Foundations for the Logical Difference of $\mathcal{EL}$-TBoxes[*]

Shasha Feng[1], Michel Ludwig[2], and Dirk Walther[2]

[1] Jilin University, China
fengss@jlu.edu.cn

[2] Theoretical Computer Science
TU Dresden, Germany
{michel,dirk}@tcs.inf.tu-dresden.de

Abstract

We investigate the logical difference problem between general $\mathcal{EL}$-TBoxes. The logical difference is the set of concept subsumptions that are logically entailed by a first TBox but not by a second one. We show how the logical difference between two $\mathcal{EL}$-TBoxes can be reduced to fixpoint reasoning w.r.t. $\mathcal{EL}$-TBoxes. Entailments of the first TBox can be represented by subsumptions of least fixpoint concepts by greatest fixpoint concepts, which can then be checked w.r.t. the second TBox. We present the foundations for a dedicated procedure based on a hypergraph representation of the fixpoint concepts without the use of automata-theoretic techniques, avoiding possible complexity issues of a reduction to modal μ-calculus reasoning. The subsumption checks are based on checking for the existence of simulations between the hypergraph representations of the fixpoint concepts and the TBoxes.

1 Introduction

Ontologies are widely used to represent domain knowledge. They contain specifications of objects, concepts and relationships that are often formalised using a logic-based language over a vocabulary that is particular to an application domain. Ontology languages based on description logics [2] have been widely adopted, e.g., description logics are underlying the Web Ontology Language (OWL) and its profiles.[1] Numerous ontologies have already been developed, in particular, in knowledge intensive areas such as the biomedical domain, and they are made available in dedicated repositories such as the NCBO bioportal.[2]

Ontologies constantly evolve, they are regularly extended, corrected and refined. As the size of ontologies increases, their continued development and maintenance becomes more challenging as well. In particular, the need to have automated tool support for detecting and representing differences between versions of an ontology is growing in importance for ontology engineering.

---

[1]http://www.w3.org/TR/owl2-overview/
[2]http://bioportal.bioontology.org

The *logical difference* is taken to be the set of queries that produce different answers when evaluated over distinct versions of an ontology. The language and the vocabulary of the queries can be adapted in such a way that exactly the differences of interest become visible, which can be independent of the syntactic representation of the ontologies. We consider ontologies formulated in the lightweight description logic $\mathcal{EL}$ [1, 3] and queries that are $\mathcal{EL}$-concept inclusions. The relevance of $\mathcal{EL}$ for ontologies is emphasised by the fact that many ontologies are largely, or even entirely, formulated in $\mathcal{EL}$. For instance, the dataset of the OWL Reasoner Evaluation workshop (ORE) in 2014 comprises 8 805 OWL-EL ontologies.[3]

The logical difference problem was introduced in [10] and investigated for $\mathcal{EL}$-terminologies in [9]. A hypergraph-based approach for $\mathcal{EL}$-terminologies was presented in [5], which was subsequently extended to $\mathcal{EL}$-terminologies with additional role inclusions, and domain and range restrictions of roles in [11]. In this paper we present an extension of the hypergraph-based approach for general $\mathcal{EL}$-TBoxes. Clearly, such an extension needs to account for the additional expressivity of general TBoxes w.r.t. terminologies. After normalisation, a terminology may contain at most one axiom of the form $\exists r.A \sqsubseteq X$ or $A_1 \sqcap \ldots \sqcap A_n \sqsubseteq X$ (with $n \geq 2$) for any concept name X, whereas a general TBox does not impose such a restriction.

For deciding the logical difference between two $\mathcal{EL}$-TBoxes $\mathcal{T}_1$, $\mathcal{T}_2$ w.r.t. a finite set of terms (signature) Σ, we show that one can represent all the concept inclusions which only use terms from Σ and which follow from $\mathcal{T}_1$ as finitely many subsumptions α_i between least and greatest fixpoint concepts. Subsequently, one has to check whether each subsumption α_i follows from $\mathcal{T}_2$. Fixpoint reasoning w.r.t. TBoxes has been studied in [4] based on automata-theoretic techniques. For our purposes, however, it is not immediate how a practical algorithm can be derived from the decidability and complexity results for the modal μ-calculus stated in the literature. In this paper we extend our approach to general $\mathcal{EL}$-TBoxes motivated by the success of our hypergraph-based procedure for the logical difference problem of $\mathcal{EL}$-terminologies [5,6,11] and by the fact that cycles in TBox definitions can be innately handled by simulations between hypergraph representations of the TBoxes.

We proceed as follows. In the next section we start with reviewing the DL $\mathcal{EL}$ together with its extensions with disjunction and the least and greatest fixpoint operators as well as with defining some auxiliary notions. In Section 3 we introduce a reduction of the logical difference problem for general $\mathcal{EL}$-TBoxes to fixpoint reasoning. However, this reduction is exponential, and fixpoint reasoning w.r.t. TBoxes is known to be ExpTime-complete [4]. Hence, the reduction only yields a double exponential upper bound for deciding the logical difference problem, whose complexity has been shown to be ExpTime-complete for $\mathcal{EL}$-TBoxes [12]. We therefore develop a hypergraph-based approach which avoids the complexity gap in Section 4. Our approach involves constructing hypergraph representations for least fixpoint concepts and graph representations for greatest fixpoint concepts. Subsumption between the fixpoints concepts w.r.t. a TBox can then be decided by checking for the existence of certain simulations between the (hyper)graph representations and by taking the TBox into account. The (hyper)graph representations can be built in exponential time, whereas the existence of the simulations can be checked in polynomial time. In Section 5 we then present an algorithm for solving the logical difference problem, which makes use of the notions developed in Section 4. Finally, we conclude the paper in Section 6.

---

[3] http://dl.kr.org/ore2014/

2 Preliminaries

Let $\mathsf{N_C}$, $\mathsf{N_R}$, and $\mathsf{N_V}$ be mutually disjoint sets of concept names, role names, and variable names, respectively. We assume these sets to be countably infinite. We typically use A, B and X, Y, Z to denote concept names, r, s, t to indicate role names, and x, y to denote variable names.

The set of $\mathcal{ELU}_{\perp,\mu,\nu}$-concepts $\mathcal{C}$ are built according to the following grammar rule:

$$\mathcal{C} ::= \perp \mid \top \mid A \mid \mathcal{C} \sqcap \mathcal{C} \mid \mathcal{C} \sqcup \mathcal{C} \mid \exists r.\mathcal{C} \mid x \mid \mu x.\mathcal{C} \mid \nu x.\mathcal{C}$$

where $A \in \mathsf{N_C}$, $r \in \mathsf{N_R}$, and $x \in \mathsf{N_V}$. We denote with $\mathcal{ELU}_{\perp,\mu,\nu}$ the set of all $\mathcal{ELU}_{\perp,\mu,\nu}$-concepts. For an $\mathcal{ELU}_{\perp,\mu,\nu}$-concept $\mathcal{C}$, the set of *free variables* in $\mathcal{C}$, denoted by $\mathrm{FV}(\mathcal{C})$, is defined inductively as follows: $\mathrm{FV}(\perp) := \emptyset$, $\mathrm{FV}(\top) := \emptyset$, $\mathrm{FV}(A) := \emptyset$, $\mathrm{FV}(\mathcal{D}_1 \sqcap \mathcal{D}_2) := \mathrm{FV}(\mathcal{D}_1) \cup \mathrm{FV}(\mathcal{D}_2)$, $\mathrm{FV}(\mathcal{D}_1 \sqcup \mathcal{D}_2) := \mathrm{FV}(\mathcal{D}_1) \cup \mathrm{FV}(\mathcal{D}_2)$, $\mathrm{FV}(\exists r.\mathcal{D}) := \mathrm{FV}(\mathcal{D})$, $\mathrm{FV}(x) := \{x\}$, $\mathrm{FV}(\mu x.\mathcal{D}) := \mathrm{FV}(\mathcal{D}) \setminus \{x\}$, and $\mathrm{FV}(\nu x.\mathcal{D}) := \mathrm{FV}(\mathcal{D}) \setminus \{x\}$. An $\mathcal{ELU}_{\perp,\mu,\nu}$-concept $\mathcal{C}$ is *closed* iff $\mathcal{C}$ does not contain free occurrences of variables, i.e. $\mathrm{FV}(\mathcal{C}) = \emptyset$; and $\mathcal{C}$ is *well-formed* if every subconcept of the form $\mu x.\mathcal{D}$ or $\nu x.\mathcal{D}$ occurring in $\mathcal{C}$ binds a fresh variable x. In the remainder of this paper we assume that every $\mathcal{ELU}_{\perp,\mu,\nu}$-concept is well-formed.

An $\mathcal{ELU}_{\perp,\mu,\nu}$-axiom is either a *concept inclusion* $\mathcal{C} \sqsubseteq \mathcal{D}$ or a *concept equation* $\mathcal{C} \equiv \mathcal{D}$, for $\mathcal{ELU}_{\perp,\mu,\nu}$-concepts $\mathcal{C}, \mathcal{D}$.

The semantics of $\mathcal{ELU}_{\perp,\mu,\nu}$ is defined using interpretations $\mathcal{I} = (\Delta^{\mathcal{I}}, \cdot^{\mathcal{I}})$, where the domain $\Delta^{\mathcal{I}}$ is a non-empty set, and $\cdot^{\mathcal{I}}$ is a function mapping each concept name A to a subset $A^{\mathcal{I}}$ of $\Delta^{\mathcal{I}}$ and every role name r to a binary relation $r^{\mathcal{I}} \subseteq \Delta^{\mathcal{I}} \times \Delta^{\mathcal{I}}$. Interpretations are extended to complex concepts using a function $\cdot^{\mathcal{I},\xi}$ that is parameterised by a *variable assignment function* that maps each variable $x \in \mathsf{N_V}$ to a set $\xi(x) \subseteq \Delta^{\mathcal{I}}$. Given an interpretation $\mathcal{I}$ and a variable assignment ξ, the extension of an $\mathcal{ELU}_{\perp,\mu,\nu}$-concept is defined inductively as follows: $\perp^{\mathcal{I},\xi} := \emptyset$, $\top^{\mathcal{I},\xi} := \Delta^{\mathcal{I}}$, $x^{\mathcal{I},\xi} := \xi(x)$ for $x \in \mathsf{N_V}$, $(\mathcal{C}_1 \sqcap \mathcal{C}_2)^{\mathcal{I},\xi} := \mathcal{C}_1^{\mathcal{I}} \cap \mathcal{C}_2^{\mathcal{I}}$, $(\exists r.\mathcal{C})^{\mathcal{I},\xi} := \{x \in \Delta^{\mathcal{I}} \mid \exists y \in \mathcal{C}^{\mathcal{I},\xi} : (x,y) \in r^{\mathcal{I}}\}$, $(\mu x.\mathcal{C})^{\mathcal{I},\xi} := \bigcap\{W \subseteq \Delta^{\mathcal{I}} \mid \mathcal{C}^{\mathcal{I},\xi[x \mapsto W]} \subseteq W\}$, and $(\nu x.\mathcal{C})^{\mathcal{I},\xi} := \bigcup\{W \subseteq \Delta^{\mathcal{I}} \mid W \subseteq \mathcal{C}^{\mathcal{I},\xi[x \mapsto W]}\}$, where $\xi[x \mapsto W]$ denotes the variable assignment ξ modified by mapping x to W.

For $\mathcal{ELU}_{\perp,\mu,\nu}$-concepts $\mathcal{C}$ and $\mathcal{D}$, an interpretation $\mathcal{I}$ *satisfies* $\mathcal{C}$, an axiom $\mathcal{C} \sqsubseteq \mathcal{D}$ or $\mathcal{C} \equiv \mathcal{D}$ if, respectively, $\mathcal{C}^{\mathcal{I},\xi_\emptyset} \neq \emptyset$, $\mathcal{C}^{\mathcal{I},\xi_\emptyset} \subseteq \mathcal{D}^{\mathcal{I},\xi_\emptyset}$, or $\mathcal{C}^{\mathcal{I},\xi_\emptyset} = \mathcal{D}^{\mathcal{I},\xi_\emptyset}$, where $\xi_\emptyset(x) = \emptyset$ for every $x \in \mathsf{N_V}$. We write $\mathcal{I} \models \alpha$ iff $\mathcal{I}$ satisfies the axiom α.

In particular we will be using the following sublanguages of $\mathcal{ELU}_{\perp,\mu,\nu}$: $\mathcal{EL}$, $\mathcal{EL}_\perp$, $\mathcal{ELU}$, $\mathcal{ELU}_\perp$, $\mathcal{ELU}_\mu$, $\mathcal{ELU}_{\perp,\mu}$, and $\mathcal{EL}_\nu$. As usual the letters $\mathcal{EL}$ indicate the presence of concept names and the concept constructors '$\top$', '$\sqcap$', and '$\exists r.\circ$'; the letter $\mathcal{U}$ stands for '$\sqcup$'. The subscript $\perp$ indicates the presence of '$\perp$', and the subscripts μ and ν indicate the availability of the least and greatest fixpoint operators '$\mu x.\circ$' and '$\nu x.\circ$', respectively, and of variable names. We use calligraphic letters $\mathcal{C}$, $\mathcal{D}$, or $\mathcal{E}$, to denote concepts that may contain a fixpoint operator, otherwise we use capital letters C, D, or E. We denote with L the set of all L-concepts, where $L \in \{\mathcal{EL}, \mathcal{EL}_\perp, \mathcal{ELU}, \mathcal{ELU}_\perp, \mathcal{ELU}_{\perp,\mu}, \mathcal{EL}_\nu\}$.

An $\mathcal{EL}$-*TBox* $\mathcal{T}$ is a finite set of axioms, where an axiom is either a *concept inclusion* $C \sqsubseteq D$, or a *concept equation* $C \equiv D$, for $\mathcal{EL}$-concepts C, D. An interpretation $\mathcal{I}$ *satisfies* a TBox $\mathcal{T}$ iff $\mathcal{I}$ satisfies all axioms in $\mathcal{T}$; in this case, we say that $\mathcal{I}$ is a *model* of $\mathcal{T}$. An $\mathcal{ELU}_{\perp,\mu,\nu}$-axiom α *follows* from an $\mathcal{EL}$-TBox $\mathcal{T}$, written $\mathcal{T} \models \alpha$, iff for all models $\mathcal{I}$ of $\mathcal{T}$, we have that $\mathcal{I} \models \alpha$. Note that for deciding $\mathcal{T} \models \alpha$ it is sufficient to only consider finite models $\mathcal{I}$ of $\mathcal{T}$ [4], i.e. interpretations $\mathcal{I} = (\Delta^{\mathcal{I}}, \cdot^{\mathcal{I}})$ where $\Delta^{\mathcal{I}}$ is finite and $\mathcal{I}$ satisfies $\mathcal{T}$. Deciding whether $\mathcal{T} \models C \sqsubseteq D$, for two $\mathcal{EL}$-concepts C and D, can be done in polynomial time in the size of $\mathcal{T}$ and C, D [1,3].

A signature Σ is a finite set of symbols from $\mathsf{N_C}$ and $\mathsf{N_R}$. The signature $\mathsf{sig}(\mathcal{C})$, $\mathsf{sig}(\alpha)$, or $\mathsf{sig}(\mathcal{T})$ of the concept $\mathcal{C}$, axiom α, or TBox $\mathcal{T}$ is the set of concept and role names occurring in $\mathcal{C}$,

α, or $\mathcal{T}$, respectively. Analogously, $\mathrm{sub}(\mathcal{C})$, $\mathrm{sub}(\alpha)$, or $\mathrm{sub}(\mathcal{T})$ denotes the set of subconcepts occurring in $\mathcal{C}$, α, or $\mathcal{T}$, respectively. We denote with L_Σ the set of L-concepts built from symbols in Σ only, where $L \in \{\mathcal{EL}, \mathcal{EL}_\perp, \mathcal{ELU}, \mathcal{ELU}_\perp, \mathcal{ELU}_{\perp,\mu}, \mathcal{EL}_\nu\}$.

An $\mathcal{EL}$-TBox $\mathcal{T}$ is *normalised* if it only consists of $\mathcal{EL}$-concept inclusions of the forms $\top \sqsubseteq B$, $A_1 \sqcap \ldots \sqcap A_n \sqsubseteq B$, $A \sqsubseteq \exists r.B$, or $\exists r.A \sqsubseteq B$, where $A, A_i, B \in \mathsf{N_C}$, $r \in \mathsf{N_R}$, and $n \geq 1$. Every $\mathcal{EL}$-TBox $\mathcal{T}$ can be normalised in polynomial time in the size of $\mathcal{T}$ with at most a linear increase in the size of the normalised TBox w.r.t. $\mathcal{T}$ such that the resulting TBox is a conservative extension of $\mathcal{T}$ [9].

$\mathcal{ELU}_{\perp,\mu}$-concepts can be simplified by pushing occurrences of $\perp$ to the top-most level, resulting either in $\perp$ or in a concept in which $\perp$ does not occur. For an $\mathcal{ELU}_{\perp,\mu}$-concept $\mathcal{C}$, we define the $\mathcal{ELU}_{\perp,\mu}$-concept $\mathrm{simp}_\perp(\mathcal{C})$ inductively as follows:

- $\mathrm{simp}_\perp(\varphi) := \varphi$ for $\varphi \in \mathsf{N_C} \cup \mathsf{N_V} \cup \{\perp, \top\}$;
- $\mathrm{simp}_\perp(\mathcal{C} \sqcap \mathcal{D}) := \perp$ if $\mathrm{simp}_\perp(\mathcal{C}) = \perp$ or $\mathrm{simp}_\perp(\mathcal{D}) = \perp$, otherwise $\mathrm{simp}_\perp(\mathcal{C} \sqcap \mathcal{D}) := \mathrm{simp}_\perp(\mathcal{C}) \sqcap \mathrm{simp}_\perp(\mathcal{D})$;
- $\mathrm{simp}_\perp(\mathcal{C} \sqcup \mathcal{D}) := \bigsqcup \{ \mathrm{simp}_\perp(\mathcal{F}) \mid \mathcal{F} \in \{\mathcal{C}, \mathcal{D}\}$ and $\mathrm{simp}_\perp(\mathcal{F}) \neq \perp \}$;
- $\mathrm{simp}_\perp(\exists r.\mathcal{C}) := \bigsqcup \{ \exists r.\mathrm{simp}_\perp(\mathcal{C}) \mid \mathrm{simp}_\perp(\mathcal{C}) \neq \perp \}$; and
- $\mathrm{simp}_\perp(\mu x.\mathcal{C}) := \perp$ if $\mathrm{simp}_\perp(\mathcal{C}) = \perp$, otherwise $\mathrm{simp}_\perp(\mu x.\mathcal{C}) := \mu x.\mathrm{simp}_\perp(\mathcal{C})$.

Example 1. *We have that $\mathrm{simp}_\perp(A \sqcap \exists r.(\perp \sqcup B)) = A \sqcap \exists r.B$, $\mathrm{simp}_\perp(A \sqcup \mu x.(\perp \sqcap x)) = A$, and $\mathrm{simp}_\perp((\exists r.\perp) \sqcup (\mu x.\perp)) = \perp$.*

An $\mathcal{ELU}_\perp$-concept C is in *disjunctive normal form* (DNF) if C is of the form $\bigsqcup_{i=1}^n C_i$, where C_i is an $\mathcal{EL}_\perp$-concept, for every $i \in \{1, \ldots, n\}$ with $n \geq 0$. Every $\mathcal{ELU}_\perp$-concept C can be transformed in an equivalent $\mathcal{ELU}_\perp$-concept in DNF in exponential time in the size of C by iteratively replacing subconcepts of the form $(D_1 \sqcup D_2) \sqcap D_3$ with $(D_1 \sqcap D_3) \sqcup (D_2 \sqcap D_3)$, and subconcepts of the form $\exists r.(D_1 \sqcup D_2)$ with $(\exists r.D_1) \sqcup (\exists r.D_2)$. We denote with $\mathrm{DNF}(C)$ the result of transforming C into DNF.

For an $\mathcal{ELU}_\perp$-concept C, we set $\mathrm{concepts}_{\mathcal{EL}}(C) = \{C_1, \ldots, C_n\}$, where $C_1, \ldots, C_n$ are $\mathcal{EL}$-concepts $(n \geq 0)$ such that $\mathrm{simp}_\perp(\mathrm{DNF}(C)) = \bigsqcup_{i=1}^n C_i$.

Definition 2 ($\mathcal{EL}$-Unfoldings of Fixpoint Concepts). *Let $L \in \{\mathcal{ELU}_{\perp,\mu}, \mathcal{EL}_\nu\}$ and let $\mathcal{C}$ be a closed L-concept. We define the L-unfolding of $\mathcal{C}$, $UF(\mathcal{C})$, as the smallest set of L-concepts closed under the following conditions:*

(i) $\mathcal{C} \in UF(\mathcal{C})$;

(ii) If $\mathcal{C}' \in UF(\mathcal{C})$ and $\tau x.\mathcal{D}$ is a subconcept of $\mathcal{C}'$ for $\tau \in \{\mu, \nu\}$, then $\mathrm{simp}_\perp(\mathcal{C}'') \in UF(\mathcal{C})$, where $\mathcal{C}''$ results from $\mathcal{C}'$ by substituting all occurrences of the subconcept $\tau x.\mathcal{D}$ in $\mathcal{C}'$ with $\mathcal{D}[x \mapsto \tau x.\mathcal{D}]$.

For an $\mathcal{ELU}_{\perp,\mu}$-concept $\mathcal{C}$ ($\mathcal{EL}_\nu$-concept $\mathcal{D}$) we define the set $UF_\mu(\mathcal{C})$ ($UF_\nu(\mathcal{D})$) to be set of all $\mathcal{ELU}_\perp$-concepts ($\mathcal{EL}$-concepts) obtained from concepts $\mathcal{C}' \in UF(\mathcal{C})$ ($\mathcal{D}' \in UF(\mathcal{D})$) by substituting every occurrence of a concept $\mu x.C$ in $\mathcal{C}'$ with $\perp$ ($\nu x.C$ in $\mathcal{D}'$ with $\top$). Finally, for an $\mathcal{ELU}_{\perp,\mu}$-concept $\mathcal{C}$ we set $UF_{\mathcal{EL}}(\mathcal{C}) := \bigcup_{C \in UF_\mu(\mathcal{C})} \mathrm{concepts}_{\mathcal{EL}}(C)$, and for an $\mathcal{EL}_\nu$-concept $\mathcal{D}$ we define $UF_{\mathcal{EL}}(\mathcal{D}) := UF_\nu(\mathcal{D})$.

Example 3. *Let $\mathcal{C} = \mu x.(A \sqcup \exists r.x)$ and $\mathcal{D} = A \sqcap \nu y.(\exists r.y)$. Then we have that $UF_{\mathcal{EL}}(\mathcal{C}) = \{A, \exists r.A, \exists r.\exists r.A, \ldots\}$ and $UF_{\mathcal{EL}}(\mathcal{D}) = \{A \sqcap \exists r.\top, A \sqcap \exists r.\exists r.\top, \ldots\}$.*

Together with the finite model property of $\mathcal{ELU}_{\perp,\mu,\nu}$ [4], an $\mathcal{ELU}_{\perp,\mu}$-concept $\mathcal{C}$ is equivalent to the (possibly infinite) disjunction of $\mathcal{EL}$-concepts in $UF_{\mathcal{EL}}(\mathcal{C})$. Analogously, an $\mathcal{EL}_\nu$-concept $\mathcal{D}$ is equivalent to the (possibly infinite) conjunction of $\mathcal{EL}$-concepts in $UF_{\mathcal{EL}}(\mathcal{D})$.

Lemma 4. *Let $\mathcal{C}$ be a closed $\mathcal{ELU}_{\perp,\mu}$-concept and let $\mathcal{D}$ be a closed $\mathcal{EL}_\nu$-concept. Additionally, let $\mathcal{I} = (\Delta^{\mathcal{I}}, \cdot^{\mathcal{I}})$ be an interpretation such that $\Delta^{\mathcal{I}}$ is finite and let ξ be a variable assignment. Then the following statements hold:*

(i) $\mathcal{C}^{\mathcal{I},\xi} = \bigcup_{C' \in UF_{\mathcal{EL}}(\mathcal{C})} (C')^{\mathcal{I},\xi}$;

(ii) $\mathcal{D}^{\mathcal{I},\xi} = \bigcap_{D' \in UF_{\mathcal{EL}}(\mathcal{D})} (D')^{\mathcal{I},\xi}$.

3   Deciding the Logical Difference between $\mathcal{EL}$-TBoxes via Fixpoint Reasoning

The logical difference between two $\mathcal{EL}$-TBoxes witnessed by concept inclusions over a signature Σ is defined as follows.

Definition 5. *The Σ-concept difference between two $\mathcal{EL}$-TBoxes $\mathcal{T}_1$ and $\mathcal{T}_2$ for a signature Σ is the set $\mathsf{cDiff}_\Sigma(\mathcal{T}_1, \mathcal{T}_2)$ of all $\mathcal{EL}$-concept inclusions α such that $\mathsf{sig}(\alpha) \subseteq \Sigma$, $\mathcal{T}_1 \models \alpha$, and $\mathcal{T}_2 \not\models \alpha$.*

Generally, the difference set $\mathsf{cDiff}_\Sigma(\mathcal{T}_1, \mathcal{T}_2)$ is infinite if it is not empty. In the case of $\mathcal{EL}$-terminologies (i.e. $\mathcal{EL}$-TBoxes of a simpler form as defined in, e.g., [9]), a "primitive witnesses" theorem from [9] states that we only have to consider two specific types of concept differences. If there is a subsumption $C \sqsubseteq D \in \mathsf{cDiff}_\Sigma(\mathcal{T}_1, \mathcal{T}_2)$ for two terminologies $\mathcal{T}_1$ and $\mathcal{T}_2$, then there exists a concept name $A \in \Sigma$ such that A occurs either on the left-hand or the right-hand side of a subsumption in $\mathsf{cDiff}_\Sigma(\mathcal{T}_1, \mathcal{T}_2)$. Thus, for checking whether $\mathsf{cDiff}_\Sigma(\mathcal{T}_1, \mathcal{T}_2) = \emptyset$, we only have to consider such simple subsumptions. If $\mathcal{T}_1$ and $\mathcal{T}_2$ are general $\mathcal{EL}$-TBoxes, however, the situation is different as illustrated by the following example.

Example 6. *Let $\mathcal{T}_1 = \{X \equiv A_1 \sqcap A_2, X \sqsubseteq \exists r.\top\}$, $\mathcal{T}_2 = \emptyset$, and let $\Sigma = \{A_1, A_2, r\}$. Note that $\mathcal{T}_1$ is not a terminology as the concept name X occurs twice on the left-hand side of an axiom. Then every inclusion $\alpha \in \mathsf{cDiff}_\Sigma(\mathcal{T}_1, \mathcal{T}_2)$ is equivalent to the inclusion $A_1 \sqcap A_2 \sqsubseteq \exists r.\top$, i.e., there does not exist a difference of the form $\psi \sqsubseteq \theta$, where ψ or θ is a concept name from Σ.*
For $\mathcal{T}_3 = \{A \sqsubseteq X, \exists r.X \sqsubseteq \exists r.Y, Y \sqsubseteq B\}$, $\mathcal{T}_4 = \{A \sqsubseteq B\}$, and $\Sigma = \{A, B, r\}$, we have that $\mathsf{cDiff}_\Sigma(\mathcal{T}_3, \mathcal{T}_4) = \emptyset$.
Finally, let $\mathcal{T}_5 = \{\exists r.A_1 \sqsubseteq B, \exists r.A_2 \sqsubseteq B\}$, $\mathcal{T}_6 = \{\exists r.X \sqsubseteq B, A_1 \sqsubseteq X, A_2 \sqsubseteq X\}$, and $\Sigma = \{A_1, A_2, B, r\}$. Then $\mathsf{cDiff}_\Sigma(\mathcal{T}_5, \mathcal{T}_6) = \emptyset$.

We need to account for a new kind of differences $C \sqsubseteq D \in \mathsf{cDiff}_\Sigma(\mathcal{T}_1, \mathcal{T}_2)$ which are induced by a concept name $X \in \mathsf{sig}(\mathcal{T}_1)$ such that $X \notin \Sigma$, $\mathcal{T}_1 \models C \sqsubseteq X$, and $\mathcal{T}_1 \models X \sqsubseteq D$. We obtain the following witness theorem for $\mathcal{EL}$-TBoxes as an extension of the witness theorem for $\mathcal{EL}$-terminologies.

Theorem 7 (Witness Theorem). *Let $\mathcal{T}_1$, $\mathcal{T}_2$ be two normalised $\mathcal{EL}$-TBoxes and let Σ be a signature. Then, $\mathsf{cDiff}_\Sigma(\mathcal{T}_1, \mathcal{T}_2) \neq \emptyset$ iff one of the following conditions is satisfied:*

(i) $\varphi \sqsubseteq A \in \mathsf{cDiff}_\Sigma(\mathcal{T}_1, \mathcal{T}_2)$ for some $\varphi \in \mathcal{EL}_\Sigma$ and $A \in \Sigma$, or

(ii) $A \sqsubseteq \psi \in \mathsf{cDiff}_\Sigma(\mathcal{T}_1, \mathcal{T}_2)$ for some $\psi \in \mathcal{EL}_\Sigma$ and $A \in \Sigma$, or

(iii) there exists $X \in \mathsf{sig}(\mathcal{T}_1) \setminus \Sigma$ and $\varphi, \psi \in \mathcal{EL}_\Sigma$ such that $\mathcal{T}_1 \models \varphi \sqsubseteq X$, $\mathcal{T}_1 \models X \sqsubseteq \psi$, and $\mathcal{T}_2 \not\models \varphi \sqsubseteq \psi$, or

(iv) $\top \sqsubseteq \psi \in \mathsf{cDiff}_\Sigma(\mathcal{T}_1, \mathcal{T}_2)$ for some $\psi \in \mathcal{EL}_\Sigma$.

The proof of the witness theorem for terminologies [9] is based on analysing a subsumption $\mathcal{T}_1 \models \varphi \sqsubseteq \psi$ for $\varphi \sqsubseteq \psi \in \mathsf{cDiff}_\Sigma(\mathcal{T}_1, \mathcal{T}_2)$ syntactically, using a sequent calculus [8]. A similar technique can be used for the proof of Theorem 7.

We use the following as a running example to illustrate our procedure for deciding the logical difference between $\mathcal{EL}$-TBoxes.

Example 8. *Let $\mathcal{T}_1$ be a normalised $\mathcal{EL}$-TBox consisting of the following axioms:*

$$
\begin{array}{llll}
A \sqsubseteq Y_1 & \exists s.Y_1 \sqsubseteq X_1 & X_1 \sqsubseteq \exists s.Z_1 & Z_1 \sqsubseteq B \\
\exists r.V_1 \sqsubseteq Y_1 & & & Z_1 \sqsubseteq \exists t.Z_1 \\
A' \sqcap Y_1 \sqsubseteq V_1 & & &
\end{array}
$$

Additionally, let $\mathcal{T}_2$ be a normalised $\mathcal{EL}$-TBox consisting of the following axioms:

$$
\begin{array}{llll}
A \sqsubseteq X_2 & \exists r.X_2 \sqsubseteq Y_2 & X_2 \sqsubseteq B & Y_2 \sqsubseteq X_2 \\
\exists r.V_2 \sqsubseteq X_2 & \exists r.Z_2 \sqsubseteq Z_2 & X_2 \sqsubseteq \exists t.W_2 & Z_2 \sqsubseteq \exists t.X_2 \\
\exists r.X_2 \sqsubseteq V_2 & A \sqsubseteq Z_2 & W_2 \sqsubseteq \exists t.X_2 & Z_2 \sqsubseteq B \\
& & W_2 \sqsubseteq B &
\end{array}
$$

Finally, let $\Sigma = \{A, A', B, r, s, t\}$ be a signature. Then $\mathsf{cDiff}_\Sigma(\mathcal{T}_1, \mathcal{T}_2) = \emptyset$.

According to Theorem 7 we have to consider four types of differences for deciding whether $\mathsf{cDiff}_\Sigma(\mathcal{T}_1, \mathcal{T}_2) = \emptyset$ in the case of two general $\mathcal{EL}$-TBoxes $\mathcal{T}_1$ and $\mathcal{T}_2$. In this paper we focus on differences of Type (iii); differences of Type (i), (ii), and (iv) can be handled similarly.

For a concept name X in $\mathsf{sig}(\mathcal{T}_1) \setminus \Sigma$, we say that X *witnesses* a Σ-concept difference if there exists a $C \sqsubseteq D \in \mathsf{cDiff}_\Sigma(\mathcal{T}_1, \mathcal{T}_2)$ with $\mathcal{T}_1 \models C \sqsubseteq X$ and $\mathcal{T}_1 \models X \sqsubseteq D$. For checking whether X is indeed a difference witness, we need to consider all such subsumptions w.r.t. X that may occur in $\mathsf{cDiff}_\Sigma(\mathcal{T}_1, \mathcal{T}_2)$. To this end we construct two Σ-concepts (using symbols from Σ only) $\mathrm{B}_{\mathcal{T}_1}^\Sigma(X)$ and $\mathrm{F}_{\mathcal{T}_1}^\Sigma(X)$ formulated respectively in $\mathcal{ELU}_{\perp,\mu}$ and in $\mathcal{EL}_\nu$. The least fixpoint concept $\mathrm{B}_{\mathcal{T}_1}^\Sigma(X)$ describes the disjunction of all Σ-concepts that are subsumed by (or that entail) X w.r.t. $\mathcal{T}_1$. Analogously the greatest fixpoint concept $\mathrm{F}_{\mathcal{T}_1}^\Sigma(X)$ describes the conjunction of all Σ-concepts that subsume (or that are entailed by) X w.r.t. $\mathcal{T}_1$. The use of fixpoint operators enables us to obtain a finite concept description that is equivalent to an infinite disjunction or conjunction.

Before we can give a formal definition for the concept $\mathrm{B}_{\mathcal{T}_1}^\Sigma(X)$, we have to introduce the following auxiliary notion to handle concept names X in the definition of $\mathrm{B}_{\mathcal{T}_1}^\Sigma(X)$ for which there exist axioms of the form $Z_1 \sqcap \ldots \sqcap Z_n \sqsubseteq Z$ in a normalised TBox $\mathcal{T}_1$ such that $\mathcal{T}_1 \models Z \sqsubseteq X$. Intuitively, given a concept name X, we construct a set $\mathrm{Conj}_{\mathcal{T}_1}(X)$ consisting of sets of concept names which has the property that for every $\mathcal{EL}$-concept D with $\mathcal{T}_1 \models D \sqsubseteq X$, there exists a set $S = \{Y_1, \ldots, Y_m\} \in \mathrm{Conj}_{\mathcal{T}_1}(X)$ such that $\mathcal{T}_1 \models D \sqsubseteq Y_i$, for all $i \in \{1, \ldots, m\}$, follows without involving any axioms of the form $Z_1 \sqcap \ldots \sqcap Z_n \sqsubseteq Z$. Nested implications between such axioms also have to be taken into account.

Definition 9 (Sets of Conjuncts). *Let $\mathcal{T}$ be a normalised $\mathcal{EL}$-TBox and let $X \in \mathsf{N_C}$. We define the set $\mathrm{Conj}_{\mathcal{T}}(X) \subseteq 2^{\mathsf{sig}(\mathcal{T}) \cap \mathsf{N_C}}$ to be smallest set inductively defined as follows:*

- *$\{X\} \in \mathrm{Conj}_{\mathcal{T}}(X)$;*

- *if $S \in \mathrm{Conj}_{\mathcal{T}}(X)$, $Y \in S$, and $Z_1 \sqcap \ldots \sqcap Z_n \sqsubseteq Z \in \mathcal{T}$ such that $n \geq 2$ and $\mathcal{T} \models Z \sqsubseteq Y$, then $S \setminus \{Y\} \cup \{Z_1, \ldots, Z_n\} \in \mathrm{Conj}_{\mathcal{T}}(X)$.*

Example 10. *Let $\mathcal{T} = \{A_1 \sqcap Y \sqsubseteq X,\ A_2 \sqcap Z \sqsubseteq Y,\ A_3 \sqsubseteq Z,\ B_1 \sqcap B_2 \sqsubseteq X\}$. Then $\mathrm{Conj}_{\mathcal{T}}(X) = \{\{X\}, \{Y, A_1\}, \{A_1, A_2, Z\}, \{B_1, B_2\}\}$.*

Note that for every concept name X the set $\mathrm{Conj}_{\mathcal{T}}(X)$ is finite as $\mathsf{sig}(\mathcal{T}) \cap \mathsf{N_C}$ is finite.

Definition 11 (Σ-Subsumees Representation). *Let $\mathcal{T}$ be a normalised $\mathcal{EL}$-TBox and let Σ be a signature. For $X \in \mathsf{sig}(\mathcal{T})$ and for a mapping $\zeta \colon \mathsf{N_C} \to \mathsf{N_V}$, we define a closed $\mathcal{ELU}_{\bot,\mu,\Sigma}$-concept $B_{\mathcal{T}}^{\Sigma}(X, \zeta)$ as follows. We set $B_{\mathcal{T}}^{\Sigma}(X, \zeta) = \top$ if $\mathcal{T} \models \top \sqsubseteq X$; otherwise $B_{\mathcal{T}}^{\Sigma}(X, \zeta)$ is defined recursively in the following way:*

- *If $X \in \mathsf{dom}(\zeta)$, then $B_{\mathcal{T}}^{\Sigma}(X, \zeta) := \zeta(X)$*

- *If $X \notin \mathsf{dom}(\zeta)$, we set $B_{\mathcal{T}}^{\Sigma}(X, \zeta) := \mu x. \displaystyle\bigsqcup_{\substack{S \in \mathrm{Conj}_{\mathcal{T}}(X) \\ S = \{Y_1, \ldots, Y_m\}}} (Y_1' \sqcap \ldots \sqcap Y_m')$*

where x is a fresh variable, and Y_i' $(1 \leq i \leq n)$ is defined as follows for $\zeta' := \zeta \cup \{X \mapsto x\}$:

$$Y_i' = \bigsqcup_{\substack{\mathcal{T} \models B \sqsubseteq Y_i \\ B \in \Sigma}} B \sqcup \bigsqcup_{\substack{\exists r.Z \sqsubseteq Y \in \mathcal{T} \\ r \in \Sigma \\ \mathcal{T} \models Y \sqsubseteq Y_i}} \exists r.B_{\mathcal{T}}^{\Sigma}(Z, \zeta')$$

Finally, we set $B_{\mathcal{T}}^{\Sigma}(X) := B_{\mathcal{T}}^{\Sigma}(X, \emptyset)$.

Intuitively, the construction of $\mathrm{B}_{\mathcal{T}}^{\Sigma}(X)$ starts from X and recursively collects all the concept names contained in Σ and all the left-hand sides of axioms in $\mathcal{T}$ that could be relevant for X to be entailed by a Σ-concept w.r.t. $\mathcal{T}$. By taking into account all possible axioms that could lead to a logical entailment, it is guaranteed that we capture every Σ-concept from which X follows w.r.t. $\mathcal{T}$. Reasoning involving axioms of the form $Z_1 \sqcap \ldots \sqcap Z_n \sqsubseteq Z$ is handled by the set $\mathrm{Conj}_{\mathcal{T}}(X)$. Infinite recursion over concepts of the form $\exists r.C$ is avoided by keeping track of the concept names that have been visited already using the mapping ζ.

Example 12. *Let $\mathcal{T}_1$ and Σ be defined as in Example 8. Then $B_{\mathcal{T}_1}^{\Sigma}(A) = \mu x.A$, $B_{\mathcal{T}_1}^{\Sigma}(A') = \mu x.A'$, $B_{\mathcal{T}_1}^{\Sigma}(B) = \mu x.B$, $B_{\mathcal{T}_1}^{\Sigma}(Y_1) = \mu y_1.(A \sqcup \exists r.\mu v_1.(A' \sqcap y_1))$, $B_{\mathcal{T}_1}^{\Sigma}(V_1) = \mu v_1.(A' \sqcap \mu y_1.(A \sqcup \exists r.v_1))$, and $B_{\mathcal{T}_1}^{\Sigma}(X_1) = \mu x_1.(\exists s.\mu y_1.(A \sqcup \exists r.\mu v_1.(A' \sqcap y_1)))$, and $B_{\mathcal{T}_1}^{\Sigma}(Z_1) = \mu x.\bot$.*

$\mathrm{B}_{\mathcal{T}_1}^{\Sigma}(X)$ represents the disjunction of all Σ-concepts that are subsumed by X w.r.t. $\mathcal{T}_1$.

Theorem 13. *Let $\mathcal{T}$ be a normalised $\mathcal{EL}$-TBox, let Σ be a signature, and let $X \in \mathsf{sig}(\mathcal{T})$. Then for every $\mathcal{EL}_{\Sigma}$-concept C it holds that $\mathcal{T} \models C \sqsubseteq X$ iff $\models C \sqsubseteq B_{\mathcal{T}}^{\Sigma}(X)$.*

The concept $\mathrm{F}_{\mathcal{T}_1}^{\Sigma}(X)$ can be defined in a way such that it represents the conjunction of all the Σ-concepts that subsume X w.r.t. $\mathcal{T}_1$.

Definition 14 (Σ-Subsumers Representation). *Let $\mathcal{T}$ be a normalised $\mathcal{EL}$-TBox, let Σ be a signature. For $X \in \mathsf{sig}(\mathcal{T})$ and for a mapping $\eta \colon \mathsf{N_C} \to \mathsf{N_V}$, we define a closed $\mathcal{EL}_{\nu,\Sigma}$-concept $F_{\mathcal{T}}^{\Sigma}(X, \eta)$ as follows:*

- *If $X \in \mathsf{dom}(\eta)$, then*

$$F_{\mathcal{T}}^{\Sigma}(X, \eta) := \eta(X)$$

- *If $X \notin \mathsf{dom}(\eta)$, we set*

$$F_{\mathcal{T}}^{\Sigma}(X, \eta) := \nu x. \bigsqcap_{\substack{\mathcal{T} \models X \sqsubseteq B \\ B \in \Sigma}} B \sqcap \bigsqcap_{\substack{Y \sqsubseteq \exists r.Z \in \mathcal{T} \\ r \in \Sigma \\ \mathcal{T} \models X \sqsubseteq Y}} \exists r.F_{\mathcal{T}}^{\Sigma}(Z, \eta')$$

　　　where x is a fresh variable and $\eta' := \eta \cup \{X \mapsto x\}$.

Finally, we set $F_{\mathcal{T}}^{\Sigma}(X) := F_{\mathcal{T}}^{\Sigma}(X, \emptyset)$.

Example 15. *Let $\mathcal{T}_1$ and Σ be defined as in Example 8. Then $F_{\mathcal{T}_1}^{\Sigma}(A) = \nu x.A$, $F_{\mathcal{T}_1}^{\Sigma}(B) = \nu x.B$, $F_{\mathcal{T}_1}^{\Sigma}(V_1) = \nu x.\top$, $F_{\mathcal{T}_1}^{\Sigma}(Y_1) = \nu x.\top$, $F_{\mathcal{T}_1}^{\Sigma}(Z_1) = \nu z_1.(B \sqcap \exists t.z_1)$, and $F_{\mathcal{T}_1}^{\Sigma}(X_1) = \nu x_1.\exists s.\nu z_1.(B \sqcap \exists t.z_1)$.*

Theorem 16. *Let $\mathcal{T}$ be a normalised $\mathcal{EL}$-TBox, let Σ be a signature, and let $X \in \mathsf{sig}(\mathcal{T})$. Then for every $\mathcal{EL}_{\Sigma}$-concept C it holds that $\mathcal{T} \models X \sqsubseteq C$ iff $\models F_{\mathcal{T}}^{\Sigma}(X) \sqsubseteq C$.*

　　Given the fixpoint descriptions of the set of subsumees and subsumers of $X \notin \Sigma$ w.r.t. $\mathcal{T}_1$, we can then decide the difference witness status of X by checking whether $\mathcal{T}_2 \models B_{\mathcal{T}_1}^{\Sigma}(X) \sqsubseteq F_{\mathcal{T}_1}^{\Sigma}(X)$. These subsumptions hold in the context of $\mathcal{T}_2$ and they use the concepts $B_{\mathcal{T}_1}^{\Sigma}(X)$ and $F_{\mathcal{T}_1}^{\Sigma}(X)$ that internalise relevant parts of $\mathcal{T}_1$. By handling the differences of Type (i), (ii), and (iv) in a similar way, we obtain an algorithm for deciding the existence of a Σ-concept difference between two $\mathcal{EL}$-TBoxes; see [7].

Theorem 17. *Let $\mathcal{T}_1, \mathcal{T}_2$ be normalised $\mathcal{EL}$-TBoxes, let Σ be a signature, and let $X \in \mathsf{sig}(\mathcal{T}_1) \backslash \Sigma$. Then $\mathcal{T}_2 \models B_{\mathcal{T}_1}^{\Sigma}(X) \sqsubseteq F_{\mathcal{T}_1}^{\Sigma}(X)$ iff for every $C, D \in \mathcal{EL}_{\Sigma}$ with $\mathcal{T}_1 \models C \sqsubseteq X$ and $\mathcal{T}_1 \models X \sqsubseteq D$ it holds that $\mathcal{T}_2 \models C \sqsubseteq D$.*

　　By internalising the TBox $\mathcal{T}_2$ into a μ-calculus formula involving greatest fixpoint constructors, the subsumption $\mathcal{T}_2 \models B_{\mathcal{T}_1}^{\Sigma}(X) \sqsubseteq F_{\mathcal{T}_1}^{\Sigma}(X)$ can be polynomially reduced to μ-calculus satisfiability, which is known to be EXPTIME-complete [4]. However, since in general the concepts $B_{\mathcal{T}_1}^{\Sigma}(X)$ and $F_{\mathcal{T}_1}^{\Sigma}(X)$ can be of exponential size w.r.t. the size of $\mathcal{T}_1$, we only obtain a double exponential bound for the running time of our algorithm (as sketched above) for deciding the existence of logical differences between $\mathcal{EL}$-TBoxes in this way. This is in contrast to the fact that the existence of a logical difference between $\mathcal{EL}$-TBoxes can be decided in exponential time in the size of the input TBoxes and the signature [12]. We will not investigate the existence of a polynomial reduction of the logical difference problem to μ-calculus satisfiability further. Instead, we show how to construct (hyper)graph representations of $B_{\mathcal{T}_1}^{\Sigma}(X)$ and $F_{\mathcal{T}_1}^{\Sigma}(X)$ without having to build the respective fixpoint concepts first. In the following section we will develop a hypergraph-based technique for the subsumption problem w.r.t. $\mathcal{T}_2$ between sets of concepts representing the subsumees and the subsumers of X w.r.t. $\mathcal{T}_1$.

4　Subsumption between Sets of Concepts

In this section we develop a method for deciding the subsumption problem between certain sets of concepts, which are possibly infinite. More concretely, we consider sets of subsumees or subsumers of an $\mathcal{EL}$-concept w.r.t. an $\mathcal{EL}$-TBox, or unfoldings of fixpoint concepts.

　　We recall from the previous section that for two normalised $\mathcal{EL}$-TBoxes $\mathcal{T}_1$, $\mathcal{T}_2$, and a signature Σ, a concept name $X \in \mathsf{sig}(\mathcal{T}_1) \setminus \Sigma$ is not a Type-(iii) witness iff for every $C \in \mathcal{EL}_{\Sigma}$ and for every $D \in \mathcal{EL}_{\Sigma}$ with $\mathcal{T}_1 \models C \sqsubseteq X$ and $\mathcal{T}_1 \models X \sqsubseteq D$ it follows that $\mathcal{T}_2 \models C \sqsubseteq D$. Our approach now consists in first representing all the $\mathcal{EL}_{\Sigma}$-concepts C with $\mathcal{T}_1 \models C \sqsubseteq X$ in a Σ-subsumee hypergraph $\exp_{\leftarrow}^{\mathcal{T},\Sigma}(X)$ and all the $\mathcal{EL}_{\Sigma}$-concepts D with $\mathcal{T}_1 \models X \sqsubseteq D$ in a Σ-subsumer graph $\exp_{\rightarrow}^{\mathcal{T},\Sigma}(X)$. More precisely, the unfoldings of $\exp_{\leftarrow}^{\mathcal{T},\Sigma}(X)$ and $\exp_{\rightarrow}^{\mathcal{T},\Sigma}(X)$ (i.e. the set of $\mathcal{EL}_{\Sigma}$-concepts represented by $\exp_{\leftarrow}^{\mathcal{T},\Sigma}(X)$ and $\exp_{\rightarrow}^{\mathcal{T},\Sigma}(X)$) respectively correspond to the set of $\mathcal{EL}_{\Sigma}$-concepts that are subsumed by X w.r.t. $\mathcal{T}_1$, and the set of $\mathcal{EL}_{\Sigma}$-concepts that subsume X w.r.t. $\mathcal{T}_1$. Note that $\exp_{\leftarrow}^{\mathcal{T},\Sigma}(X)$ and $\exp_{\rightarrow}^{\mathcal{T},\Sigma}(X)$ respectively stand for the

least fixpoint concept $\mathrm{B}_{\mathcal{T}_1}^{\Sigma}(X)$ and for the greatest fixpoint concept $\mathrm{F}_{\mathcal{T}_1}^{\Sigma}(X)$ (as defined in the previous section). Subsequently, we use these (hyper)graph representations in the context of $\mathcal{T}_2$ to check whether every unfolding of $\exp_{\leftarrow}^{\mathcal{T},\Sigma}(X)$ entails every unfolding of $\exp_{\rightarrow}^{\mathcal{T},\Sigma}(X)$ w.r.t. $\mathcal{T}_2$.

The remainder of the section is organised as follows. First, we introduce our hypergraph representation for sets Φ of subsumees in Subsection 4.1, and we show in Subsection 4.2 how it can be used to decide whether $\mathcal{T} \models C \sqsubseteq E$ holds for every $C \in \Phi$, where E is an $\mathcal{EL}$-concept and $\mathcal{T}$ is a normalised $\mathcal{EL}$-TBox. Alternatively, the method allows us to decide subsumptions of the form $\mathcal{T} \models \mathcal{C} \sqsubseteq E$, where $\mathcal{C}$ is an $\mathcal{ELU}_{\perp,\mu}$-concept. Subsequently, we turn to our graph representation of sets Ψ of subsumers in Subsection 4.3, and we present its application to deciding whether $\mathcal{T} \models E \sqsubseteq D$ holds for every $D \in \Psi$ in Subsection 4.4. Moreover, our method enables us to decide subsumptions of the form $\mathcal{T} \models E \sqsubseteq \mathcal{D}$, where $\mathcal{D}$ an $\mathcal{EL}_\nu$-concept. Finally, in Subsection 4.5 we combine these two methods to decide whether $\mathcal{T} \models C \sqsubseteq D$ holds for every $C \in \Phi$ and for every $D \in \Psi$, or subsumptions of the form $\mathcal{T} \models \mathcal{C} \sqsubseteq \mathcal{D}$.

4.1 Hypergraph Representation of Subsumee Sets

We now introduce our main hypergraph-based notion for representing sets of concepts.

Definition 18 (Concept Set Hypergraph). *A concept set hypergraph is a finite, labelled, directed, hypergraph $(\mathcal{V}, \mathcal{E}, \mathcal{L}, \mathcal{R})$ with a dedicated set of root nodes $\mathcal{R}$, where*

- $\mathcal{V}$ *is a finite, non-empty set of nodes;*
- $\mathcal{E} \subseteq 2^{\mathcal{V}} \times \mathcal{V}$ *is a set of directed hyperedges;*
- $\mathcal{L} \colon \mathcal{V} \cup \mathcal{E} \to 2^{\mathsf{N_C}} \cup \{\top\} \cup 2^{\mathsf{N_R}}$ *is a labelling function, mapping nodes $v \in \mathcal{V}$ to subsets $\mathcal{L}(v) \subseteq \mathsf{N_C} \cup \{\top\}$, and mapping edges $e \in \mathcal{E}$ to non-empty sets of role names $\mathcal{L}(e) \subseteq \mathsf{N_R}$;*
- $\mathcal{R} \subseteq \mathcal{V}$.

A concept set hypergraph is admissible *iff for every $v, v' \in \mathcal{V}$, $\mathcal{L}(v) = \mathcal{L}(v')$ implies $v = v'$.*

We use concept set hypergraphs to represent sets of subsumees or subsumers of fixpoints concepts, or of $\mathcal{EL}$-concepts w.r.t. a TBox. We now introduce the subsumee unfolding semantics of concept set hypergraphs.

Definition 19 ($\mathcal{EL}$-Subsumee Unfoldings of Concept Set Hypergraphs). *Let $\mathcal{G} = (\mathcal{V}, \mathcal{E}, \mathcal{L}, \mathcal{R})$ be a concept set hypergraph. First, let $\mathit{Unfold}_{\mathcal{G}}^{\leftarrow} \subseteq \mathcal{V} \times \mathcal{EL}$ be the smallest set closed under the following conditions:*

- *if $v \in \mathcal{V}$, $\varphi = A \in \mathcal{L}(v)$ or $\varphi = \top \in \mathcal{L}(v)$, then $(v, \varphi) \in \mathit{Unfold}_{\mathcal{G}}^{\leftarrow}$;*
- *if $e = (\{v_1, \ldots, v_n\}, v) \in \mathcal{E}$, $r \in \mathcal{L}(e)$, $(v_i, C_i) \in \mathit{Unfold}_{\mathcal{G}}^{\leftarrow}$ for every $1 \le i \le n$, then $(v, \exists r. \prod_{i=1}^{n} C_i) \in \mathit{Unfold}_{\mathcal{G}}^{\leftarrow}$.*

We set $\mathit{Unfold}_{\leftarrow}(\mathcal{G}) = \{ \prod_{v \in \mathcal{R}} C_v \mid (v, C_v) \in \mathit{Unfold}_{\mathcal{G}}^{\leftarrow} \}$.

Intuitively, under the subsumee unfolding semantics, hyperedges represent conjunctions and the node labels are disjunctively connected.

We now recall from [6] that we can compute dedicated concept set hypergraphs that serve as a representation of the subsumees of an $\mathcal{EL}$-concept w.r.t. a normalised $\mathcal{EL}$-TBox, or the subsumees of an $\mathcal{ELU}_{\perp,\mu}$-concept. In the following we write $S_1 \equiv S_2$, for two sets S_1, S_2 of $\mathcal{EL}$-concepts, to denote that for every concept $C_1 \in S_1$ there exists a concept $C_2 \in S_2$ with $\models C_1 \sqsubseteq C_2$, and that for every $D_2 \in S_2$ there exists $D_1 \in S_1$ with $\models D_2 \sqsubseteq D_1$.

The set of subsumees of an $\mathcal{EL}$-concept D w.r.t. an $\mathcal{EL}$-TBox $\mathcal{T}$, i.e. the set of all $\mathcal{EL}$-concepts E such that $\mathcal{T} \models E \sqsubseteq D$, can be represented in a dedicated concept set hypergraph $\exp_{\leftarrow}^{\mathcal{T}}(D)$ (termed *expansion hypergraph of D w.r.t. $\mathcal{T}$* in [6]).

Theorem 20 (See [6]). *Let $\mathcal{T}$ be a normalised $\mathcal{EL}$-TBox and let D be an $\mathcal{EL}$-concept. Then there exist a concept set hypergraph $\exp_{\leftarrow}^{\mathcal{T}}(D) = (\mathcal{V}, \mathcal{E}, \mathcal{L}, \mathcal{R})$, called* subsumee hypergraph of D *w.r.t. $\mathcal{T}$, which can be computed in exponential time in the size of D and $\mathcal{T}$ such that*

(i) $Unfold_{\leftarrow}(\exp_{\leftarrow}^{\mathcal{T}}(D)) \equiv \{ E \in \mathcal{EL} \mid \mathcal{T} \models E \sqsubseteq D \}$, and

(ii) for every $r \in \mathsf{N_R}$ and for every $e, e' \in \mathcal{E}$ with $r \in \mathcal{L}(e) \cap \mathcal{L}(e')$, it holds that $e = e'$.

Property (ii) states that for every node v and for every $r \in \mathsf{N_R}$ there exists at most one incoming hyperedge in $\exp_{\leftarrow}^{\mathcal{T}}(D)$ labelled with r. This property will be necessary to establish Theorem 27 below.

Alternatively, concept set hypergraphs can also represent the set $\mathrm{UF}_{\mathcal{EL}}(\mathcal{C})$ of unfoldings of a closed $\mathcal{ELU}_{\perp,\mu}$-concept $\mathcal{C}$. We use the fact that $\mathcal{C}$ is equivalent to the (potentially) infinite disjunction over its unfoldings contained in the set $\mathrm{UF}_{\mathcal{EL}}(\mathcal{C})$, and in particular, every unfolding in $\mathrm{UF}_{\mathcal{EL}}(\mathcal{C})$ is a subsumee of $\mathcal{C}$ (w.r.t. the empty TBox).

Theorem 21 (See [6]). *Let $\mathcal{C}$ be a closed $\mathcal{ELU}_{\perp,\mu}$-concept. Then there exist a concept set hypergraph $\exp_{\leftarrow}(\mathcal{C}) = (\mathcal{V}, \mathcal{E}, \mathcal{L}, \mathcal{R})$, called* subsumee hypergraph of $\mathcal{C}$, *which can be computed in exponential time in the size of $\mathcal{C}$ such that*

(i) $Unfold_{\leftarrow}(\exp_{\leftarrow}(\mathcal{C})) \equiv UF_{\mathcal{EL}}(\mathcal{C})$, and

(ii) for every $r \in \mathsf{N_R}$ and for every $e, e' \in \mathcal{E}$ with $r \in \mathcal{L}(e) \cap \mathcal{L}(e')$, it holds that $e = e'$.

The subsumee hypergraph $\exp_{\leftarrow}(\mathcal{C})$ of $\mathcal{C}$ was called *expansion hypergraph for $\mathcal{C}$* in [6].

As we are interested in sets of concepts using symbols from a given signature Σ only, we introduce the following operation on concept set hypergraphs to restrict the set of unfoldings to concepts formulated in $\mathcal{EL}_{\Sigma}$.

Definition 22 (Σ-Reduct). *Let $\mathcal{G} = (\mathcal{V}, \mathcal{E}, \mathcal{L}, \mathcal{R})$ be a concept set hypergraph and let Σ be a signature.*

We define the Σ-reduct of $\mathcal{G}$, denoted with $reduct_{\Sigma}(\mathcal{G})$, to be the concept set hypergraph $reduct_{\Sigma}(\mathcal{G}) := (\mathcal{V}_{\Sigma}, \mathcal{E}_{\Sigma}, \mathcal{L}_{\Sigma}, \mathcal{R}_{\Sigma})$, where

$$
\begin{aligned}
\mathcal{V}_{\Sigma} &:= \mathcal{V}, \\
\mathcal{E}_{\Sigma} &:= \{ e \in \mathcal{E} \mid (\mathcal{L}(e) \cap \Sigma) \neq \emptyset \}, \\
\mathcal{L}_{\Sigma} &:= \{ (v, S) \mid v \in \mathcal{V}, \ S = \mathcal{L}(v) \cap (\Sigma \cup \{\top\}) \} \\
&\quad \cup \{ (e, S) \mid e \in \mathcal{E}_{\Sigma}, \ S = \mathcal{L}(e) \cap \Sigma \}, \\
\mathcal{R}_{\Sigma} &:= \mathcal{R}.
\end{aligned}
$$

Theorem 23. *Let $\mathcal{T}$ be a normalised $\mathcal{EL}$-TBox and let $X \in \mathsf{N_C}$ be a concept name. Then for the concept set hypergraph $\exp_{\leftarrow}^{\mathcal{T},\Sigma}(X) := reduct_{\Sigma}(\exp_{\leftarrow}^{\mathcal{T}}(X))$ it holds that*

$$
Unfold_{\leftarrow}(\exp_{\leftarrow}^{\mathcal{T},\Sigma}(X)) \equiv \{ E \in \mathcal{EL}_{\Sigma} \mid \mathcal{T} \models E \sqsubseteq X \}.
$$

The concept set hypergraph $\exp_{\leftarrow}^{\mathcal{T},\Sigma}(X)$ can be computed in exponential time in the size of $\mathcal{T}$ and Σ.

4.2 Subsumee Set Subsumption w.r.t. TBoxes

We proceed with deciding whether $\mathcal{T} \models C \sqsubseteq E$ holds for every $C \in \Phi$, where $\mathcal{T}$ is a normalised $\mathcal{EL}$-TBox, $\Phi \subseteq \mathcal{EL}$ is a set of *foreign* subsumees, and E is an $\mathcal{EL}$-concept, by using the technique developed in [6], which we briefly recall here. Foreign subsumees are either subsumees of an
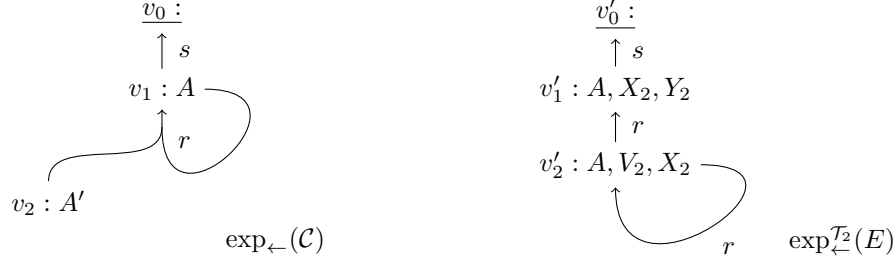
Figure 1: Purged subsumee hypergraphs for Example 26

$\mathcal{EL}$-concept w.r.t. an $\mathcal{EL}$-TBox, possibly different from E and $\mathcal{T}$, or the unfoldings of an $\mathcal{ELU}_{\perp,\mu}$-concept. Hence, by taking $\Phi = \mathrm{UF}_{\mathcal{EL}}(\mathcal{C})$, it becomes possible to decide subsumptions of the form $\mathcal{T} \models \mathcal{C} \sqsubseteq E$ for closed $\mathcal{ELU}_{\perp,\mu}$-concepts $\mathcal{C}$.

In [6] it was also established that every concept set hypergraph $\mathcal{G} = (\mathcal{V}, \mathcal{E}, \mathcal{L}, \mathcal{R})$ can be transformed into a *purged* concept set hypergraph $\mathcal{G}' = (\mathcal{V}', \mathcal{E}', \mathcal{L}', \mathcal{R}')$ (by removing superfluous nodes and hyperedges that do not produce $\mathcal{EL}$-concept unfoldings) such that $\mathrm{Unfold}_{\leftarrow}(\mathcal{G}) = \emptyset$ iff $|\mathcal{V}'| = 1$, $\mathcal{E}' = \emptyset$, $\mathcal{L}' = \{v \mapsto \emptyset\}$ for $\mathcal{V}' = \{v\}$, and $\mathcal{R}' = \mathcal{V}'$.

Theorem 24 (See [6])**.** *Let $\mathcal{G} = (\mathcal{V}, \mathcal{E}, \mathcal{L}, \mathcal{R})$ be a concept set hypergraph. Then there exists a concept set hypergraph $purge(\mathcal{G}) = (\mathcal{V}', \mathcal{E}', \mathcal{L}', \mathcal{R}')$ such that*

- *$\mathrm{Unfold}_{\leftarrow}(purge(\mathcal{G})) \equiv \mathrm{Unfold}_{\leftarrow}(\mathcal{G})$; and*

- *if $\mathrm{Unfold}_{\leftarrow}(\mathcal{G}) = \emptyset$, then $|\mathcal{V}'| = 1$, $\mathcal{E}' = \emptyset$, $\mathcal{L}' = \{v \mapsto \emptyset\}$ for $\mathcal{V}' = \{v\}$, $\mathcal{R}' = \mathcal{V}'$; and*

- *if $\mathrm{Unfold}_{\leftarrow}(\mathcal{G}) \neq \emptyset$, then for every $v \in \mathcal{V}'$ it holds that $(v, \varphi) \in \mathrm{Unfold}_{\mathcal{G}}^{\leftarrow}$ for some $\varphi \in \mathcal{EL}$.*

Given two purged concept set hypergraphs $\mathcal{G}_1$ and $\mathcal{G}_2$, we can decide whether for every $C' \in \mathrm{Unfold}_{\leftarrow}(\mathcal{G}_1)$ there exists $D' \in \mathrm{Unfold}_{\leftarrow}(\mathcal{G}_2)$ with $\models C' \sqsubseteq D'$ by trying to construct an *subsumee hypergraph simulation*, which is formally defined as follows.

Definition 25 (Subsumee Hypergraph Simulation)**.** *Let $\mathcal{G}_1 = (\mathcal{V}_1, \mathcal{E}_1, \mathcal{L}_1, \mathcal{R}_1)$ and let $\mathcal{G}_2 = (\mathcal{V}_2, \mathcal{E}_2, \mathcal{L}_2, \mathcal{R}_2)$ be two concept set hypergraphs. We say that $\mathcal{G}_1$ can be subsumee simulated by $\mathcal{G}_2$, written $sim_{\leftarrow}(\mathcal{G}_1, \mathcal{G}_2)$, iff there exists a binary relation $S \subseteq \mathcal{V}_1 \times \mathcal{V}_2$ which fulfills the following conditions:*

(i) *if $(v_1, v_2) \in S$ and $\top \notin \mathcal{L}_2(v_2)$, then $\mathcal{L}_1(v_1) \cup \{\top\} \subseteq \mathcal{L}_2(v_2)$;*

(ii) *if $(v_1, v_2) \in S$, $\top \notin \mathcal{L}_2(v_2)$, $e_1 = (H_1, v_1) \in \mathcal{E}_1$, and $r \in \mathcal{L}_1(e_1)$, then there exists $e_2 = (H_2, v_2) \in \mathcal{E}_2$ such that $r \in \mathcal{L}_2(e_2)$ and for every $v_2' \in H_2$ there exists $v_1' \in H_1$ with $(v_1', v_2') \in S$; and*

(iii) *for every $v_2 \in \mathcal{R}_2$ there exists $v_1 \in \mathcal{R}_1$ such that $(v_1, v_2) \in S$.*

S *is called a* a subsumee hypergraph simulation between $\mathcal{G}_1$ and $\mathcal{G}_2$.

For a detailed explanation of the simulation conditions we refer the reader to [5, 6, 11].

Example 26. *Let $\mathcal{T}_1$, $\mathcal{T}_2$ and Σ be defined as in Example 8. Additionally, let $\mathcal{C} = B_{\mathcal{T}_1}^{\Sigma}(X_1) = \mu x_1.(\exists s.\mu y_1.(A \sqcup \exists r.\mu v_1.(A' \sqcap y_1)))$ and let $E = \exists s.(X_2 \sqcup Y_2)$. Then $\mathcal{T}_2 \models \mathcal{C} \sqsubseteq E$. The purged concept set hypergraph $\exp_{\leftarrow}(\mathcal{C})$ for $\mathcal{C}$ and the purged concept set hypergraph $\exp_{\leftarrow}^{\mathcal{T}_2}(E)$ for E w.r.t. $\mathcal{T}_2$ are shown in Figure 1. We have that:*

103

$$Unfold_{\leftarrow}(\exp_{\leftarrow}(\mathcal{C})) = \{\exists s.A, \exists s.\exists r.(A' \sqcap A), \exists s.\exists r.(A' \sqcap \exists r.(A' \sqcap A)), \dots\},$$
$$Unfold_{\leftarrow}(\exp_{\leftarrow}^{\mathcal{T}_2}(E)) = \{\exists s.A, \exists s.X_2, \exists s.Y_2, \exists s.\exists r.A, \exists s.\exists r.V_2, \exists s.\exists r.X_2, \exists s.\exists r.\exists r.A, \dots\}.$$

Moreover, it holds that $S = \{(v_0, v_0'), (v_1, v_1'), (v_1, v_2')\}$ is a subsumee hypergraph simulation between $\exp_{\leftarrow}(\mathcal{C})$ and $\exp_{\leftarrow}^{\mathcal{T}_2}(E)$, i.e. $sim_{\leftarrow}(\exp_{\leftarrow}(\mathcal{C}), \exp_{\leftarrow}^{\mathcal{T}_2}(E))$.

Theorem 27. Let $\mathcal{G}_1 = (\mathcal{V}_1, \mathcal{E}_1, \mathcal{L}_1, \mathcal{R}_1)$ and let $\mathcal{G}_2 = (\mathcal{V}_2, \mathcal{E}_2, \mathcal{L}_2, \mathcal{R}_2)$ be two concept set hypergraphs. Then it holds that $sim_{\leftarrow}(purge(\mathcal{G}_1), purge(\mathcal{G}_2))$ iff for every $C \in Unfold_{\leftarrow}(\mathcal{G}_1)$ there exists $D \in Unfold_{\leftarrow}(\mathcal{G}_2)$ with $\models C \sqsubseteq D$.

We now recall the main result from [6], connecting subsumption with the existence of a subsumee hypergraph simulation between purged subsumee hypergraphs.

Corollary 28 (See [6]). Let $\mathcal{T}$ be a normalised $\mathcal{EL}$-TBox, let $\mathcal{C}$ be a closed $\mathcal{ELU}_{\perp,\mu}$-concept, and let E be an $\mathcal{EL}$-concept. Then it holds that:

$$\mathcal{T} \models \mathcal{C} \sqsubseteq E \quad iff \quad sim_{\leftarrow}(purge(\exp_{\leftarrow}(\mathcal{C})), purge(\exp_{\leftarrow}^{\mathcal{T}}(E)))$$

It can be decided in polynomial time w.r.t. the size of $\exp_{\leftarrow}(\mathcal{C})$ and $\exp_{\leftarrow}^{\mathcal{T}}(E)$ whether $purge(\exp_{\leftarrow}(\mathcal{C}))$ can be simulated by $purge(\exp_{\leftarrow}^{\mathcal{T}}(E))$. Consequently, the subsumption $\mathcal{T} \models \mathcal{C} \sqsubseteq E$ can be decided in exponential time in the size of $\mathcal{T}$, $\mathcal{C}$, and E.

4.3   Graph Representations of Subsumer Sets

We now introduce our graph-based representation for sets of $\mathcal{EL}$-concepts that stand for the subsumers of a concept name w.r.t. a normalised $\mathcal{EL}$-TBox, or of an $\mathcal{EL}_\nu$-concept.

Definition 29 (Concept Set Graph). A concept set graph is a concept set hypergraph $(\mathcal{V}, \mathcal{E}, \mathcal{L}, \mathcal{R})$ such that for every $(H, v) \in \mathcal{E}$ it holds that $|H| = 1$, $|\mathcal{R}| = 1$, and $\mathcal{L}(v) \subseteq \mathsf{N_C}$ for every $v \in \mathcal{V}$.

In the following we use a tuple $(\mathcal{V}, \mathcal{E}, \mathcal{L}, \rho)$ to denote a concept set graph $(\mathcal{V}, \mathcal{E}, \mathcal{L}, \mathcal{R})$ with $\mathcal{R} = \{\rho\}$. Moreover, for $(H, v') \in \mathcal{E}$ with $H = \{v\}$ we simply write $(v, v') \in \mathcal{E}$. Note that in a concept set graph $\top$ does not occur in node labels.

We now define the subsumer unfolding semantics for concept set graphs. In contrast to the subsumee unfolding semantics of concept set hypergraphs, under the subsumer unfolding semantics we view a concept set graph $\mathcal{G}$ as a disjunctive normal form representation of a set of $\mathcal{EL}$-concepts, denoted with $Unfold_{\rightarrow}(\mathcal{G})$. Concept names in node labels are conjunctively connected and source nodes in hyperedges are disjunctively connected. However, as hyperedges in concept set graphs are simple directed edges, any disjunction consists of only one disjunct.

Definition 30 ($\mathcal{EL}$-Subsumer Unfoldings of Concept Set Graphs). Let $\mathcal{G} = (\mathcal{V}, \mathcal{E}, \mathcal{L}, \rho)$ be a concept set graph. First, let $Unfold_{\mathcal{G}}^{\rightarrow} \subseteq \mathcal{V} \times \mathcal{EL}$ be the smallest set closed under the following conditions:

- if $v \in \mathcal{V}$, then $(v, \top) \in Unfold_{\mathcal{G}}^{\rightarrow}$;
- if $v \in \mathcal{V}$, $\{v' \mid (v, v') \in \mathcal{E}\} = \{v_1', \dots, v_m'\}$ with $m \geq 0$, $(v_i', C_i) \in Unfold_{\mathcal{G}}^{\rightarrow}$ for every $1 \leq i \leq m$, then $(v, \prod_{A \in \mathcal{L}(v)} A \sqcap \prod_{i=1}^{m} \prod_{r \in \mathcal{L}(v, v_i')} \exists r.C_i) \in Unfold_{\mathcal{G}}^{\rightarrow}$.

We set $Unfold_{\rightarrow}(\mathcal{G}) = \{C \mid (\rho, C) \in Unfold_{\mathcal{G}}^{\rightarrow}\}$.

Intuitively, the set $Unfold_{\rightarrow}(\mathcal{G})$ can be seen to represent a (potentially) infinite conjunction over the concepts contained in it. It holds that $Unfold_{\rightarrow}(\mathcal{G}) = \{\top\}$ iff $\mathcal{L}(\rho) = \emptyset$ and for all

$(v, v') \in \mathcal{E}$ it holds that $v \neq \rho$. Note that for our purposes simple directed edges (instead of complex hyperedges) are sufficient as we do not need disjunctions to represent sets of $\mathcal{EL}$-subsumers.

Similarly to Subsection 4.1, one can compute a dedicated concept set hypergraph to represent the set of subsumers of a concept name w.r.t. a normalised $\mathcal{EL}$-TBox.

Definition 31 (Subsumer Graph for a Concept Name w.r.t. a TBox). *Let $\mathcal{T}$ be a normalised $\mathcal{EL}$-TBox and let $X \in (\mathrm{sub}(\mathcal{T}) \cap \mathsf{N_C}) \cup \{\top\}$. The subsumer graph of X w.r.t. $\mathcal{T}$, denoted with $\exp^{\mathcal{T}}_{\rightarrow}(X)$, is the concept set graph $(\mathcal{V}, \mathcal{E}, \mathcal{L}, \rho)$, where*

$$\mathcal{V} := \{\, v_\varphi \mid \varphi \in (\mathrm{sub}(\mathcal{T}) \cap \mathsf{N_C}) \cup \{\top\} \,\},$$
$$\mathcal{E} := \{\, (v_\varphi, v_Y) \mid \exists r.Y \in \mathrm{sub}(\mathcal{T}), \mathcal{T} \models \varphi \sqsubseteq \exists r.Y \,\},$$
$$\mathcal{L} := \{\, (v_\varphi, S) \mid v_\varphi \in \mathcal{V}, S = \{\, Y \in \mathrm{sub}(\mathcal{T}) \cap \mathsf{N_C} \mid \mathcal{T} \models \varphi \sqsubseteq Y \,\} \,\}$$
$$\cup \{\, (e, S) \mid e = (v_\varphi, v_Y), S = \{\, r \mid \exists r.Y \in \mathrm{sub}(\mathcal{T}), \mathcal{T} \models \varphi \sqsubseteq \exists r.Y \,\} \,\},$$
$$\rho := v_X.$$

Note that in contrast to the subsumee hypergraph $\exp^{\mathcal{T}}_{\leftarrow}(X)$ of X w.r.t. $\mathcal{T}$, the subsumer graph $\exp^{\mathcal{T}}_{\rightarrow}(X)$ of X w.r.t. $\mathcal{T}$ can be computed in polynomial time in the size of $\mathcal{T}$.

In the following it will also be convenient to represent a concept set graph $\mathcal{G}$ in the form of a dedicated TBox $\mathcal{T}(\mathcal{G})$.

Definition 32 (TBox Encoding of Concept Set Graphs). *Let $\mathcal{G} = (\mathcal{V}, \mathcal{E}, \mathcal{L}, \rho)$ be a concept set graph. We define the subsumee TBox encoding $\mathcal{T}_{\leftarrow}(\mathcal{G})$ and the subsumer TBox encoding $\mathcal{T}_{\rightarrow}(\mathcal{G})$ of $\mathcal{G}$ as follows:*

$$\mathcal{T}_{\leftarrow}(\mathcal{G}) := \{\, \bigsqcap_{A \in \mathcal{L}(v)} A \sqcap \bigsqcap_{\substack{e = (v, v') \in \mathcal{E} \\ r \in \mathcal{L}(e)}} \exists r.X_{v'} \sqsubseteq X_v \mid v \in \mathcal{V} \,\},$$
$$\mathcal{T}_{\rightarrow}(\mathcal{G}) := \bigcup_{v \in \mathcal{V}} \big(\{\, X_v \sqsubseteq A \mid A \in \mathcal{L}(v) \,\} \cup \{\, X_v \sqsubseteq \exists r.X_{v'} \mid e = (v, v') \in \mathcal{E}, r \in \mathcal{L}(e) \,\}\big).$$

Finally, we set $\mathcal{T}(\mathcal{G}) := \mathcal{T}_{\leftarrow}(\mathcal{G}) \cup \mathcal{T}_{\rightarrow}(\mathcal{G})$.

We obtain the following properties of subsumer graphs $\exp^{\mathcal{T}}_{\rightarrow}(X)$ w.r.t. its subsumer unfoldings.

Theorem 33. *Let $\mathcal{T}$ be a normalised $\mathcal{EL}$-TBox and let $X \in (\mathrm{sub}(\mathcal{T}) \cap \mathsf{N_C}) \cup \{\top\}$. Additionally, let $\exp^{\mathcal{T}, \Sigma}_{\rightarrow}(X) := \mathrm{reduct}_\Sigma(\exp^{\mathcal{T}}_{\rightarrow}(X))$. Then it holds that:*

(i) $\mathrm{Unfold}_{\rightarrow}(\exp^{\mathcal{T}}_{\rightarrow}(X)) \equiv \{\, E \in \mathcal{EL} \mid \mathcal{T} \models X \sqsubseteq E \,\}$, and

(ii) $\mathrm{Unfold}_{\rightarrow}(\exp^{\mathcal{T}, \Sigma}_{\rightarrow}(X)) \equiv \{\, E \in \mathcal{EL}_\Sigma \mid \mathcal{T} \models X \sqsubseteq E \,\}$.

The concept set graph $\exp^{\mathcal{T}, \Sigma}_{\rightarrow}(X)$ can be computed in polynomial time in the size of $\mathcal{T}$ and Σ.

Similarly to the case of $\mathcal{ELU}_{\perp, \mu}$-concepts, given an $\mathcal{EL}_\nu$-concept $\mathcal{D}$, it is possible to construct a concept set graph $\exp_{\rightarrow}(\mathcal{D})$, called *subsumer graph for $\mathcal{D}$*, in polynomial time which represents the potentially infinitely many unfoldings of $\mathcal{D}$.

Theorem 34. *Let $\mathcal{D}$ be a closed $\mathcal{EL}_\nu$-concept. Then there exists a concept set graph $\exp_{\rightarrow}(\mathcal{D})$ such that*

$$\mathrm{Unfold}_{\rightarrow}(\exp_{\rightarrow}(\mathcal{D})) \equiv UF_{\mathcal{EL}}(\mathcal{D}).$$

The concept set graph $\exp_{\rightarrow}(\mathcal{D})$ can be computed in polynomial time in the size of $\mathcal{T}$ and Σ.
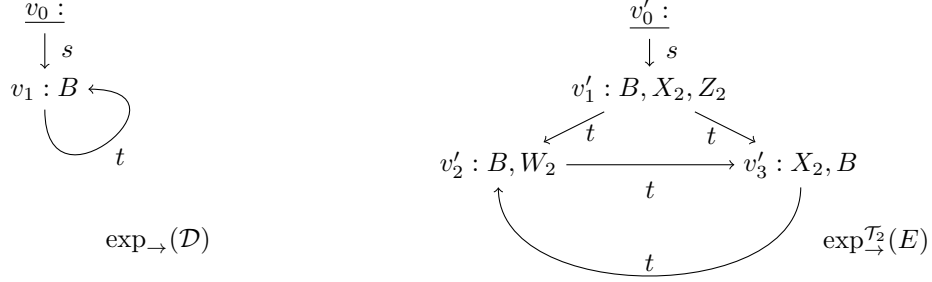
Figure 2: Subsumer graphs for Example 36

4.4   Subsumer Set Subsumption w.r.t. TBoxes

The problem whether $\mathcal{T} \models E \sqsubseteq D$ holds for every $D \in \Psi$, where $\mathcal{T}$ is a normalised $\mathcal{EL}$-TBox, E is an $\mathcal{EL}$-concept, and $\Psi \subseteq \mathcal{EL}$ is a set of subsumers, can be characterised in terms of the existence of a subsumer graph simulation between the concept set graphs representing Ψ and the subsumers of E w.r.t. $\mathcal{T}$. We employ the simulation notion for left-hand side witnesses developed in [9]. Similarly, by taking $\Psi = \mathrm{UF}_{\mathcal{EL}}(\mathcal{D})$, we can decide subsumptions of the form $\mathcal{T} \models E \sqsubseteq \mathcal{D}$ for closed $\mathcal{EL}_\nu$-concepts $\mathcal{D}$.

Definition 35 (Subsumer Graph Simulation). *Let $\mathcal{G}_1 = (\mathcal{V}_1, \mathcal{E}_1, \mathcal{L}_1, \rho_1)$ and $\mathcal{G}_2 = (\mathcal{V}_2, \mathcal{E}_2, \mathcal{L}_2, \rho_2)$ be two concept set graphs. Moreover, let $u_1 \in \mathcal{V}_1$ and $u_2 \in \mathcal{V}_2$.*
*   We say that u_2 subsumer simulates u_1, denoted with $sim_\rightarrow([\mathcal{G}_1, u_1], [\mathcal{G}_2, u_2])$, iff there exists a relation $S \subseteq \mathcal{V}_1 \times \mathcal{V}_2$ such that*

(i) if $(v_1, v_2) \in S$, then $\mathcal{L}_1(v_1) \subseteq \mathcal{L}_2(v_2)$;

(ii) if $(v_1, v_2) \in S$, $e_1 = (v_1, v_1') \in \mathcal{E}_1$, and $r \in \mathcal{L}_1(e_1)$, then there exists $e_2 = (v_2, v_2') \in \mathcal{E}_2$ such that $r \in \mathcal{L}_2(e_2)$ and $(v_1', v_2') \in S$; and

(iii) $(u_1, u_2) \in S$.

S is called a a subsumer graph simulation between $[\mathcal{G}_1, u_1]$ and $[\mathcal{G}_2, u_2]$. We write $sim_\rightarrow(\mathcal{G}_1, \mathcal{G}_2)$ iff $sim_\rightarrow([\mathcal{G}_1, \rho_1], [\mathcal{G}_2, \rho_2])$ holds.

For a detailed explanation of the simulation conditions we refer the reader again to [5,6,11].

Example 36. *Let $\mathcal{T}_1$, $\mathcal{T}_2$ and Σ be defined as in Example 8. Additionally, let $\mathcal{D} = F_{\mathcal{T}_1}^\Sigma(X_1) = \nu x_1.\exists s.\nu z_1.(B \sqcap \exists t.z_1)$ and let $E = \exists s.(X_2 \sqcap Z_2)$. Then $\mathcal{T}_2 \models E \sqsubseteq \mathcal{D}$. The subsumer graph $\exp_\rightarrow(\mathcal{D})$ for $\mathcal{D}$ and the subsumer graph $\exp_\rightarrow^{\mathcal{T}_2}(E)$ for E w.r.t. $\mathcal{T}_2$ are shown in Figure 2.*
*   We have that:*

$Unfold_\rightarrow(\exp_\rightarrow(\mathcal{D})) = \{\top, \exists s.\top, \exists s.(B \sqcap \exists t.\top), \exists s.(B \sqcap \exists t.(B \sqcap \exists t.\top)), \dots\},$

$Unfold_\rightarrow(\exp_\rightarrow^{\mathcal{T}_2}(E)) = \{\top, \exists s.\top, \exists s.(B \sqcap X_2 \sqcap Z_2 \sqcap (\exists t.\top) \sqcap (\exists t.\top)),$
$$\exists s.(B \sqcap X_2 \sqcap Z_2 \sqcap (\exists t.(W_2 \sqcap \exists t.\top)) \sqcap (\exists t.\top)), \dots\}.$$

Moreover, it holds that $S = \{(v_0, v_0'), (v_1, v_1'), (v_1, v_2'), (v_1, v_3')\}$ is a subsumer graph simulation between $\exp_\rightarrow(\mathcal{D})$ and $\exp_\rightarrow^{\mathcal{T}_2}(E)$, i.e. $sim_\rightarrow(\exp_\rightarrow(\mathcal{D}), \exp_\rightarrow^{\mathcal{T}_2}(E))$.

We obtain the following result.

Theorem 37. *Let $\mathcal{G}_1 = (\mathcal{V}_1, \mathcal{E}_1, \mathcal{L}_1, \rho_1)$ and let $\mathcal{G}_2 = (\mathcal{V}_2, \mathcal{E}_2, \mathcal{L}_2, \rho_2)$ be two concept set graphs. Then it holds that $sim_\rightarrow(\mathcal{G}_1, \mathcal{G}_2)$ iff for every $C \in Unfold_\rightarrow(\mathcal{G}_1)$ there exists $D \in Unfold_\rightarrow(\mathcal{G}_2)$ with $\models D \sqsubseteq C$.*

Corollary 38. *Let $\mathcal{T}$ be a normalised $\mathcal{EL}$-TBox, let E be an $\mathcal{EL}$-concept, and let $\mathcal{D}$ be a closed $\mathcal{EL}_\nu$-concept. Additionally, let $\mathcal{T}'$ be the normalisation of the TBox $\mathcal{T} \cup \{X_E \equiv E\}$ for a fresh concept name X_E. Then the following two statements are equivalent:*

(i) $\mathcal{T} \models E \sqsubseteq \mathcal{D}$;

(ii) $sim_\rightarrow(\exp_\rightarrow(\mathcal{D}), \exp_\rightarrow^{\mathcal{T}'}(X_E))$.

The subsumption $\mathcal{T} \models E \sqsubseteq \mathcal{D}$ can be decided in exponential time in the size of $\mathcal{T}$, E, and $\mathcal{D}$.

4.5 Subsumee–Subsumer Set Subsumption w.r.t. TBoxes

We now combine the previous methods for deciding whether $\mathcal{T} \models C \sqsubseteq D$ holds for every $C \in \Phi$ and for every $D \in \Psi$, where $\mathcal{T}$ is a normalised $\mathcal{EL}$-TBox, $\Phi \subseteq \mathcal{EL}$ is a set of subsumees and $\Psi \subseteq \mathcal{EL}$ is a set of subsumers. By setting $\Phi = \mathrm{UF}_{\mathcal{EL}}(\mathcal{C})$ and $\Psi = \mathrm{UF}_{\mathcal{EL}}(\mathcal{D})$ we obtain a decision procedure for subsumptions of the form $\mathcal{T} \models \mathcal{C} \sqsubseteq \mathcal{D}$, where $\mathcal{C}$ is an $\mathcal{ELU}_{\perp,\mu}$-concept and $\mathcal{D}$ is an $\mathcal{EL}_\nu$-concept.

Our procedure for deciding whether all the unfoldings of a subsumee hypergraph $\mathcal{G}_{\mathcal{C}}$ entail all the unfoldings of a subsumer graph $\mathcal{G}_{\mathcal{D}}$ w.r.t. a TBox $\mathcal{T}$ is based on the following observation regarding subsumptions on the concept level. If $\mathrm{Unfold}_\rightarrow(\mathcal{G}_{\mathcal{D}})$ is finite, it would be sufficient to build the conjunction $F_{\mathcal{G}_{\mathcal{D}}} := \bigsqcap_{\varphi \in \mathrm{Unfold}_\rightarrow(\mathcal{G}_{\mathcal{D}})} \varphi$ and to check whether $\mathcal{T} \models C \sqsubseteq F_{\mathcal{G}_{\mathcal{D}}}$ holds for every $C \in \mathrm{Unfold}_\leftarrow(\mathcal{G}_{\mathcal{C}})$. Hence, using the standard technique of encoding the concept $F_{\mathcal{G}_{\mathcal{D}}}$ in a TBox (by adding an axiom of the form $X_\rho \equiv F_{\mathcal{G}_{\mathcal{D}}}$ for a fresh concept name X_ρ), or, equivalently, encoding the graph $\mathcal{G}_{\mathcal{D}}$ in the TBox $\mathcal{T}_{\mathcal{G}_{\mathcal{D}}}^\leftarrow$, we obtain that for every $C \in \mathrm{Unfold}_\leftarrow(\mathcal{G}_{\mathcal{C}})$ and for every $D \in \mathrm{Unfold}_\rightarrow(\mathcal{G}_{\mathcal{D}})$, the subsumption $\mathcal{T} \models C \sqsubseteq D$ (†) is equivalent to

$$\mathcal{T} \cup \mathcal{T}_\leftarrow(\mathcal{G}_{\mathcal{D}}) \models C \sqsubseteq X_\rho$$

for every $C \in \mathrm{Unfold}_\leftarrow(\mathcal{G}_{\mathcal{C}})$, where ρ is the root node of $\mathcal{G}_{\mathcal{D}}$. In the case where $\mathrm{Unfold}_\rightarrow(\mathcal{G}_{\mathcal{D}})$ is finite, the subsumption (†) can thus be decided using the techniques developed in Section 4.2. If $\mathrm{Unfold}_\rightarrow(\mathcal{G}_{\mathcal{D}})$ is infinite, it is no longer to possible to construct the conjunction $F_{\mathcal{G}_{\mathcal{D}}}$. Our technique now relies on identifying appropriate concept names $X_{\eta,v}$ in $\mathcal{T}$, also called *hooks* in $\mathcal{T}$, for every cyclic node v in $\mathcal{G}_{\mathcal{D}}$. By adding the hooks in the form of axioms $X_{\eta,v} \sqsubseteq X_v$ to a TBox $\mathcal{T}_\eta$ for every cyclic node $v \in \mathcal{G}_{\mathcal{D}}$ we will establish that the subsumption (†) is equivalent to

$$\mathcal{T} \cup \mathcal{T}_\leftarrow(\mathcal{G}_{\mathcal{D}}) \cup \mathcal{T}_\eta \models C \sqsubseteq X_\rho$$

for every $C \in \mathrm{Unfold}_\leftarrow(\mathcal{G}_{\mathcal{C}})$, which can then be checked again via the techniques developed in Section 4.2. In the following we will provide a formal characterisation of the notion of a TBox hook and of the TBox $\mathcal{T}_\eta$.

We start with introducing the notion of cyclic nodes in concept set graphs.

Definition 39 (Cyclic Nodes). *Let $\mathcal{G} = (\mathcal{V}, \mathcal{E}, \mathcal{L}, \rho)$ be a concept set graph. We set $cyclic(\mathcal{G}) := \{ v \in \mathcal{V} \mid (v,v) \in \mathcal{E}^+ \}$, where $\mathcal{E}^+$ is the transitive closure of $\mathcal{E}$.*

The following theorem will form the basis of our decision procedure for Type-(iii) witnesses. It establishes that an unfolding D' of a subsumer graph $\mathcal{G}$ subsumes an $\mathcal{EL}$-concept C w.r.t. an $\mathcal{EL}$-TBox $\mathcal{T}$ iff the concept name X_ρ corresponding to the root node of $\mathcal{G}$ subsumes C w.r.t. the union of $\mathcal{T}$, the subsumee TBox encoding $\mathcal{T}(\mathcal{G}(\mathcal{D}))$ of $\mathcal{G}$ and the *hook TBox $\mathcal{T}_\eta$*. The TBox

$\mathcal{T}_\eta$ contains axioms of the form $\varphi \sqsubseteq X_v$ for every cyclic node v in $\mathcal{G}$, where X_v represents the encoding of v in $\mathcal{T}(\mathcal{G}(\mathcal{D}))$ and φ ranges over all the concept names in $\mathcal{T}$ or $\varphi = \top$ such that $[\exp^{\mathcal{T}}_\rightarrow(\varphi), v_\varphi]$ subsumer simulates $[\mathcal{G}, v]$. Recall that $\mathcal{T}(\mathcal{G}) = \mathcal{T}_\leftarrow(\mathcal{G}) \cup \mathcal{T}_\rightarrow(\mathcal{G})$.

Theorem 40. *Let $\mathcal{T}$ be a normalised $\mathcal{EL}$-TBox and let $\mathcal{G} = (\mathcal{V}, \mathcal{E}, \mathcal{L}, \rho)$ be a concept set graph. Then for the mapping $\eta\colon cyclic(\mathcal{G}) \to 2^{(\mathrm{sub}(\mathcal{T}) \cap \mathsf{N_C}) \cup \{\top\}}$ defined for every $v \in cyclic(\mathcal{G})$ as*

$$\eta(v) := \{\, \varphi \in (\mathrm{sub}(\mathcal{T}) \cap \mathsf{N_C}) \cup \{\top\} \mid sim_\rightarrow([\mathcal{G}, v], [\exp^{\mathcal{T}}_\rightarrow(\varphi), v_\varphi]) \,\},$$

it holds for every $C \in \mathcal{EL}$ that the following two statements are equivalent:

(i) $\mathcal{T} \models C \sqsubseteq D'$ *for every* $D' \in \mathit{Unfold}_\rightarrow(\mathcal{G})$;

(ii) $\mathcal{T} \cup \mathcal{T}(\mathcal{G}) \cup \mathcal{T}_\eta \models C \sqsubseteq X_\rho$, *where*

$$\mathcal{T}_\eta := \{\, \varphi \sqsubseteq X_v \mid v \in cyclic(\mathcal{G}),\ \varphi \in \eta(v) \,\}.$$

Intuitively, the correctness of the previous theorem follows from the fact that the concept names X_v corresponding to $v \in cyclic(\mathcal{G})$ actually represent greatest fixpoints, but a model of $\mathcal{T}(\mathcal{G})$ interprets every node X_v as an *arbitrary* fixpoint, i.e. not necessarily as a greatest fixpoint. The purpose of the TBox is $\mathcal{T}_\eta$ is now to ensure that the considered models interpret the concepts X_v as fixpoints that are greater (subset-wise) than any compatible fixpoint concept name present in $\mathcal{T}$. The compatibility of fixpoint concept names in $\mathcal{T}$ is established via the existence of subsumer simulations.

More formally, the direction $(i) \Rightarrow (ii)$ follows by using the characterisation of the subsumption $\mathcal{T} \models C \sqsubseteq D'$, for every $D' \in \mathit{Unfold}_\rightarrow(\mathcal{G})$ using a subsumer simulation. The TBox $\mathcal{T}_\eta$ contains all possible simulation partners for cyclic nodes v in $\mathcal{G}$, and thus a simulation can be found for every subsumee $C \in \mathcal{EL}$. For the direction $(ii) \Rightarrow (i)$ it suffices to interpret every concept name $X_v \in cyclic(\mathcal{G})$ as its respective greatest fixpoint.

By combining the previous theorem with theorems 20 and 27, we obtain the following main result.

Corollary 41. *Let $\mathcal{T}$ be a normalised $\mathcal{EL}$-TBox, let $\mathcal{G_C}$ be a concept set hypergraph and let $\mathcal{G_D}$ be a concept set graph. Then for the mapping $\eta\colon cyclic(\mathcal{G}) \to 2^{(\mathrm{sub}(\mathcal{T}) \cap \mathsf{N_C}) \cup \{\top\}}$ defined for every $v \in cyclic(\mathcal{G})$ as*

$$\eta(v) := \{\, \varphi \in (\mathrm{sub}(\mathcal{T}) \cap \mathsf{N_C}) \cup \{\top\} \mid sim_\rightarrow([\mathcal{G}, v], [\exp^{\mathcal{T}}_\rightarrow(\varphi), v_\varphi]) \,\},$$

it holds that the following two statements are equivalent:

(i) $\mathcal{T} \models C' \sqsubseteq D'$ *for every* $C' \in \mathit{Unfold}_\leftarrow(\mathcal{G_C})$ *and for every* $D' \in \mathit{Unfold}_\rightarrow(\mathcal{G_D})$;

(ii) $\mathcal{T}_{intp} \models C' \sqsubseteq X_\rho$ *for every* $C' \in \mathit{Unfold}_\leftarrow(\mathcal{G_C})$;

(iii) $sim_\leftarrow(purge(\mathcal{G_C}), purge(\exp^{\mathcal{T}_{intp}}_\leftarrow(X_\rho)))$;

where $\mathcal{T}_{intp} := \mathcal{T} \cup \mathcal{T}(\mathcal{G_D}) \cup \mathcal{T}_\eta$ *and* $\mathcal{T}_\eta := \{\, \varphi \sqsubseteq X_v \mid v \in cyclic(\mathcal{G}),\ \varphi \in \eta(v) \,\}$.

The equivalence of statements (i) and (ii) follows from Theorem 40, and the equivalence of statements (ii) and (iii) is entailed by theorems 20 and 27.

Example 42. *Let $\mathcal{T}_1$, $\mathcal{T}_2$, and Σ be defined as in Example 8. Additionally, let $\mathcal{C} = B^\Sigma_{\mathcal{T}_1}(X_1) = \mu x_1.(\exists s.\mu y_1.(A \sqcup \exists r.\mu v_1.(A' \sqcap y_1)))$ and let $\mathcal{D} = F^\Sigma_{\mathcal{T}_1}(X_1) = \nu x_1.\exists s.\nu z_1.(B \sqcap \exists t.z_1)$. Finally,*

let $\mathcal{G}_\mathcal{C}$ and $\mathcal{G}_\mathcal{D}$ be defined as in Figures 1 and 2, respectively. Then $cyclic(\mathcal{G}_\mathcal{D}) = \{v_1\}$ and for $\eta = \{v_1 \mapsto \{W_2, X_2, Y_2, Z_2\}\}$, we obtain $\mathcal{T}_{intp} = \mathcal{T}_2 \cup \mathcal{T}_\leftarrow(\mathcal{G}_\mathcal{D}) \cup \mathcal{T}_\rightarrow(\mathcal{G}_\mathcal{D}) \cup \mathcal{T}_\eta$, where

$$\mathcal{T}_\leftarrow(\mathcal{G}_\mathcal{D}) = \{\exists s.X_{v_1} \sqsubseteq X_{v_0}, B \sqcap \exists t.X_{v_1} \sqsubseteq X_{v_1}\},$$
$$\mathcal{T}_\rightarrow(\mathcal{G}_\mathcal{D}) = \{X_{v_0} \sqsubseteq \exists s.X_{v_1}, X_{v_1} \sqsubseteq B, X_{v_1} \sqsubseteq \exists t.X_{v_1}\},$$
$$\mathcal{T}_\eta = \{W_2 \sqsubseteq X_{v_1}, X_2 \sqsubseteq X_{v_1}, Y_2 \sqsubseteq X_{v_1}, Z_2 \sqsubseteq X_{v_1}\}.$$

It holds that $\mathcal{T}_{intp} \models C' \sqsubseteq X_{v_0}$ for every $C' \in Unfold(\mathcal{G}_\mathcal{C})$, where v_0 is the root of $\mathcal{G}_\mathcal{D}$. Note that concept set hypergraph $\mathcal{G}_\mathcal{C}$ is already purged (that is $purge(\mathcal{G}_\mathcal{C}) = \mathcal{G}_\mathcal{C}$, cf. Theorem 24). The subsumee hypergraph $\exp^{\mathcal{T}_{intp}}_\leftarrow(X_{v_0})$ can be constructed as in Theorem 20. It can then be verified that $\mathcal{G}_\mathcal{C}$ can be subsumee simulated by $purge(\exp^{\mathcal{T}_{intp}}_\leftarrow(X_{v_0}))$ (cf. Definition 25).

5 Deciding the Existence of Difference Witnesses

We are now ready to formulate an algorithm for deciding whether a concept name is a Type-(iii) witness using the techniques developed in the previous section.

Given two normalised $\mathcal{EL}$-TBoxes $\mathcal{T}_1$, $\mathcal{T}_2$, a signature Σ, and a concept name $X \notin \Sigma$, our procedure for deciding whether X is a Type-(iii) witness is now shown in Algorithm 1. The algorithm operates as follows. First, the concept set hypergraph $\mathcal{B}_\Sigma$ is computed by purging the Σ-reduct of the subsumee hypergraph $\exp^{\mathcal{T}_1}_\leftarrow(X)$ for X w.r.t. $\mathcal{T}_1$. Then, it is checked whether $\mathcal{B}_\Sigma$ corresponds to $\bot$ (cf. Theorem 24), in which case X is not a Type-(iii) witness and the algorithm returns '**no**'. Subsequently, the concept set graph $\mathcal{F}_\Sigma$ is computed by purging the Σ-reduct of the subsumer graph $\exp^{\mathcal{T}_1}_\rightarrow(X)$ for X w.r.t. $\mathcal{T}_1$. If $\mathcal{F}_\Sigma$ is equivalent to $\top$ (cf. Definition 30), one can infer that X is not a Type-(iii) witness and the algorithm returns '**no**'. Otherwise, the cyclic nodes in $\mathcal{F}_\Sigma$ are determined, and after initialising the "fixpoint hook" TBox $\mathcal{T}_\eta$ with $\emptyset$, the algorithm iterates over all the cyclic nodes v_i in the concept set graph $\mathcal{F}_\Sigma$. First the hook TBox $\mathcal{T}_{v_i}$ for the node v_i is initialised with $\emptyset$ before the algorithm iterates over every concept name φ contained in the signature of $\mathcal{T}_2$ and over $\varphi = \top$. In each inner iteration the subsumer graph $\mathcal{F}^{\mathcal{T}_2}_\varphi$ of φ is first computed before it is checked whether the node x_φ in $\mathcal{F}^{\mathcal{T}_2}_\varphi$ can subsumer simulate the cyclic node v_i in $\mathcal{F}_\Sigma$. If a simulation can be found, φ is a valid hook for v_i in $\mathcal{T}_2$ and the axiom $\varphi \sqsubseteq X_{v_i}$ is added to $\mathcal{T}_{v_i}$. After all the concept names $\varphi \in \mathsf{sig}(\mathcal{T}_2)$ have been checked, the hooks for the cyclic node v_i contained in the TBox $\mathcal{T}_{v_i}$ are added to the global fixpoint hook TBox $\mathcal{T}_\eta$. Subsequently, after the search for hooks for every cyclic node in $\mathcal{F}_\Sigma$ has been completed, the TBox $\mathcal{T}_{intp}$ is computed by taking the union of $\mathcal{T}_2$ with the TBox encoding $\mathcal{T}(\mathcal{F}_\Sigma)$ of $\mathcal{F}_\Sigma$ and $\mathcal{T}_\eta$. Next, the concept set hypergraph $\mathcal{G}_{intp}$ is computed by purging the subsumee hypergraph for X_ρ w.r.t. $\mathcal{T}_2$. Note that the concept name X_ρ represents the root node of the graph $\mathcal{F}_\Sigma$. If the hypergraph $\mathcal{B}_\Sigma$ cannot be backward simulated by $\mathcal{G}_{intp}$, we can conclude that X is a Type-(iii) witness and the algorithm returns '**yes**'. Otherwise, it holds that X is not a Type-(iii) witness and '**no**' is returned.

We obtain the following theorem which establishes the correctness of Algorithm 1.

Theorem 43. *Let $\mathcal{T}_1$, $\mathcal{T}_2$ be normalised $\mathcal{EL}$-TBoxes, and let $X \in (\mathsf{sig}(\mathcal{T}_1) \cap \mathsf{N_C}) \setminus \Sigma$ be a concept name. Then Algorithm 1 applied on $\mathcal{T}_1$, $\mathcal{T}_2$, Σ, and X returns '**yes**' iff there exist two $\mathcal{EL}_\Sigma$-concepts C and D such that $\mathcal{T}_1 \models C \sqsubseteq X$, $\mathcal{T}_1 \models X \sqsubseteq D$ and $\mathcal{T}_2 \not\models C \sqsubseteq D$. Algorithm 1 runs in exponential time in the size of $\mathcal{T}_1$, $\mathcal{T}_2$, and Σ.*

The correctness of Algorithm 1 follows from Corollary 41.

Regarding computational complexity, we observe that subsumee hypergraphs for an $\mathcal{EL}$-TBox $\mathcal{T}$ and a concept name X can be computed in exponential time in the size of $\mathcal{T}$ and Σ,

---

Algorithm 1 Deciding Type-(iii) Witnesses

Input: Normalised $\mathcal{EL}$-TBoxes $\mathcal{T}_1$ and $\mathcal{T}_2$, signature Σ, concept name $X \in (\mathsf{sig}(\mathcal{T}_1) \cap \mathsf{N_C}) \setminus \Sigma$
Output: 'yes' or 'no'
 1: Let $\mathcal{B}_\Sigma := \mathrm{purge}(\mathrm{reduct}_\Sigma(\mathrm{exp}_{\leftarrow}^{\mathcal{T}_1}(X)))$
 2: **if** $\mathcal{B}_\Sigma \equiv \bot$ **then**
 3:     **return no**
 4: **end if**
 5: Let $\mathcal{F}_\Sigma = (\mathcal{V}, \mathcal{E}, \mathcal{L}, \rho) := \mathrm{reduct}_\Sigma(\mathrm{exp}_{\rightarrow}^{\mathcal{T}_1}(X))$
 6: **if** $\mathcal{F}_\Sigma \equiv \top$ **then**
 7:     **return no**
 8: **end if**
 9: Let $\{v_1, \ldots, v_\ell\} := \mathrm{cyclic}(\mathcal{F}_\Sigma)$
10: Let $\mathcal{T}_\eta := \emptyset$
11: **for** $i = 1$ to ℓ **do**
12:     Let $\mathcal{T}_{v_i} := \emptyset$
13:     **for every** $\varphi \in (\mathrm{sub}(\mathcal{T}_2) \cap \mathsf{N_C}) \cup \{\top\}$ **do**
14:         Let $\mathcal{F}_\varphi^{\mathcal{T}_2} := \mathrm{exp}_{\rightarrow}^{\mathcal{T}_2}(\varphi)$
15:         **if not** $\mathrm{sim}_{\rightarrow}([\mathcal{F}_\Sigma, v_i], [\mathcal{F}_\varphi^{\mathcal{T}_2}, v_\varphi])$ **then**
16:             **continue**
17:         **end if**
18:         $\mathcal{T}_{v_i} := \mathcal{T}_{v_i} \cup \{\varphi \sqsubseteq X_{v_i}\}$
19:     **end for**
20:     **if** $\mathcal{T}_{v_i} = \emptyset$ **then**
21:         **return** 'yes'
22:     **end if**
23:     $\mathcal{T}_\eta := \mathcal{T}_\eta \cup \mathcal{T}_{v_i}$
24: **end for**
25: Let $\mathcal{T}_{\mathrm{intp}} := \mathcal{T}_2 \cup \mathcal{T}(\mathcal{F}_\Sigma) \cup \mathcal{T}_\eta$
26: Let $\mathcal{G}_{\mathrm{intp}} := \mathrm{purge}(\mathrm{exp}_{\leftarrow}^{\mathcal{T}_{\mathrm{intp}}}(X_\rho))$
27: **if not** $\mathrm{sim}_{\leftarrow}(\mathcal{B}_\Sigma, \mathcal{G}_{\mathrm{intp}})$ **then**
28:     **return** 'yes'
29: **end if**
30: **return** 'no'

---

whereas the subsumer graph for X w.r.t. $\mathcal{T}$ can be constructed in polynomial time in the size of $\mathcal{T}$ and Σ. As the computation of the Σ-reduct and as the purging can each be computed in polynomial time in the size of the input, we obtain that the subsumee hypergraph $\mathcal{B}_\Sigma$ can be computed in exponential time in the size of $\mathcal{T}_1$ and Σ. Furthermore, the subsumer graph $\mathcal{F}_\Sigma$ is of polynomial size w.r.t. $\mathcal{T}_1$ and Σ, i.e. the number ℓ of cyclic nodes is polynomial in the size of $\mathcal{T}_1$ and Σ. Consequently, the for-loops will be executed polynomially many times in the size of $\mathcal{T}_1$, $\mathcal{T}_2$, and Σ, and $\mathcal{T}_\eta$ is also of polynomial size w.r.t. $\mathcal{T}_1$, $\mathcal{T}_2$, and Σ. We can infer that the subsumee hypergraph $\mathcal{G}_{\mathrm{intp}}$ can be computed in exponential time in the size of $\mathcal{T}_1$, $\mathcal{T}_2$, and Σ. As checking for the existence of simulations can be performed in polynomial time in the size of the two input graphs, we obtain that in the worst case Algorithm 1 hence runs in exponential time in the size of $\mathcal{T}_1$, $\mathcal{T}_2$, and Σ.

6 Conclusion

We have presented foundations for a hypergraph-based approach for deciding the existence of a logical difference between general $\mathcal{EL}$-TBoxes $\mathcal{T}_1$ and $\mathcal{T}_2$ w.r.t. a signature Σ, which can be reduced to checking for the existence of four types of difference witnesses. In this paper we focused on Type-(iii) witnesses, i.e. concept names $X \in \mathsf{sig}(\mathcal{T}_1) \setminus \Sigma$ for which there exists $C \in \mathcal{EL}_\Sigma$ and $D \in \mathcal{EL}_\Sigma$ with $\mathcal{T}_1 \models C \sqsubseteq X$ and $\mathcal{T}_1 \models X \sqsubseteq D$ but $\mathcal{T}_2 \not\models C \sqsubseteq D$. Our procedure for deciding the existence of Type-(iii) witnesses is based on (hyper)graph representations of sets of such concepts C and D combined with simulations between such (hyper)graphs to characterise subsumption. The same technique can be used to decide subsumptions between least and greatest fixpoint concepts w.r.t. TBoxes.

It will be necessary to evaluate our method on versions of ontologies that are used in practice. Moreover, in order to minimise the computational effort required to obtain the hypergraph representation, the sets of subsumees might need to be represented in disjunctive normal form, contrary to the current conjunctive normal form version from [6].

Finally, there appear to be several ways to continue this line of research. Based on the notions developed in this paper, one could extend the results to devise algorithms for reasoning with more general fixpoint concepts. One could also investigate possible applications of our hypergraph-based approach to the problem of deciding the existence of uniform interpolants of $\mathcal{EL}$-TBoxes. A uniform interpolant of an $\mathcal{EL}$-TBox $\mathcal{T}$ w.r.t. a signature Σ is a TBox $\mathcal{T}'$ such that $\mathsf{sig}(\mathcal{T}') \subseteq \Sigma$ and $\mathsf{cDiff}_\Sigma(\mathcal{T}, \mathcal{T}') = \mathsf{cDiff}_\Sigma(\mathcal{T}', \mathcal{T}) = \emptyset$.

References

[1] Baader, F., Brandt, S., Lutz, C.: Pushing the EL envelope. In: Proceedings of the 19th International Joint Conference on Artificial Intelligence (IJCAI-05). Morgan-Kaufmann Publishers, Edinburgh, UK (2005)

[2] Baader, F., Calvanese, D., McGuinness, D.L., Nardi, D., Patel-Schneider, P.F. (eds.): The description logic handbook: theory, implementation, and applications. Cambridge University Press, 2 edn. (June 2010)

[3] Brandt, S.: Polynomial time reasoning in a description logic with existential restrictions, GCI axioms, and—what else? In: Proceedings of the 16th European Conference on Artificial Intelligence (ECAI-2004). pp. 298–302. IOS Press (2004)

[4] De Giacomo, G., Lenzerini, M.: A uniform framework for concept definitions in description logics. Journal of Artificial Intelligence Research (JAIR) 6, 87–110 (1997)

[5] Ecke, A., Ludwig, M., Walther, D.: The concept difference for EL-terminologies using hypergraphs. In: Proceedings of the International workshop on (Document) Changes: modeling, detection, storage and visualization (DChanges 2013). CEUR-WS, vol. 1008 (2013)

[6] Feng, S., Ludwig, M., Walther, D.: Deciding subsumers of least fixpoint concepts w.r.t. EL-TBoxes. In: Proceedings of the 38th edition of the German Conference on Artificial Intelligence (KI-15). CEUR workshop proceedings (2015)

[7] Feng, S., Ludwig, M., Walther, D.: The logical difference for EL: from terminologies towards TBoxes. In: Proceedings of the 1st International Workshop on Semantic Technologies (IWOST-15). CEUR workshop proceedings (2015)

[8] Hofmann, M.: Proof-theoretic approach to description-logic. In: Proceedings of the 20th Annual IEEE Symposium on Logic in Computer Science (LICS-05). pp. 229–237. IEEE Computer Society, Washington, DC, USA (2005)

[9] Konev, B., Ludwig, M., Walther, D., Wolter, F.: The logical difference for the lightweight description logic EL. Journal of Artificial Intelligence Research (JAIR) 44, 633–708 (2012)

[10] Konev, B., Walther, D., Wolter, F.: The logical difference problem for description logic terminologies. In: Proceedings of the 4th International Joint Conference on Automated Reasoning (IJCAR-08). pp. 259–274. Springer-Verlag, Berlin, Heidelberg (2008)

[11] Ludwig, M., Walther, D.: The logical difference for ELHr-terminologies using hypergraphs. In: Proceedings of the 21st European Conference on Artifical Intelligence (ECAI 2014). Frontiers in Artificial Intelligence and Applications, vol. 263, pp. 555–560. IOS Press (2014)

[12] Lutz, C., Wolter, F.: Deciding inseparability and conservative extensions in the description logic EL. Journal of Symbolic Computation 45(2), 194–228 (Feb 2010)