat{\mathcal{J}}^k_{(d_C,d_D)}} &:= r^{\mathcal{J}^k_{(d_C,d_D)}} \cup \bigcup_{\sigma \in \mathsf{stubs}(\Delta^{\mathcal{J}^k_{(d_C,d_D)}})} \bigcup_{\mathcal{J} \in \Upsilon(\sigma)} r^{\mathcal{J}}, \text{ for all } r \in N_R
\end{aligned} \tag{9}$$

We show that the resulting interpretation $\widehat{\mathcal{J}}^k_{(d_C,d_D)}$ has the desired properties.

Lemma 23. *Let E be a concept and $K = X^k(\mathcal{I}_{C,\mathcal{T}} \times \mathcal{I}_{D,\mathcal{T}}, (d_C, d_D))$. There is a functional simulation Z from $\mathcal{J}^k_{(d_C,d_D)}$ to $(\mathcal{I}_{K,\mathcal{T}}, d_K)$ with $Z((d_C, d_D)) = d_K$ such that for all $\sigma \in \Delta^{\mathcal{J}^k_{(d_C,d_D)}} \cap \Delta^{\widehat{\mathcal{J}}^k_{(d_C,d_D)}}$ it holds that $\sigma \in E^{\widehat{\mathcal{J}}^k_{(d_C,d_D)}}$ implies $Z(\sigma) \in E^{\mathcal{I}_{K,\mathcal{T}}}$.*

Proof. There exists a functional simulation Z satisfying the following properties:

- $Z \subseteq \Delta^{\mathcal{J}^k_{(d_C,d_D)}} \times \Delta^{\mathcal{I}_{K,\mathcal{T}}}$ with $Z((d_C, d_D)) = d_K$.
- Let $\sigma \in \mathsf{stubs}(\mathcal{J}^k_{(d_C,d_D)})$, such that σ satisfies property 1 of Definition 22 w.r.t. $\exists r.F$. Since $Z(\sigma)$ simulates σ it is implied (by $(S1)$ of Definition 5) that $Z(\sigma) \in (\exists r.F)^{\mathcal{I}_{K,\mathcal{T}}}$. Therefore $Z(\sigma)$ simulates the tree $\mathcal{J}_{\sigma r(d_F,d_F)} \in \Upsilon(\sigma)$.
- All stubs in $\mathsf{stubs}(\mathcal{J}^k_{(d_C,d_D)})$ satisfying the second property of Definition 22 are matched by Z.

Since all tree unravelings appended to a stub σ are simulated by $Z(\sigma)$, there exists a simulation $\mathcal{S} \subseteq \Delta^{\widehat{\mathcal{J}}^k_{(d_C,d_D)}} \times \Delta^{\mathcal{I}_{K,\mathcal{T}}}$ that is an extension of Z such that $((d_C, d_D), d_K) \in \mathcal{S}$ and $\mathcal{S}$ is functional on $\Delta^{\mathcal{J}^k_{(d_C,d_D)}} \cap \Delta^{\widehat{\mathcal{J}}^k_{(d_C,d_D)}}$.

Now we can show the claim as follows:

Let $\sigma \in \Delta^{\mathcal{J}^k_{(d_C,d_D)}} \cap \Delta^{\widehat{\mathcal{J}}^k_{(d_C,d_D)}}$ with $\sigma \in E^{\widehat{\mathcal{J}}^k_{(d_C,d_D)}}$. From Lemma 6.4 it follows that

$$(\mathcal{I}_{E,\emptyset}, d_E) \lesssim (\widehat{\mathcal{J}}^k_{(d_C,d_D)}, \sigma) \lesssim (\mathcal{I}_{K,\mathcal{T}}, Z(\sigma)).$$

Again by Lemma 6.4 it holds that $Z(\sigma) \in E^{\mathcal{I}_{K,\mathcal{T}}}$. □

Lemma 24. *Let $K = X^k(\mathcal{I}_{C,\mathcal{T}} \times \mathcal{I}_{D,\mathcal{T}}, (d_C, d_D))$. The interpretation $\widehat{\mathcal{J}}^k_{(d_C,d_D)}$ has the following properties:*

1. *$\widehat{\mathcal{J}}^k_{(d_C,d_D)}$ is a model of $\mathcal{T}$;*
2. *$\widehat{\mathcal{J}}^k_{(d_C,d_D)} \simeq (\mathcal{I}_{K,\mathcal{T}}, d_K)$.*

Proof of point 1. Let $F \sqsubseteq G \in \mathcal{T}$. It has to be shown that $F^{\widehat{\mathcal{J}}^k_{(d_C,d_D)}} \subseteq G^{\widehat{\mathcal{J}}^k_{(d_C,d_D)}}$. Assume that $F \sqsubseteq G$ is of the form

$$A_1 \sqcap \cdots \sqcap A_n \sqcap \exists r_1.E_1 \sqcap \cdots \sqcap \exists r_m.E_m \sqsubseteq B_1 \sqcap \cdots \sqcap B_{n'} \sqcap \exists s_1.H_1 \sqcap \cdots \sqcap \exists s_{m'}.H_{m'}$$

For any $\sigma \in \Delta^{\widehat{\mathcal{J}}^k_{(d_C,d_D)}} \subseteq \Delta^{\mathcal{J}_{(d_C,d_D)}}$ it holds

$$\sigma \in A^{\widehat{\mathcal{J}}^k_{(d_C,d_D)}} \text{ iff } \sigma \in A^{\mathcal{J}_{(d_C,d_D)}}, \text{ for all } A \in N_C. \tag{10}$$

Let $\sigma \in F^{\widehat{\mathcal{J}}^k_{(d_C,d_D)}}$. It has to be shown that $\sigma \in G^{\widehat{\mathcal{J}}^k_{(d_C,d_D)}}$. To do this we distinguish the following cases for $\sigma \in \Delta^{\widehat{\mathcal{J}}^k_{(d_C,d_D)}}$:

1. $|\sigma| < k$;
2. $|\sigma| = k$ and $\sigma \in \mathsf{stubs}(\mathcal{J}^k_{(d_C,d_D)})$;
3. $|\sigma| = k$ and $\sigma \notin \mathsf{stubs}(\mathcal{J}^k_{(d_C,d_D)})$;
4. $|\sigma| > k$.

1. Case: $\sigma \in \Delta^{\mathcal{J}^k_{(d_C,d_D)}}$ with $|\sigma| < k$. From $\sigma \in F^{\widehat{\mathcal{J}}^k_{(d_C,d_D)}}$ it follows $\sigma \in F^{\mathcal{J}_{(d_C,d_D)}}$. Since $\mathcal{J}_{(d_C,d_D)}$ is a model of $\mathcal{T}$, it is implied that $\sigma \in G^{\mathcal{J}_{(d_C,d_D)}}$. We want to show that $\sigma \in G^{\widehat{\mathcal{J}}^k_{(d_C,d_D)}}$ also holds. We first consider the concept names $B_i (i = 1, ..., n')$ on top-level of G. Since $\mathcal{J}_{(d_C,d_D)}$ is a model of $\mathcal{T}$, it holds $\sigma \in (\bigsqcap_{i=1,\dots,n'} B_i)^{\mathcal{J}_{(d_C,d_D)}}$ and therefore $\sigma \in (\bigsqcap_{i=1,\dots,n'} B_i)^{\widehat{\mathcal{J}}^k_{(d_C,d_D)}}$ (by (10)).

   Now we consider a top-level conjunct $\exists s.H$ of G. Let $\mathsf{tail}(\sigma) = (d_L, d_{L'})$. Since $\sigma \in (\exists s.H)^{\mathcal{J}_{(d_C,d_D)}}$ it is implied that $L \sqsubseteq_{\mathcal{T}} \exists s.H$ and $L' \sqsubseteq_{\mathcal{T}} \exists s.H$. Therefore the element $\sigma s(d_H, d_H) \in \Delta^{\mathcal{J}_{(d_C,d_D)}}$ and with $\sigma < k$, we have also $\sigma s(d_H, d_H) \in \Delta^{\mathcal{J}^k_{(d_C,d_D)}}$. Next we show that the elements on level k, that are reachable from $\sigma s(d_H, d_H)$ in $\mathcal{J}^k_{(d_C,d_D)}$ are stubs according to the second property in Definition 22. Let $\mathcal{F}$ be an arbitrary functional simulation from $\mathcal{J}^k_{(d_C,d_D)}$ to $(\mathcal{I}_{K,\mathcal{T}}, d_K)$. Assume that there exists at least one element δ in $\mathcal{J}^k_{(d_C,d_D)}$ at depth k reachable from $\sigma s(d_H, d_H)$ such that δ is not matched by $\mathcal{F}$, i.e. $\delta \notin \mathcal{M}(\mathcal{F})$. Otherwise, if no such δ and $\mathcal{F}$ exist, then it is implied that all successors of $\sigma s(d_H, d_H)$ at level k are stubs, and we get $\sigma \in (\exists s.H)^{\widehat{\mathcal{J}}^k_{(d_C,d_D)}}$ directly.

   There exists a functional simulation $\mathcal{Z}$ with $\mathcal{M}(\mathcal{F}) \subseteq \mathcal{M}(\mathcal{Z})$ and $\mathcal{M}(\mathcal{Z})$ is maximal. By Definition 22 it is implied that, all stubs satisfying the second property of Definition 22 are matched by $\mathcal{Z}$. Therefore $\mathcal{Z}$ satisfies the properties, that are required in Lemma 23. Assume that $\delta \notin \mathcal{M}(\mathcal{Z})$. We show that this assumption leads to a contradiction. Since $\sigma \in F^{\widehat{\mathcal{J}}^k_{(d_C,d_D)}}$ it is implied by Lemma 23 that $\mathcal{Z}(\sigma) \in F^{\mathcal{I}_{K,\mathcal{T}}}$. Because $F \sqsubseteq G \in \mathcal{T}$ and $\exists s.H$ is a conjunct in G, we get $\mathcal{Z}(\sigma) \in (\exists s.H)^{\mathcal{I}_{K,\mathcal{T}}}$, i.e. d_H is an s-successor of $\mathcal{Z}(\sigma)$ in $\mathcal{I}_{K,\mathcal{T}}$ by Definition of the canonical models. There exists a function $\widehat{\mathcal{Z}}$ that maps $\sigma s(d_H, d_H)$ to d_H and all successors of $\sigma s(d_H, d_H)$ in $\mathcal{J}^k_{(d_C,d_D)}$ to an element d_Y such that Y is the tail concept of the successor of $\sigma s(d_H, d_H)$. For all other elements in $\mathcal{J}^k_{(d_C,d_D)}$, $\widehat{\mathcal{Z}}$ coincides with $\mathcal{Z}$. By construction of $\widehat{\mathcal{Z}}$ it is implied that $\delta \in \mathcal{M}(\widehat{\mathcal{Z}})$ and $\mathcal{M}(\mathcal{Z}) \subseteq \mathcal{M}(\widehat{\mathcal{Z}})$. Since $\mathcal{M}(\mathcal{Z})$ is maximal, it is implied that $\mathcal{M}(\mathcal{Z}) = \mathcal{M}(\widehat{\mathcal{Z}})$, which contradicts the assumption $\delta \notin \mathcal{M}(\mathcal{Z})$. Therefore $\sigma \in (\exists s.H)^{\widehat{\mathcal{J}}^k_{(d_C,d_D)}}$ and $\sigma \in G^{\widehat{\mathcal{J}}^k_{(d_C,d_D)}}$.

2. Case: $|\sigma| = k$ and $\sigma \in \mathsf{stubs}(\mathcal{J}^k_{(d_C,d_D)})$.

   (a) Assume that σ satisfies the first property of Definition 22 but not the second one. Let $M := \bigsqcap_{\sigma \in A^{\widehat{\mathcal{J}}^k_{(d_C,d_D)}}} A$. By assumption we have $\sigma \in F^{\widehat{\mathcal{J}}^k_{(d_C,d_D)}}$ and since (10) holds the names $A_1, ..., A_n$ on top-level of F and $B_1, ..., B_{n'}$ on top-level of G are contained in M, because $\mathcal{J}_{(d_C,d_D)}$ is a model of $\mathcal{T}$. From $\sigma \in F^{\widehat{\mathcal{J}}^k_{(d_C,d_D)}}$ it follows that $\sigma \in (\exists r_i.E_i)^{\widehat{\mathcal{J}}^k_{(d_C,d_D)}}$ for all $1 \le i \le m$. By definition of $\widehat{\mathcal{J}}^k_{(d_C,d_D)}$ there exists for all i

$\exists r_i.E'_i \in \mathsf{sub}(\mathcal{T})$ with $M \sqsubseteq_{\mathcal{T}} \exists r_i.E'_i$ and $E'_i \sqsubseteq_{\mathcal{T}} E_i$. It is implied that $M \sqsubseteq_{\mathcal{T}} \exists r_i.E_i$ for all $i = 1, ..., m$ and therefore $M \sqsubseteq_{\mathcal{T}} F$. Since $F \sqsubseteq_{\mathcal{T}} G$ we have also $M \sqsubseteq_{\mathcal{T}} \exists s_j.H_j$ for all $j = 1, ..., m'$. By construction of $\widehat{\mathcal{J}}^k_{(d_C,d_D)}$ it holds that the tree model $\mathcal{J}_{\sigma s_j d_{H_j}}$ was added to $\mathcal{J}^k_{(d_C,d_D)}$ such that $\sigma \in (\exists s_j.H_j)^{\widehat{\mathcal{J}}^k_{(d_C,d_D)}}$. Finally we get $\sigma \in G^{\widehat{\mathcal{J}}^k_{(d_C,d_D)}}$.

(b) Assume that σ satisfies the second property of Definition 22. In this case $\mathcal{J}_\sigma \in \Upsilon(\sigma)$ is a model of $\mathcal{T}$ and a subtree of $\widehat{\mathcal{J}}^k_{(d_C,d_D)}$.

3. Case: $|\sigma| = k$ and $\sigma \notin \mathsf{stubs}(\mathcal{J}^k_{(d_C,d_D)})$. The element σ has no successors in $\widehat{\mathcal{J}}^k_{(d_C,d_D)}$. Therefore $rd(F) = 0$, i.e. F is a conjunction of concept names. Since $\sigma \notin \mathsf{stubs}(\mathcal{J}^k_{(d_C,d_D)})$, G is also a conjunction of concept names and the claim follows directly from (10).

4. In the remaining case, σ is part of a tree model, that was added to $\mathcal{J}^k_{(d_C,d_D)}$ during the construction of $\widehat{\mathcal{J}}^k_{(d_C,d_D)}$. Since these trees are models of $\mathcal{T}$, the GCIs are satisfied in this tree model.

□

Proof of point 2. This is a direct consequence of Lemma 23 and point 1. □

Having this representation of the canonical model of the k-lcs$_{\mathcal{T}}(C, D)$ we first show another sufficient condition for the existence of the lcs.

Corollary 25. *If $(\mathcal{I}_{C,\mathcal{T}} \times \mathcal{I}_{D,\mathcal{T}}, (d_C, d_D))$ only has cycles reachable from (d_C, d_D) consisting of synchronous elements, then the lcs$_{\mathcal{T}}(C, D)$ exists.*

Proof. Consider the tree unraveling $\mathcal{J}_{(d_C,d_D)}$. There exists a number n such that all paths in $\mathcal{J}_{(d_C,d_D)}$ have maximal a finite prefix of asynchronous elements of length $\leq n-1$ and has from position n on only synchronous elements. Now consider the interpretation $\mathcal{J}^n_{(d_C,d_D)}$ and an element σ on level n with tail concept E. Assume E has role-depth m. Now we unravel $\mathcal{J}^n_{(d_C,d_D)}$ further up to depth m such that we get $\mathcal{J}^{n+m}_{(d_C,d_D)}$. It is implied that the corresponding model $\widehat{\mathcal{J}}^{n+m}_{(d_C,d_D)}$ contains all paths from $\mathcal{J}_{(d_C,d_D)}$ that have the prefix $(d_C, d_D) \cdots \sigma$. Therefore with $n := |\Delta^{\mathcal{I}_{C,\mathcal{T}} \times \mathcal{I}_{D,\mathcal{T}}}|$ and $u := \mathsf{max}(\{rd(F) \mid F \in \mathsf{sub}(\mathcal{T}) \cup \{C, D\}\})$ we get $\widehat{\mathcal{J}}^{n+u}_{(d_C,d_D)} = \mathcal{J}_{(d_C,d_D)}$. From Lemma 24 and Corollary 19 it follows that $X^{n+u}(\mathcal{I}_{C,\mathcal{T}} \times \mathcal{I}_{D,\mathcal{T}}, (d_C, d_D))$ is the lcs. □

As seen in Example 18 for $\mathcal{T}_2$, this is not a necessary condition for the existence of the lcs.

Another consequence of Lemma 24 is, that if the product model $(\mathcal{I}_{C,\mathcal{T}} \times \mathcal{I}_{D,\mathcal{T}}, (d_C, d_D))$ has only asynchronous cycles reachable from (d_C, d_D), then the lcs$_{\mathcal{T}}(C, D)$ does not exist. Since in this case $\mathcal{J}_{(d_C,d_D)}$ is infinite but $\widehat{\mathcal{J}}^k_{(d_C,d_D)}$ is finite for all $k \in \mathbb{N}$, a simulation from $(\mathcal{I}_{C,\mathcal{T}} \times \mathcal{I}_{D,\mathcal{T}}, (d_C, d_D))$ to $\widehat{\mathcal{J}}^k_{(d_C,d_D)}$ never exists for all k. For instance, this case applies to Example 18 w.r.t. to $\mathcal{T}_1$.

The interesting case is where we have both asynchronous and synchronous cycles reachable from (d_C, d_D) in the product model. In this case we choose a k that is large enough and then check whether the canonical model of $X^k(\mathcal{I}_{C,\mathcal{T}} \times \mathcal{I}_{D,\mathcal{T}}, (d_C, d_D))$ w.r.t. $\mathcal{T}$ simulates the product model.

We show in the next Lemma that the role-depth of the lcs$_{\mathcal{T}}(C, D)$, if it exists, can be bounded by a polynomial, that is quadratic in the size of the product model.

$$
\begin{tikzcd}
p = & d_0 \arrow[r, "r_1"] \arrow[d, "\mathcal{S}"] & d_1 \arrow[r, "r_2"] \arrow[d, "\mathcal{S}"] & d_2 \arrow[r, "r_3"] \arrow[d, "\mathcal{S}"] & d_3 \arrow[r, "r_4"] \arrow[d, "\mathcal{S}"] & \cdots \\
p_\ell = & \sigma_0 \arrow[r, "r_1"] & \sigma_1 \arrow[r, "r_2"] & \sigma_2 \arrow[r, "r_3"] & \sigma_3 \arrow[r, "r_4"] & \cdots
\end{tikzcd}
$$

Figure 2: simulation chain of p and p_ℓ

Lemma 26. *Let*

$$
\begin{aligned}
&n := |\Delta^{\mathcal{I}_{C,\mathcal{T}} \times \mathcal{I}_{D,\mathcal{T}}}| \textit{ and}\\
&m := \mathit{max}(\{rd(F) \mid F \in \mathit{sub}(\mathcal{T}) \cup \{C, D\}\}) \textit{ and}\\
&k = n^2 + m + 1.
\end{aligned}
$$

If $lcs_{\mathcal{T}}(C, D)$ *exists then* $(\mathcal{I}_{C,\mathcal{T}} \times \mathcal{I}_{D,\mathcal{T}}, (d_C, d_D)) \lesssim \widehat{\mathcal{J}}^k_{(d_C,d_D)}$.

We outline the proof of this claim as follows: Assume $\mathrm{lcs}_{\mathcal{T}}(C, D)$ exists. From Corollary 19 and Lemma 24 it follows that there exists a number ℓ such that

$$(\mathcal{I}_{C,\mathcal{T}} \times \mathcal{I}_{D,\mathcal{T}}, (d_C, d_D)) \lesssim \widehat{\mathcal{J}}^\ell_{(d_C,d_D)}. \tag{11}$$

Every path in $\widehat{\mathcal{J}}^\ell_{(d_C,d_D)}$ has a maximal asynchronous prefix of length $\leq \ell$. From depth $\ell + 1$ on there are only synchronous elements in the tree $\widehat{\mathcal{J}}^\ell_{(d_C,d_D)}$. From (11) it follows that every path p in $(\mathcal{I}_{C,\mathcal{T}} \times \mathcal{I}_{D,\mathcal{T}}, (d_C, d_D))$ starting in (d_C, d_D), is simulated by a corresponding path p_ℓ in $\widehat{\mathcal{J}}^\ell_{(d_C,d_D)}$ starting in (d_C, d_D). The *simulation chain* of p and p_ℓ is depicted in Figure 2. The idea is to use the simulating path p_ℓ to construct a simulating path in $\widehat{\mathcal{J}}^\ell_{(d_C,d_D)}$ (also starting in (d_C, d_D)) with a maximal asynchronous prefix of length $\leq n^2$. n^2 is the number of pairs of elements from $\Delta^{\mathcal{I}_{C,\mathcal{T}} \times \mathcal{I}_{D,\mathcal{T}}}$. Intuitively, if p_ℓ has a maximal asynchronous prefix that is longer than n^2, then there are pairs in the simulation chain that occur more than once. This is used to construct step wise a simulating path with a shorter maximal asynchronous prefix such that all pairs consisting of asynchronous elements in the simulation chain are pairwise distinct. Therefore we need only asynchronous elements from $\widehat{\mathcal{J}}^\ell_{(d_C,d_D)}$ up to depth n^2 to simulate the product model. Then $m + 1$ was added to n^2 to ensure that $\widehat{\mathcal{J}}^{n^2+m+1}_{(d_C,d_D)}$ contains *all* paths from $\mathcal{J}_{(d_C,d_D)}$ starting in (d_C, d_D), that have a maximal asynchronous prefix of length $\leq n^2$. As argued above $\widehat{\mathcal{J}}^{n^2+m+1}_{(d_C,d_D)}$ simulates $(\mathcal{I}_{C,\mathcal{T}} \times \mathcal{I}_{D,\mathcal{T}}, (d_C, d_D))$.

Proof of Lemma 26. Assume the lcs E of C and D w.r.t. $\mathcal{T}$ exists with $rd(E) = \ell$. It is implied by Lemma 24 that $(\mathcal{I}_{E,\mathcal{T}}, d_E) \simeq \widehat{\mathcal{J}}^\ell_{(d_C,d_D)}$ and $(\mathcal{I}_{K,\mathcal{T}}, d_K) \simeq \widehat{\mathcal{J}}^k_{(d_C,d_D)}$ for $K = k\text{-lcs}_{\mathcal{T}}(C, D)$. For $\ell \leq k$ the claim follows directly. We consider the case $\ell > k$.

By assumption it holds that $(\mathcal{I}_{C,\mathcal{T}} \times \mathcal{I}_{D,\mathcal{T}}, (d_C, d_D)) \lesssim \widehat{\mathcal{J}}^\ell_{(d_C,d_D)}$. Let $\mathcal{S} \subseteq \Delta^{\mathcal{I}_{C,\mathcal{T}} \times \mathcal{I}_{D,\mathcal{T}}} \times \Delta^{\widehat{\mathcal{J}}^\ell_{(d_C,d_D)}}$ be the maximal simulation with $((d_C, d_D), (d_C, d_D)) \in \mathcal{S}$.

Consider a path p with elements d_i, $i = 0, 1, 2, ...$ in $(\mathcal{I}_{C,\mathcal{T}} \times \mathcal{I}_{D,\mathcal{T}}, (d_C, d_D))$ starting in $d_0 = (d_C, d_D)$ with an asynchronous prefix of length $> n^2$. There exists a path p_ℓ in $\widehat{\mathcal{J}}^\ell_{(d_C,d_D)}$ with elements σ_i, $i = 0, 1, 2, ...$ such that $\sigma_0 = (d_C, d_D)$ and p is simulated by p_ℓ in $\mathcal{S}$, which means that $(d_i, d_{i+1}) \in r_{i+1}^{\mathcal{I}_{C,\mathcal{T}} \times \mathcal{I}_{D,\mathcal{T}}}$, $(\sigma_i, \sigma_{i+1}) \in r_{i+1}^{\widehat{\mathcal{J}}^\ell_{(d_C,d_D)}}$ and $(d_i, \sigma_i) \in \mathcal{S}$ for all $i = 0, 1, 2, ...$ p_ℓ has a maximal asynchronous prefix of length $\leq \ell$ and we assume that this prefix has a length $> n^2$. p and p_ℓ form a simulation chain as depicted in Figure 2.

Now we construct step wise a path in $\widehat{\mathcal{J}}^\ell_{(d_C,d_D)}$ that simulates the path p in $\mathcal{S}$ and has a maximal asynchronous prefix of length $\leq n^2$. Let $p^0_\ell = p_\ell$.

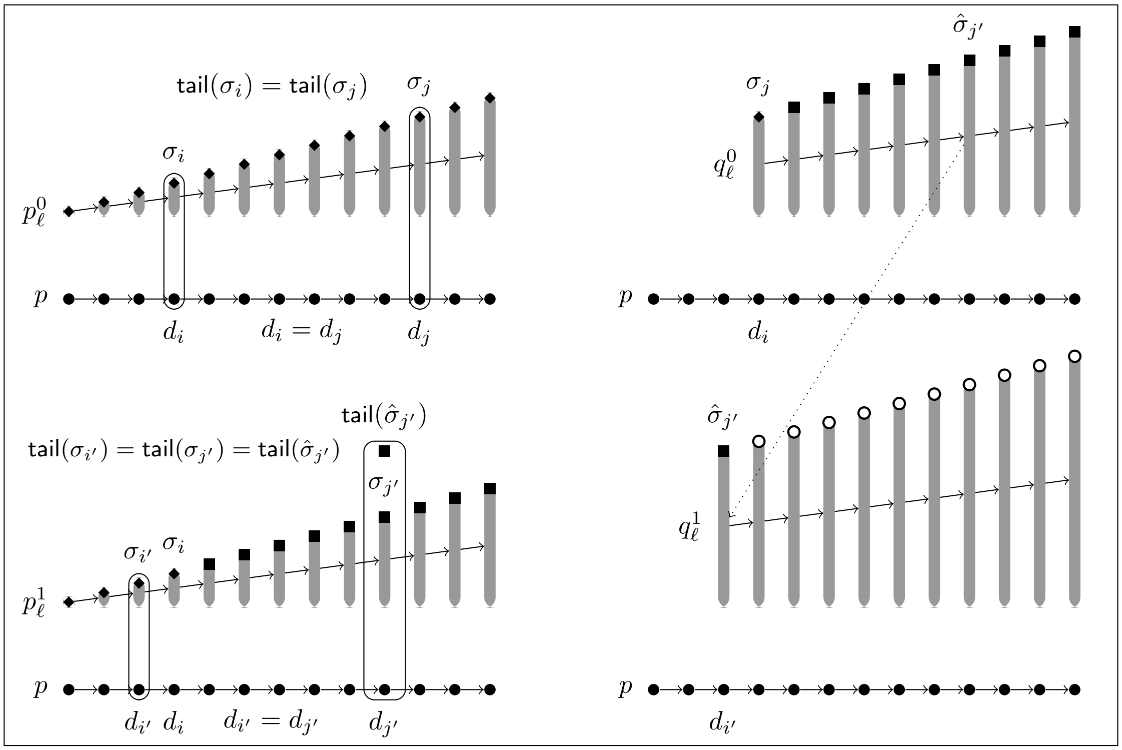

Figure 3: Visualization of the proof idea of Lemma 26

Since p and p_ℓ have a maximal asynchronous prefix that is longer than n^2, there are indices $i < j$ such that $(d_i, \mathsf{tail}(\sigma_i)) = (d_j, \mathsf{tail}(\sigma_j))$ and the asynchronous pairs

$$(d_0, \mathsf{tail}(\sigma_0)), (d_1, \mathsf{tail}(\sigma_1)), \cdots, (d_i, \mathsf{tail}(\sigma_i)) \tag{12}$$

are pairwise distinct. This is depicted in the first diagram in Figure 3 where the tails of the elements of p_ℓ have a diamond shape. Since $d_i = d_j$ and d_j is simulated by σ_j in $\mathcal{S}$, the path fragment $p[d_i \cdots]$ of p starting in d_i is also simulated by a path $q_\ell^0 = \sigma_j r_{j+1} \hat{\sigma}_{j+1} r_{j+2} \cdots$ in $\widehat{\mathcal{J}}^\ell_{(d_C, d_D)}$ starting in σ_j. This is depicted in the second diagram in Figure 3. Note that the successors of σ_j in q_ℓ can have different tails than the successors of σ_j in p_ℓ. This is depicted by the tails with square shape. Since σ_i and σ_j with $|\sigma_i| < |\sigma_j|$ are copies of the same element (which means they have the same tail) of the product model, there is also a path in $\widehat{\mathcal{J}}^\ell_{(d_C, d_D)}$ starting in σ_i such that the tails of the successors of σ_i in this path are equal to the tails of the successors of σ_j in q_ℓ. Therefore the simulating path p_ℓ^0 can be modified to a simulating path p_ℓ^1 in $\widehat{\mathcal{J}}^\ell_{(d_C, d_D)}$ as sketched in the third diagram of Figure 3. p_ℓ^1 is the result of the first construction step. p_ℓ^1 has the following form

$$p_\ell^1 = \sigma_0 r_1 \sigma_1 r_2 ... r_i \sigma_i r_{i+1} \sigma'_{i+1} r_{i+2} \sigma'_{i+2} r_{i+3} \cdots$$

If the asynchronous part of the corresponding simulation chain p_ℓ^1 and p consists only of pairwise distinct pairs (w.r.t. to the tails), then we are finished. Assume that this does not hold for p_ℓ^1. Then as for p_ℓ there are indices $i' < j'$ with $(d_{i'}, \mathsf{tail}(\sigma_{i'})) = (d_{j'}, \mathsf{tail}(\sigma_{j'}))$. Because (12) holds it is implied that $j' > i$. This means w.r.t. the sketch in figure 3, that $\mathsf{tail}(\sigma_{j'})$ has a square shape. It is implied that $d_{i'}$ is also simulated by a successor $\hat{\sigma}_{j'}$ of σ_j in q_ℓ. Analogous to the first step, the path fragment $p[d_{i'} \cdots]$ of p is simulated by a path q_ℓ^1 in $\widehat{\mathcal{J}}^\ell_{(d_C, d_D)}$ starting with $\hat{\sigma}_{j'}$.

Since it holds that $|\sigma_j| < |\hat{\sigma}_{j'}|$, it is implied that only a finite number of this replacement steps can be executed. By construction as sketched above we finally get a path p_{fin} starting with

(d_C, d_D) in $\widehat{\mathcal{J}}^{\ell}_{(d_C,d_D)}$, that has a maximal asynchronous prefix of length at most n^2. We show that p_{fin} is also contained in $\widehat{\mathcal{J}}^{k}_{(d_C,d_D)}$. Consider the synchronous element δ at position n^2+1 of p_{fin}. Let F be the tail concept of δ. It is then implied by Lemma 20 that the path fragment of p_{fin} starting at position n^2+1 belongs to the canonical model of F. It holds that $rd(F) \leq m$. Since $k = n^2 + m + 1$, the prefix of length k of p_{fin} is contained in $\mathcal{J}^{k}_{(d_C,d_D)}$. It is implied that $\delta \in F^{\mathcal{J}^{k}_{(d_C,d_D)}}$ and therefore also $\delta \in F^{\widehat{\mathcal{J}}^{k}_{(d_C,d_D)}}$. Since $\widehat{\mathcal{J}}^{k}_{(d_C,d_D)}$ is a model of $\mathcal{T}$, it follows from Lemma 6.4 that $(\mathcal{I}_{F,\mathcal{T}}, d_F) \lesssim (\widehat{\mathcal{J}}^{k}_{(d_C,d_D)}, \delta)$. The other direction also holds since we have added to subtree of $\mathcal{J}^{k}_{(d_C,d_D)}$ with root δ only synchronous elements that belong to the tree unraveling of $(\mathcal{I}_{F,\mathcal{T}}, d_F)$. Therefore p_{fin} is also contained in $\widehat{\mathcal{J}}^{k}_{(d_C,d_D)}$. It is implied that there exists a simulation such that $(\mathcal{I}_{C,\mathcal{T}} \times \mathcal{I}_{D,\mathcal{T}}, (d_C, d_D)) \lesssim \widehat{\mathcal{J}}^{k}_{(d_C,d_D)}$. □

Using Lemma 17 and Lemma 26 we can now show the main result of this section.

Theorem 27. *Let C, D be concepts and $\mathcal{T}$ a general TBox. It is decidable in polynomial time whether the $lcs_{\mathcal{T}}(C, D)$ exists. If the $lcs_{\mathcal{T}}(C, D)$ exists it can be computed in polynomial time.*

Proof. First we compute the bound k as given in Lemma 26 and then the k-characteristic concept K of $(\mathcal{I}_{C,\mathcal{T}} \times \mathcal{I}_{D,\mathcal{T}}, (d_C, d_D))$. The canonical model of K is build according to Definition 3 in polynomial time. Next we check whether $(\mathcal{I}_{C,\mathcal{T}} \times \mathcal{I}_{D,\mathcal{T}}, (d_C, d_D)) \lesssim (\mathcal{I}_{K,\mathcal{T}}, d_K)$ holds, which can be done in polynomial time. If yes K is the lcs by Lemma 17 and if not the lcs doesn't exist by Lemma 26. □

The results from this section can be easily generalized to the lcs of an arbitrary set of concepts $M = \{C_1, ..., C_m\}$ w.r.t. a TBox $\mathcal{T}$. But in this case the size of the lcs is already exponential w.r.t. an empty TBox [BKM99]. In this general case we have to take the product model

$$(\mathcal{I}_{C_1,\mathcal{T}} \times \cdots \times \mathcal{I}_{C_m,\mathcal{T}}, (d_{C_1}, \cdots, d_{C_m})),$$

which size is exponential in the size of M and $\mathcal{T}$, as input for the methods introduced in this section. Then the same steps as for the binary version can be applied.

4 Existence of the Most Specific Concept

We show now that the results obtained for the lcs, can be easily applied to the existence problem of the msc.

Example 28 (From [KM02])**.** The msc of the individual a w.r.t. the following KB

$$\mathcal{K}_1 = (\emptyset, \mathcal{A}_1), \text{ with } \mathcal{A}_1 = \{r(a, a)\}$$

doesn't exist, whereas w.r.t. the modified KB

$$\mathcal{K}_2 = (\{C \sqsubseteq \exists r.C\}, \mathcal{A}_2), \text{ with } \mathcal{A}_2 = \mathcal{A}_1 \cup \{C(a)\}$$

C is the msc of a.

To decide existence of the msc of an individual a w.r.t. a KB $\mathcal{K} = (\mathcal{T}, \mathcal{A})$, we again start with defining the set of msc-candidates for the msc by taking the k-characteristic concept of the canonical model $(\mathcal{I}_{\mathcal{K}}, d_a)$.

Lemma 29. *Let k be a natural number.*

1. $\mathcal{K} \models X^k(\mathcal{I}_\mathcal{K}, d_a)(a)$.

2. Let E be a concept with $rd(E) \leq k$ and $\mathcal{K} \models E(a)$.
   It holds that $X^k(\mathcal{I}_\mathcal{K}, d_a) \sqsubseteq_\mathcal{T} E$.

Proof of 1. We show the claim by induction on k.

$k = 0:$ It holds that
$$X^0(\mathcal{I}_\mathcal{K}, d_a) = \bigsqcap\{A \in N_C \mid d_a \in A^{\mathcal{I}_\mathcal{K}}\}. \tag{13}$$

For any concept name A in this conjunction it holds that $\mathcal{K} \models A(a)$ by definition of $\mathcal{I}_\mathcal{K}$ and therefore $\mathcal{K} \models X^0(\mathcal{I}_\mathcal{K}, d_a)(a)$.

$k > 0:$ By definition
$$X^k(\mathcal{I}_\mathcal{K}, d_a) = X^0(\mathcal{I}_\mathcal{K}, d_a) \sqcap \bigsqcap_{r \in N_R} \bigsqcap\{\exists r.X^{k-1}(\mathcal{I}_\mathcal{K}, d_y) \mid (d_a, d_y) \in r^{\mathcal{I}_\mathcal{K}}\}. \tag{14}$$

By I.H. (and Lemma 13 if y is a concept) it is implied
$$d_y \in (X^{k-1}(\mathcal{I}_\mathcal{K}, d_y))^{\mathcal{I}_\mathcal{K}}$$
and therefore $d_a \in (\exists r.X^{k-1}(\mathcal{I}_\mathcal{K}, d_y))^{\mathcal{I}_\mathcal{K}}$ and $d_a \in (X^k(\mathcal{I}_\mathcal{K}, d_a))^{\mathcal{I}_\mathcal{K}}$.

□

Proof of 2. The claim is proven by induction on the role-depth of an arbitrary concept E with $rd(E) \leq k$ and $\mathcal{K} \models E(a)$.

$rd(E) = 0:$ E is a conjunction of concept names of the form $\bigsqcap_{i=1,\ldots,n} A_i$. We show that the concept names A_i are contained in the top-level conjunction of $X^k(\mathcal{I}_\mathcal{K}, d_a)$. Since $\mathcal{K} \models E(a)$ and Lemma 7 holds it is implied that $d_a \in E^{\mathcal{I}_\mathcal{K}}$. It follows that $d_a \in A_i^{\mathcal{I}_\mathcal{K}}$ for all i. It is implied that $X^0(\mathcal{I}_\mathcal{K}, d_a) \sqsubseteq_\mathcal{T} E$.

$rd(E) = n:$ E is of the form
$$A_1 \sqcap \ldots \sqcap A_n \sqcap \exists r_1.E'_1 \sqcap \ldots \sqcap \exists r_m.E'_m$$

It can be shown like in the base case that the conjunction $A_1 \sqcap \ldots \sqcap A_n$ subsumes $X^k(\mathcal{I}_\mathcal{K}, d_a)$. Let $\exists r_j.E'_j$ be a top-level conjunct of E. Since $d_a \in (\exists r_j.E'_j)^{\mathcal{I}_\mathcal{K}}$ there is an r_j - successor d_y in $\mathcal{I}_\mathcal{K}$ of d_a with $d_y \in (E'_j)^{\mathcal{I}_\mathcal{K}}$ By induction hypothesis and $rd(E'_j) \leq k-1$ we now have that $X^{k-1}(\mathcal{I}_\mathcal{K}, d_y) \sqsubseteq_\mathcal{T} E'_j$ and therefore also $\exists r_j.X^{k-1}(\mathcal{I}_\mathcal{K}, d_y) \sqsubseteq_\mathcal{T} \exists r_j.E'_j$. It follows that $X^k(\mathcal{I}_\mathcal{K}, d_a) \sqsubseteq_\mathcal{T} \exists r_j.E'_j$ by (14).

□

Therefore $X^k(\mathcal{I}_\mathcal{K}, d_a) \equiv_\mathcal{T} k$-msc$_\mathcal{K}(a)$.

Now we use the canonical model of $X^k(\mathcal{I}_\mathcal{K}, d_a)$ w.r.t. the TBox component $\mathcal{T}$ of $\mathcal{K}$ and the model $(\mathcal{I}_\mathcal{K}, d_a)$ to check whether $X^k(\mathcal{I}_\mathcal{K}, d_a)$ is the msc.

Lemma 30. *Let $\mathcal{J}_{d_a}$ be the tree unraveling of $(\mathcal{I}_\mathcal{K}, d_a)$ in d_a and K the k-msc$_\mathcal{K}(a)$. It holds that $\mathcal{J}^k_{d_a} \precsim (\mathcal{I}_{K,\mathcal{T}}, d_K)$.*

Proof. a is an instance of the concept $X^k(\mathcal{I}_\mathcal{K}, d_a)$ w.r.t. $\mathcal{K}$ by Lemma 29. Since $X^k(\mathcal{I}_\mathcal{K}, d_a)$ has role depth k, we have $K \sqsubseteq_\mathcal{T} X^k(\mathcal{I}_\mathcal{K}, d_a)$ and therefore $d_K \in (X^k(\mathcal{I}_\mathcal{K}, d_a))^{\mathcal{I}_{K,\mathcal{T}}}$ by point 2 and 5 of Lemma 6. With Lemma 10 we have now $\mathcal{J}^k_{d_a} \lesssim (\mathcal{I}_{K,\mathcal{T}}, d_K)$. □

Lemma 31. *The concept C is the most specific concept of a w.r.t. $\mathcal{K}$ iff $(\mathcal{I}_\mathcal{K}, d_a) \simeq (\mathcal{I}_{C,\mathcal{T}}, d_C)$.*

Proof. "$\Rightarrow$":
Assume that C is the most specific concept of a w.r.t. $\mathcal{K}$. By Lemma 7 it is implied that $d_a \in C^{\mathcal{I}_\mathcal{K}}$. Since $\mathcal{I}_\mathcal{K}$ is a model of $\mathcal{T}$, it follows $(\mathcal{I}_{C,\mathcal{T}}, d_C) \lesssim (\mathcal{I}_\mathcal{K}, d_a)$ from Lemma 6.4.

We need to show $(\mathcal{I}_\mathcal{K}, d_a) \lesssim (\mathcal{I}_{C,\mathcal{T}}, d_C)$. The proof is analogous to the proof of Lemma 17.

"$\Leftarrow$":
Assume $(\mathcal{I}_\mathcal{K}, d_a) \simeq (\mathcal{I}_{C,\mathcal{T}}, d_C)$ for a concept C. Since $(\mathcal{I}_{C,\mathcal{T}}, d_C) \lesssim (\mathcal{I}_\mathcal{K}, d_a)$ and $\mathcal{I}_\mathcal{K}$ is a model of $\mathcal{T}$, a is an instance of C w.r.t. $\mathcal{K}$ by Lemma 6. It has to be shown that C is the *most specific* concept. Let C' be an arbitrary concept with $\mathcal{K} \models C'(a)$. It is implied that $d_a \in C'^{\mathcal{I}_\mathcal{K}}$ by Lemma 7. By Lemma 6.4 it holds that $(\mathcal{I}_{C',\mathcal{T}}, d_{C'}) \lesssim (\mathcal{I}_\mathcal{K}, d_a)$. Together with the assumption $(\mathcal{I}_\mathcal{K}, d_a) \lesssim (\mathcal{I}_{C,\mathcal{T}}, d_C)$ and transitivity of $\lesssim$ we get $(\mathcal{I}_{C',\mathcal{T}}, d_{C'}) \lesssim (\mathcal{I}_{C,\mathcal{T}}, d_C)$. It is implied by Lemma 6 that $C \sqsubseteq_\mathcal{T} C'$. Therefore C is the most specific concept of a. □

By this Lemma the existence of the msc can be characterized as follows.

Corollary 32. *The $msc_\mathcal{K}(C, D)$ exists iff there exists a k such that the canonical model of $X^k(\mathcal{I}_\mathcal{K}, d_a)$ w.r.t. $\mathcal{T}$ simulates $(\mathcal{I}_\mathcal{K}, d_a)$.*

To decide whether an appropriate k exists such that $X^k(\mathcal{I}_\mathcal{K}, d_a)$ simulates $(\mathcal{I}_\mathcal{K}, d_a)$ we further examine the structure of $(\mathcal{I}_\mathcal{K}, d_a)$. In Example 28 d_a has a self-loop in the model $(\mathcal{I}_{\mathcal{K}_1}, d_a)$, but the canonical models of $X^k(\mathcal{I}_{\mathcal{K}_1}, d_a)$ are finite for all $k \in \mathbb{N}$, because the TBox is empty. Therefore a simulation never exist. In comparison, the model $(\mathcal{I}_{\mathcal{K}_2}, d_a)$ has additionally a self-loop at d_C and the canonical models of $X^k(\mathcal{I}_{\mathcal{K}_2}, d_a)$ w.r.t. $\mathcal{T}_2$ also contain this loop.

The elements in $d_x \in \Delta^{\mathcal{I}_\mathcal{K}}$ with $x = b$ (for $b \in N_{I,\mathcal{A}}$) are *asynchronous* elements and the elements with $x = C$ for some concept C are *synchronous* elements. The model $(\mathcal{I}_\mathcal{K}, d_a)$ has an analogous structure compared to the product model $(\mathcal{I}_{C,\mathcal{T}} \times \mathcal{I}_{D,\mathcal{T}}, (d_C, d_D))$ in the sense that synchronous elements in $\Delta^{\mathcal{I}_\mathcal{K}}$ only have synchronous successor elements. Therefore analogous arguments as presented in Section 3.2 can be used to show, that a representation of the canonical model of $X^k(\mathcal{I}_\mathcal{K}, d_a)$ as a subtree of the tree unraveling of $(\mathcal{I}_\mathcal{K}, d_a)$ can be obtained. This representation is denoted by $\widehat{\mathcal{J}}^k_{d_a}$. This model is used to show an upper bound on the role-depth k of the msc.

Lemma 33. *Let $m := \mathsf{max}(\{rd(F) \mid F \in \mathsf{sub}(\mathcal{K})\})$ and $n := |N_{I,\mathcal{A}}|$. If the $msc_\mathcal{K}(a)$ exists, then $(\mathcal{I}_\mathcal{K}, d_a) \lesssim \widehat{\mathcal{J}}^{n^2+m+1}_{d_a}$.*

Proof. This Lemma can be proven using analogous arguments as in the proof of Lemma 26. □

The results of this section can be summarized in the following theorem.

Theorem 34. *Let $\mathcal{K} = (\mathcal{T}, \mathcal{A})$ be a KB and $a \in N_{I,\mathcal{A}}$. It is decidable in polynomial time whether the $msc_\mathcal{K}(a)$ exists. If the $msc_\mathcal{K}(a)$ exists, it can be computed in polynomial time.*

Proof sketch. First we compute the bound k as given in Lemma 33 and then the k-characteristic concept $X^k(\mathcal{I}_\mathcal{K}, d_a)$. The canonical model of K can be build according to Definition 3 in polynomial time [BBL05]. Then we check whether $(\mathcal{I}_\mathcal{K}, d_a) \lesssim (\mathcal{I}_{K,\mathcal{T}}, d_K)$ holds, which can be done in polynomial time. If yes, K is the msc and if no, the msc doesn't exist by Corollary 32. □

5 Conclusions

In this paper we have studied the conditions for the existence of the lcs and of the msc, if computed w.r.t. general TBoxes or cyclic ABoxes, respectively, written in the DL $\mathcal{EL}$. In this setting neither the lcs nor the msc need to exist. It was an open problem to give necessary and sufficient conditions for their existence. We showed that the existence problem of the msc and the lcs of two concepts is decidable in polynomial time. Furthermore, we showed that the role-depth of these most specific generalizations can be bounded by a polynomial. This upper bound k can be used to compute the msc or lcs, if it exists. Otherwise the computed concept can still serve as an approximation [PT11b].

Future work on the practical side includes to improve the described procedure in order to obtain a practical algorithm such that an appropriate implementation can be integrated into existing tools [ET12] for computing generalizations. On the theoretical side, we would like extend the results towards knowledge bases formulated in more expressive Horn-DLs than $\mathcal{EL}$.

References

[Baa03] Baader, Franz: Least Common Subsumers and Most Specific Concepts in a Description Logic with Existential Restrictions and Terminological Cycles. In: Gottlob, Georg (Hrsg.) ; Walsh, Toby (Hrsg.): *Proceedings of the 18th International Joint Conference on Artificial Intelligence*, Morgan Kaufman, 2003, S. 319–324

[Baa04] Baader, Franz: A Graph-Theoretic Generalization of the Least Common Subsumer and the Most Specific Concept in the Description Logic $\mathcal{EL}$. In: Hromkovic, J. (Hrsg.) ; Nagl, M. (Hrsg.): *Proceedings of the 30th International Workshop on Graph-Theoretic Concepts in Computer Science (WG 2004)* Bd. 3353, Springer-Verlag, 2004 (Lecture Notes in Computer Science), S. 177–188

[BBL05] Baader, Franz ; Brandt, Sebastian ; Lutz, Carsten: Pushing the $\mathcal{EL}$ Envelope. In: *Proceedings of the Nineteenth International Joint Conference on Artificial Intelligence IJCAI-05*, Morgan-Kaufmann Publishers, 2005

[BCM+03] Baader, F. (Hrsg.) ; Calvanese, D. (Hrsg.) ; McGuinness, D. (Hrsg.) ; Nardi, D. (Hrsg.) ; Patel-Schneider, P.F. (Hrsg.): *The Description Logic Handbook: Theory, Implementation, and Applications*. 2003

[BKM99] Baader, F. ; Küsters, R. ; Molitor, R.: Computing Least Common Subsumers in Description Logics with Existential Restrictions. Stockholm, Sweden, 1999, S. 96–101

[BST07] Baader, Franz ; Sertkaya, Barış ; Turhan, Anni-Yasmin: Computing the Least Common Subsumer w.r.t. a Background Terminology. In: *Journal of Applied Logic* 5 (2007), Nr. 3, S. 392–420

[ET12] Ecke, Andreas ; Turhan, Anni-Yasmin: Optimizations for the Role-Depth Bounded Least Common Subsumer in $\mathcal{EL}$+. In: Klinov, Pavel (Hrsg.) ; Horridge, Matthew (Hrsg.): *Proceedings of OWL: Experiences and Directions Workshop* Bd. 849, CEUR-WS.org, 2012 (CEUR Workshop Proceedings)

[KM02] Küsters, R. ; Molitor, R.: Approximating Most Specific Concepts in Description Logics with Existential Restrictions. In: *AI Communications* 15 (2002), Nr. 1, S. 47–59

[LPW10] Lutz, Carsten ; Piro, Robert ; Wolter, Frank: Enriching $\mathcal{EL}$-Concepts with Greatest Fixpoints. In: *Proceedings of the 19th European Conference on Artificial Intelligence (ECAI10)*, IOS Press, 2010

[LW10a] Lutz, Carsten ; Wolter, Frank: Deciding inseparability and conservative extensions in the description logic $\mathcal{EL}$. In: *Journal of Symbolic Computation* 45 (2010), Nr. 2, S. 194–228

[LW10b] Lutz, Carsten ; Wolter, Frank: Deciding inseparability and conservative extensions in the description logic $\mathcal{EL}$. In: *Journal of Symbolic Computation* 45 (2010), Nr. 2, S. 194 – 228. `http://dx.doi.org/DOI: 10.1016/j.jsc.2008.10.007`. – DOI DOI: 10.1016/j.jsc.2008.10.007. – ISSN 0747–7171

[PT11a] Peñaloza, R. ; Turhan, A.-Y.: A Practical Approach for Computing Generalization Inferences in $\mathcal{EL}$. In: Grobelnik, Marko (Hrsg.) ; Simperl, Elena (Hrsg.): *Proceedings of the 8th European Semantic Web Conference (ESWC'11)*, Springer-Verlag, 2011 (Lecture Notes in Computer Science)

[PT11b] Peñaloza, R. ; Turhan, A.-Y.: A Practical Approach for Computing Generalization Inferences in $\mathcal{EL}$, 2011

[SdH+07] Sioutos, Nicholas ; de Coronado, Sherri ; Haber, Margaret W. ; Hartel, Frank W. ; Shaiu, Wen-Ling ; Wright, Lawrence W.: NCI Thesaurus: A semantic model integrating cancer-related clinical and molecular information. In: *J. of Biomedical Informatics* 40 (2007), Nr. 1, 30–43. `http://dx.doi.org/10.1016/j.jbi.2006.02.013`. – DOI 10.1016/j.jbi.2006.02.013. – ISSN 1532–0464

[Spa00] Spackman, K.: Managing clinical terminology hierarchies using algorithmic calculation of subsumption: Experience with SNOMED-RT. In: *Journal of the American Medical Informatics Assoc.* (2000). – Fall Symposium Special Issue.

[W3C09] W3C OWL Working Group: *OWL 2 Web Ontology Language Document Overview*. W3C Recommendation, 27th October 2009. – `http://www.w3.org/TR/2009/REC-owl2-overview-20091027/`