Pushing the $\mathcal{EL}$ Envelope Further

Franz Baader[1], Sebastian Brandt[2], Carsten Lutz[1]

[1]: TU Dresden [2]: Univ. of Manchester

Abstract. We extend the description logic $\mathcal{EL}^{++}$ with reflexive roles and range restrictions, and show that subsumption remains tractable if a certain syntactic restriction is adopted. We also show that subsumption becomes PSPACE-hard (resp. undecidable) if this restriction is weakened (resp. dropped). Additionally, we prove that tractability is lost when symmetric roles are added: in this case, subsumption becomes EXPTIME-hard.

1 Introduction

The W3C recommendation OWL is currently being revisited by a W3C working group with the goal of refining and extending the existing version OWL 1.0 [1]. Although the main aim is to produce a new version of the OWL language with the working title OWL 1.1, the group is also discussing a number of prominent fragments of OWL that may or may not become part of the upcoming 1.1 W3C recommendation. These fragments trade expressive power for favourable properties that are not shared by the full OWL language. In particular, several fragments are a subset of OWL DL, the description logic (DL) dialect of OWL, and aim at high efficiency for reasoning tasks such as subsumption, classfication, and satisfiability.

A notable fragment of this kind is the description logic $\mathcal{EL}^{++}$, which has been introduced in [3]. The advantage of $\mathcal{EL}^{++}$ is that it combines tractability of the afore mentioned reasoning problems with expressive power that is sufficient for many important applications of ontologies. In particular, $\mathcal{EL}^{++}$ is well-suited for the design of life science ontologies, and many existing such ontologies are formulated in this language. Examples include SNOMED CT [22], the Gene ontology [24], and large parts of GALEN [19]. As witnesses by publications such as [20, 23], serious ontology projects are currently picking up on it—including commercial ones such as SNOMED CT. Tractability of reasoning in $\mathcal{EL}^{++}$, which has been established in [3], is in stark contrast to the NEXPTIME-completeness (and thus high intractability) of reasoning in full OWL DL 1.0. Moreover, the CEL system [6] and the publications [5, 7] have demonstrated that reasoning in $\mathcal{EL}^{++}$ can be implemented in an extremely efficient way. Other reasoners are currently under construction or already available. For example, the Terminology Development Environment (TDE) of Apelon Corporation includes a reasoner for a fragment of $\mathcal{EL}^{++}$.

The basis of $\mathcal{EL}^{++}$ is the description logic $\mathcal{EL}$, which provides as concept constructors the top concept (i.e., owl:Thing), conjunction (i.e., objectIntersectionOf), and existential restriction (i.e., objectSomeValuesFrom). In $\mathcal{EL}^{++}$, this

somewhat frugal list is extended with a number of constructors such as the bottom concept (i.e., owl:Nothing) and nominals (i.e., objectOneOf with a single argument), to name only two. Additional features of $\mathcal{EL}^{++}$ include GCIs (i.e., unrestricted subClassOf) and complex role inclusions, which allow to express role hierarchies (i.e., subObjectPropertyExpression), transitive roles, and right identities. A complete description of $\mathcal{EL}^{++}$ is given in Section 2, and, in a more OWL-ish way, at [2].

One of the most prominent constructors *not* included in $\mathcal{EL}^{++}$ is universal value restriction (i.e., objectAllValuesFrom). The reason for omitting it is that, as shown in [3], universal value restrictions cause ExpTime-completeness of reasoning already when added to $\mathcal{EL}$. In the current paper, we extend $\mathcal{EL}^{++}$ with reflexive roles (i.e., reflexiveObjectProperty) and range restrictions (i.e. objectPropertyRange), a very important special case of universal value restrictions. This extension allows to capture additional ontologies in $\mathcal{EL}^{++}$ such as (certain versions of) the thesaurus of the US national cancer institute (NCI), which is intended to become the reference terminology for cancer research [21]. In this paper, we prove that reasoning in the extended version of $\mathcal{EL}^{++}$ remains tractable if a certain and rather natural syntactic condition is adopted. In particular, this restriction is satisfied by the NCI thesaurus. We also investigate the effects of weakening and dropping the restriction, showing that this leads to PSpace-hardness and undecidability of reasoning, respectively. Finally, we investigate the option of adding symmetric roles to $\mathcal{EL}^{++}$ and show that already in $\mathcal{EL}$ with symmetric roles, reasoning is ExpTime-complete. Based on the approach in this paper, the CEL reasoner is currently being extended to support both range restrictions and symmetric roles, see http://lat.inf.tu-dresden.de/systems/cel.

2 Introducing $\mathcal{EL}^{++}$

In $\mathcal{EL}^{++}$, *concepts* are inductively defined from a set $\mathsf{N_C}$ of *concept names*, a set $\mathsf{N_R}$ of *role names*, and a set $\mathsf{N_I}$ of *individual names*, using the constructors shown in the top five rows of Table 1. As usual, we use C and D to refer to concepts, r to refer to a role name, and a and b to refer to individual names. The semantics of $\mathcal{EL}^{++}$-concepts is defined in terms of an *interpretation* $\mathcal{I} = (\Delta^{\mathcal{I}}, \cdot^{\mathcal{I}})$. The *domain* $\Delta^{\mathcal{I}}$ is a non-empty set of individuals and the *interpretation function* $\cdot^{\mathcal{I}}$ maps each concept name $A \in \mathsf{N_C}$ to a subset $A^{\mathcal{I}}$ of $\Delta^{\mathcal{I}}$, each role name $r \in \mathsf{N_R}$ to a binary relation $r^{\mathcal{I}}$ on $\Delta^{\mathcal{I}}$, and each individual name $a \in \mathsf{N_I}$ to an individual $a^{\mathcal{I}} \in \Delta^{\mathcal{I}}$. The extension of $\cdot^{\mathcal{I}}$ to arbitrary concept descriptions is inductively defined as shown in the third column of Table 1.

The logic $\mathcal{EL}^{++}$ can be parameterized by one or more concrete domains $\mathcal{D}_1, \ldots, \mathcal{D}_n$, which correspond to datatypes in OWL and permit reference to concrete data objects such as strings and integers. Formally, a *concrete domain* $\mathcal{D}$ is a pair $(\Delta^{\mathcal{D}}, \mathcal{P}^{\mathcal{D}})$ with $\Delta^{\mathcal{D}}$ a set and $\mathcal{P}^{\mathcal{D}}$ a set of *predicate names*. Each $p \in \mathcal{P}$ is associated with an arity $n > 0$ and an extension $p^{\mathcal{D}} \subseteq (\Delta^{\mathcal{D}})^n$. To provide a link between the DL and the concrete domains $\mathcal{D}_1, \ldots, \mathcal{D}_n$, we introduce a set of *feature names* $\mathsf{N_F}$ and the concrete domain constructor shown in Table 1. We use

Name	Syntax	Semantics
top	$\top$	$\Delta^{\mathcal{I}}$
bottom	$\bot$	$\emptyset$
nominal	$\{a\}$	$\{a^{\mathcal{I}}\}$
conjunction	$C \sqcap D$	$C^{\mathcal{I}} \cap D^{\mathcal{I}}$
existential restriction	$\exists r.C$	$\{x \in \Delta^{\mathcal{I}} \mid \exists y \in \Delta^{\mathcal{I}} : (x,y) \in r^{\mathcal{I}} \wedge y \in C^{\mathcal{I}}\}$
concrete domain	$p(f_1,\ldots,f_k)$ for $p \in \mathcal{P}^{\mathcal{D}_j}$	$\{x \in \Delta^{\mathcal{I}} \mid \exists y_1,\ldots,y_k \in \Delta^{\mathcal{D}_j} :$ $f_i^{\mathcal{I}}(x) = y_i$ for $1 \leq i \leq k \ \wedge (y_1,\ldots,y_k) \in p^{\mathcal{D}_j}\}$
GCI	$C \sqsubseteq D$	$C^{\mathcal{I}} \subseteq D^{\mathcal{I}}$
RI	$r_1 \circ \cdots \circ r_k \sqsubseteq r$	$r_1^{\mathcal{I}} \circ \cdots \circ r_k^{\mathcal{I}} \subseteq r^{\mathcal{I}}$
domain restriction	$\mathsf{dom}(r) \sqsubseteq C$	$r^{\mathcal{I}} \subseteq C^{\mathcal{I}} \times \Delta^{\mathcal{I}}$
range restriction	$\mathsf{ran}(r) \sqsubseteq C$	$r^{\mathcal{I}} \subseteq \Delta^{\mathcal{I}} \times C^{\mathcal{I}}$
concept assertion	$C(a)$	$a^{\mathcal{I}} \in C^{\mathcal{I}}$
role assertion	$r(a,b)$	$(a^{\mathcal{I}}, b^{\mathcal{I}}) \in r^{\mathcal{I}}$

Table 1. Syntax and semantics of $\mathcal{EL}^{++}$.

p to denotes a predicate of a concrete domain and $f_1,\ldots,f_k$ to denote feature names. The interpretation function is required to map each feature name f to a partial function from $\Delta^{\mathcal{I}}$ to $\bigcup_{1\leq i\leq n} \Delta^{\mathcal{D}_i}$. Moreover, we generally assume that $\Delta^{\mathcal{D}_i} \cap \Delta^{\mathcal{D}_j} = \emptyset$ for $1 \leq i < j \leq n$.

An $\mathcal{EL}^{++}$ knowledge base comprises two sets, the TBox and the ABox. While the TBox contains intensional knowledge defining the main notions relevant to the domain of discourse, the ABox contains extensional knowledge about individual objects in the domain. Formally, a *TBox* is a finite set of *constraints*, which can be *general concept inclusions (GCIs)*, *role inclusions (RIs)*, *domain restrictions (DRs)* and *range restrictions (RRs)*. All of them are shown in Table 1. In RIs, we allow the case where $k = 0$, written as $\varepsilon \sqsubseteq r$. An *ABox* is a finite set of *concept assertions* and *role assertions*, which are also shown in Table 1. An interpretation $\mathcal{I}$ is a *model* of a TBox $\mathcal{T}$ (resp. ABox $\mathcal{A}$) if, for each constraint (resp. assertion) it contains, the conditions given in the third column of Table 1 are satisfied.

Regarding the expressive power of $\mathcal{EL}^{++}$, we remark the following. First, our RIs generalize several means of expressivity important in ontology applications:

- *role hierarchies* $r \sqsubseteq s$ can be expressed as $r \sqsubseteq s$;
- *role equivalences* $r \equiv s$ can be expressed as $r \sqsubseteq s, s \sqsubseteq r$;
- *transitive roles* can be expressed as $r \circ r \sqsubseteq r$;
- *reflexive roles* can be expressed as $\varepsilon \sqsubseteq r$;
- *left-identity rules* can be expressed as $r \circ s \sqsubseteq s$;
- *right-identity rules* can be expressed as $r \circ s \sqsubseteq r$.

Second, the bottom concept in combination with GCIs can be used to express *disjointness* of complex concept descriptions: $C \sqcap D \sqsubseteq \bot$ says that C, D are disjoint. Finally, the *identity* of two individuals can be expressed as $\{a\} \sqsubseteq \{b\}$, and their *distinctness* as $\{a\} \sqcap \{b\} \sqsubseteq \bot$. We remark that the version of $\mathcal{EL}^{++}$ presented here extends the original version in [3] by range restrictions and reflexive roles, i.e., role inclusions of the form $\varepsilon \sqsubseteq r$.

In this paper, we are interested in the reasoning task of *subsumption* because it allows to compute the classification of a TBox, i.e., a hierarchy which shows the subconcept-superconcept relations between the concepts defined in a TBox. We say that a concept C *subsumes* a concept D w.r.t. a TBox $\mathcal{T}$, written $C \sqsubseteq_{\mathcal{T}} D$, if $C^{\mathcal{I}} \subseteq D^{\mathcal{I}}$ in every model $\mathcal{I}$ of $\mathcal{T}$. As shown in [3], subsumption in $\mathcal{EL}^{++}$ is polynomially inter-reducible with many other reasoning tasks such as concept satisfiability, ABox consistency, and instance checking.

3 A Syntactic Restriction

To avoid intractability (and even undecidability), we have to impose a restriction on the structure of TBoxes that prevents the otherwise too intricate interplay of range restrictions and role inclusions. For a TBox $\mathcal{T}$ and role names r, s, we write $\mathcal{T} \models r \sqsubseteq s$ iff $r = s$ or $\mathcal{T}$ contains role inclusions

$$r_1 \sqsubseteq r_2, \ldots, r_{n-1} \sqsubseteq r_n \text{ with } r = r_1 \text{ and } s = r_n.$$

We write $\mathcal{T} \models \mathsf{ran}(r) \sqsubseteq C$ if there is a role name s with $\mathcal{T} \models r \sqsubseteq s$ and $\mathsf{ran}(s) \sqsubseteq C \in \mathcal{T}$. Now, the mentioned restriction is as follows:

> If $r_1 \circ \cdots \circ r_n \sqsubseteq s \in \mathcal{T}$ with $n \geq 1$ and $\mathcal{T} \models \mathsf{ran}(s) \sqsubseteq C$, then $\mathcal{T} \models \mathsf{ran}(r_n) \sqsubseteq C$.

The restriction ensures that if a role inclusion $r_1 \circ \cdots \circ r_n \sqsubseteq s$, $n \geq 1$, implies a role relationship $(d, e) \in s^{\mathcal{I}}$, then the RR on s do not impose new concept memberships of e. Note that the condition is vacuously true if the role inclusion is a reflexivity statement, a role hierarchy statement, a transitivity statement, or a generalized left-identity of the form $r_1 \circ \cdots \circ r_k \sqsubseteq r_k$. In the following, we assume without further notice that this restriction is satisfied. We will show in Section 5 that dropping the restriction results in undecidability, and even if only right identities are allowed as role inclusions it still leads to PSPACE-hardness and thus intractability.

4 Deciding Subsumption in Polytime

We show that, in the version of $\mathcal{EL}^{++}$ presented in this paper, subsumption can be decided in polytime. In fact, the extension with RIs of the form $\varepsilon \sqsubseteq r$ is technically a minor one, and it is easy to see how to extend the polytime algorithm for subsumption in the original version of $\mathcal{EL}^{++}$ given in [3] to account

for them. In contrast, the extension with range restrictions is less trivial. We show that range restrictions can be eliminated in quadratic time in a way such that all (non)-subsumptions are preserved. It then remains to apply the subsumption algorithm for the original version of $\mathcal{EL}^{++}$. As a preliminary, we convert TBoxes into a normal form.

4.1 A Normal Form for TBoxes

Given a TBox $\mathcal{T}$ formulated in $\mathcal{EL}^{++}$, we use $\mathsf{BC}_\mathcal{T}$ to denote the set of *basic concept descriptions in* $\mathcal{T}$, i.e., the smallest set of concept descriptions that contains the top concept $\top$, all concept names used in $\mathcal{T}$, and all concepts of the form $\{a\}$ and $p(f_1, \ldots, f_k)$ appearing in $\mathcal{T}$, possibly as subconcepts.

Definition 1 (Normal Form for TBoxes). *A TBox $\mathcal{T}$ is in* normal form *if*

1. *all concept inclusions have one of the following forms, where $C_1, C_2 \in \mathsf{BC}_\mathcal{T}$ and $D \in \mathsf{BC}_\mathcal{T} \cup \{\bot\}$: $C_1 \sqcap \cdots \sqcap C_n \sqsubseteq D$, $C_1 \sqsubseteq \exists r.C_2$, and $\exists r.C_1 \sqsubseteq D$*
2. *for all role inclusions $r_1 \circ \cdots \circ r_k \sqsubseteq r \in \mathcal{T}$, we have $k \leq 2$;*
3. *there are no domain restrictions and all range restrictions are of the form $\mathsf{ran}(r) \sqsubseteq A$, where A is a concept name.*

By introducing new concept and role names, any TBox $\mathcal{T}$ can be turned into a normalized TBox $\mathcal{T}'$ such that every model of $\mathcal{T}'$ is also a model of $\mathcal{T}$, and every model of $\mathcal{T}$ can be extended to a model of $\mathcal{T}'$ by appropriate choice of the interpretations of the additional concept and role names. The transformation can be performed in linear time. More details can be found in [3].

4.2 Eliminating Range Restrictions

Let $\mathcal{T}$ be a TBox in normal form. For each role name r, we use $\mathsf{ran}_\mathcal{T}(r)$ to denote the set of concept names A such that $\mathcal{T} \models \mathsf{ran}(r) \sqsubseteq A$. To eliminate range restrictions, we introduce a fresh concept name $X_{r,D}$ for every GCI $C \sqsubseteq \exists r.D$ in $\mathcal{T}$. Intuitively, $X_{r,D}$ denotes the range of r intersected with the extension of D. Now, let $\mathcal{T}'$ be the TBox obtained from $\mathcal{T}$ by dropping all range restrictions and, additionally, performing the following operations:

1. exchange every CI $C \sqsubseteq \exists r.D$ with the CIs $C \sqsubseteq \exists r.X_{r,D}$, $X_{r,D} \sqsubseteq D$, and $X_{r,D} \sqsubseteq A$ for all $A \in \mathsf{ran}_\mathcal{T}(r)$;
2. if $\varepsilon \sqsubseteq r \in \mathcal{T}$, then add the CI $\top \sqsubseteq A$ for all $A \in \mathsf{ran}_\mathcal{T}(r)$;

Then we have the following.

Lemma 1. *For all concept names A, B occurring in $\mathcal{T}$, $A \sqsubseteq_\mathcal{T} B$ iff $A \sqsubseteq_{\mathcal{T}'} B$.*

Proof. The "$\Leftarrow$" direction is trivial since every model $\mathcal{I}$ of $\mathcal{T}$ can be extended to a model of $\mathcal{T}'$ by interpreting every fresh concept name $X_{r,D}$ as $\{d \in D^\mathcal{I} \mid \exists e : (e, d) \in r^\mathcal{I}\}$. Thus, we concentrate on the "$\Rightarrow$" direction.

We show the contrapositive. Let $A_0 \not\sqsubseteq_{\mathcal{T}'} B_0$. Then there is a model $\mathcal{I}'$ of $\mathcal{T}'$ such that $A_0^{\mathcal{I}'} \setminus B_0^{\mathcal{I}'} \neq \emptyset$. $\mathcal{I}'$ is not necessarily a model of $\mathcal{T}$ since it need not

satisfy the RRs in $\mathcal{T}$. To fix this problem, we remove r-edges from $\mathcal{I}'$ that violate the range restrictions. The resulting interpretation turns out to be a model of $\mathcal{T}$. More specifically, convert $\mathcal{I}'$ into a new interpretation $\mathcal{I}$ by changing the interpretation of all role names r as follows:

$$r^{\mathcal{I}} = \{(d,e) \in r^{\mathcal{I}'} \mid e \in \bigcap_{A \in \mathsf{ran}_{\mathcal{T}}(r)} A^{\mathcal{I}'}\}.$$

To show that $\mathcal{I}$ is a model of $\mathcal{T}$, we only consider those constraints in $\mathcal{T}$ that could possibly be influenced by the modification:

- Let $C \sqsubseteq \exists r.D \in \mathcal{T}$. Then $\mathcal{T}'$ contains the CIs $C \sqsubseteq \exists r.X_{r,D}$ and $X_{r,D} \sqsubseteq D$. Let $d \in C^{\mathcal{I}}$. Since $C \in \mathsf{BC}_{\mathcal{T}}$, we get $d \in C^{\mathcal{I}'} \subseteq (\exists r.X_{r,D})^{\mathcal{I}'}$. Thus there is an $e \in \Delta^{\mathcal{I}'}$ with $(d,e) \in r^{\mathcal{I}'}$ and $e \in X_{r,D}^{\mathcal{I}'} \subseteq D^{\mathcal{I}'}$. Since $D \in \mathsf{BC}_{\mathcal{T}}$, $e \in D^{\mathcal{I}}$. Since there is a CI $X_{r,D} \sqsubseteq A$ in $\mathcal{T}'$ for all $A \in \mathsf{ran}_{\mathcal{T}}(r)$, we have $e \in A^{\mathcal{I}'}$ for all these A. By definition of $\mathcal{I}$, this together with $(d,e) \in r^{\mathcal{I}'}$ yields $(d,e) \in r^{\mathcal{I}}$. It follows that $d \in (\exists r.D)^{\mathcal{I}}$ as required.
- Let $\exists r.C \sqsubseteq D \in \mathcal{T}$. Then this CI is also in $\mathcal{T}'$. Let $d \in (\exists r.C)^{\mathcal{I}}$. Then $d \in (\exists r.C)^{\mathcal{I}'}$ and thus $d \in D^{\mathcal{I}'} = D^{\mathcal{I}}$.
- Let $\varepsilon \sqsubseteq r \in \mathcal{T}$. Then this constraint is also in $\mathcal{T}'$. Let $d \in \Delta^{\mathcal{I}}$. Then $(d,d) \in r^{\mathcal{I}'}$. In $\mathcal{T}'$, there is a CI $\top \sqsubseteq A$ for all $A \in \mathsf{ran}_{\mathcal{T}}(r)$. Thus, we have $d \in A^{\mathcal{I}'}$ for all these A and the definition of $\mathcal{I}$ yields $(d,d) \in r^{\mathcal{I}}$.
- Let $s \sqsubseteq r \in \mathcal{T}$. Then this RI is also in $\mathcal{T}'$ and thus $s^{\mathcal{I}'} \subseteq r^{\mathcal{I}'}$. Since $s \sqsubseteq r \in \mathcal{T}$, we have $\mathsf{ran}_{\mathcal{T}}(r) \subseteq \mathsf{ran}_{\mathcal{T}}(s)$. By definition of $\mathcal{I}$, this together with $s^{\mathcal{I}'} \subseteq r^{\mathcal{I}'}$ yields $s^{\mathcal{I}} \subseteq r^{\mathcal{I}}$.
- Let $r_1 \circ r_2 \sqsubseteq r \in \mathcal{T}$. Then this RI is also in $\mathcal{T}'$. Let $(d,d') \in r_1^{\mathcal{I}}$ and $(d',d'') \in r_2^{\mathcal{I}}$. Then $(d,d') \in r_1^{\mathcal{I}'}$ and $(d',d'') \in r_2^{\mathcal{I}'}$, and thus $(d,d'') \in r^{\mathcal{I}'}$. Since $r_1 \circ r_2 \sqsubseteq r \in \mathcal{T}$, the syntactic restriction on $\mathcal{T}$ ensures that $\mathsf{ran}_{\mathcal{T}}(r) \subseteq \mathsf{ran}_{\mathcal{T}}(r_2)$. By definition of $\mathcal{I}$, this together with $(d',d'') \in r_2^{\mathcal{I}}$ and $(d,d'') \in r^{\mathcal{I}'}$ implies $(d,d'') \in r^{\mathcal{I}}$.
- Let $\mathsf{ran}(r) \sqsubseteq A \in \mathcal{T}$ and $(d,e) \in r^{\mathcal{I}}$. By definition of $\mathcal{I}$, $e \in A^{\mathcal{I}'} = A^{\mathcal{I}}$.

Since $A_0^{\mathcal{I}} = A_0^{\mathcal{I}'}$ and $B_0^{\mathcal{I}} = B_0^{\mathcal{I}'}$, $\mathcal{I}'$ is thus a countermodel against $A_0 \sqsubseteq_{\mathcal{T}} B_0$. ❑

Observe that eliminating range restrictions induces only a quadratic blowup in the size of the original TBox. We expect that, in practical cases, the blowup will actually only be linear.

5 More on Range Restrictions and Role Inclusions

We show that dropping the syntactic restriction adopted in $\mathcal{EL}^{++}$ leads to undecidability of subsumption, even if we allow only the concept constructors of $\mathcal{EL}$, i.e., conjunction and existential restriction. If we restrict further by admitting only right identities as role inclusions, subsumption is still at least PSPACE-hard. This shows that weakening the restriction by excluding right identities (which play an important role in medical knowledge bases) does not recover polytime reasoning. The PSPACE lower bound applies even if a well-known acyclicity condition on role inclusions is adopted.

5.1 Undecidability

We consider $\mathcal{EL}$ with RIs and range restrictions as the only constraints in TBoxes (no GCIs!). The proof is by reduction of the emptiness problem of the intersection of two context-free grammars, which is known to be undecidable [14]. Recall that a context free grammar is a tuple (Σ, N, P, S) with Σ a finite alphabet of *terminal symbols*, N a finite set of *non-terminal symbols*, $S \in N$ a *start symbol*, and $P \subseteq N \times (\Sigma \cup N)^*$ a finite set of *productions*. We denote the lanuage generated by a grammar G with $L(G)$.

Let $G = (\Sigma, N, P, S)$ and $G' = (\Sigma, N', P', S')$ be two context-free grammars. W.l.o.g., we may assume that $N \cap N' = \emptyset$. We show how to translate G and G' into a TBox $\mathcal{T}$ and concepts C and D such that $L(G) \cap L(G') = \emptyset$ iff $C \not\sqsubseteq_{\mathcal{T}} D$. In the TBox $\mathcal{T}$, we use role names r_x and r'_x for every $x \in \Sigma \cup N \cup N'$ and two concept names A and B. More precisely, we set

$$C = \bigsqcap_{a \in \Sigma} \exists r'_a.\top \text{ and } D = \exists r_S.(A \sqcap B),$$

and $\mathcal{T}$ contains the following constraints:

1. the RR $\mathsf{ran}(r'_a) \sqsubseteq \bigsqcap_{b \in \Sigma} \exists r_b.\top$ and $\mathsf{ran}(r_a) \sqsubseteq \bigsqcap_{b \in \Sigma} \exists r_b.\top$ for all $a \in \Sigma$;
2. the RIs $r_{x_1} \circ \cdots \circ r_{x_n} \sqsubseteq r_v$ and $r'_{x_1} \circ r_{x_2} \circ \cdots \circ r_{x_n} \sqsubseteq r'_v$ for every production $v \vdash x_1 \cdots x_n \in P \cup P'$,
3. the range restrictions $\mathsf{ran}(r'_S) \sqsubseteq A$ and $\mathsf{ran}(r'_{S'}) \sqsubseteq B$.

To understand the construction, let d be an instance of C in some model $\mathcal{I}$ of $\mathcal{T}$. The definition of C and the RRs in Points 1 ensure that, for every word $w = a_1 \cdots a_n \in \Sigma^*$, there is an element d_w in $\mathcal{I}$ reachable along the path $r'_{a_1} r_{a_2} \cdots r_{a_n}$ from d. The RIs in Point 2 ensure that if $w \in L(G)$, then $(d, d_w) \in {r'_S}^{\mathcal{I}}$, and likewise for G' and $r'_{S'}$. We use the role r'_S here instead of r_S to distinguish, at the elements d_w, incoming role edges that originate in d from edges originating elsewhere. The RIs in Point 3 simply mark those d_w with $w \in L(G)$ with A, and those with $w \in L(G')$ with B. It is now easy to show that $L(G) \cap L(G') = \emptyset$ iff $C \not\sqsubseteq_{\mathcal{T}} D$. We remark that similar reductions have been used in [15, 17, 16].

Theorem 1. *Subsumption in $\mathcal{EL}$ extended with role inclusions and range restrictions is undecidable.*

In [12], an acyclicity condition on role inclusions is introduced, which is expected to become part of OWL 1.1. It follows from results in [12] that unrestricted $\mathcal{EL}^{++}$ becomes decidable when this condition is adopted. As shown in the subsequent section, however, acyclicity does not suffice to guarantee tractability.

5.2 PSPACE-hardness

We show that even if RIs are restricted to acyclic right identities, subsumption in $\mathcal{EL}$ with range restrictions is at least PSPACE-hard. Although it is possible to establish the result without any GCIs, we include them to improve readability.

The proof uses acyclic role inclusions. It is by reduction of the validity of QBFs, i.e., formulas of the form $\psi = Q_1 x_1 \cdots Q_n x_n.\varphi$ where $Q_i \in \{\exists, \forall\}$ for $1 \leq i \leq n$ and φ is a propositional formula. We refer to [18] for a formal definition of QBFs and their validity. Let a QBF $\psi = Q_1 x_1 \cdots Q_n x_n.\varphi$ be given, with φ in negation normal form. In the following, we assemble a TBox $\mathcal{T}$ such that, for certain concept names A and B, we have $A \sqsubseteq_{\mathcal{T}} B$ iff ψ is valid.

We start with building up a tree of depth n rooted in A. Roles $r_1, \ldots, r_n$ are used to connect left successor in the tree, and roles $\overline{r}_1, \ldots, \overline{r}_n$ for right successors. We use concept names L_i, $1 \leq i \leq n$, to mark the different levels of the tree:

$$A \sqsubseteq L_0$$
$$L_i \sqsubseteq \exists r_{i+1}.L_{i+1} \sqcap \exists \overline{r}_{i+1}.L_{i+1} \text{ for } 1 \leq i < n$$

Intuitively, nodes in the tree represent partial truth assignments. An incoming r_i edge means that the variable x_i is true, and an incoming $\overline{r}_i$ edge that it is false. We ensure that, once we have decided on the truth value of a variable, we keep it in all descendant nodes. For $1 \leq i < j \leq n$:

$$r_i \circ r_j \sqsubseteq r_i \qquad \overline{r}_i \circ r_j \sqsubseteq \overline{r}_i$$
$$r_i \circ \overline{r}_j \sqsubseteq r_i \qquad \overline{r_i} \circ \overline{r}_j \sqsubseteq \overline{r}_i$$

Clearly, the leafs of the tree represent all possible (non-partial) truth assignments. For what is to follow, we represent these assignments not only in terms of incoming edges, but also by concept names. We use a concept name T_i to indicate that x_i is true, and F_i for false. For $1 \leq i \leq n$:

$$\mathsf{ran}(r_i) = T_i \qquad \mathsf{ran}(\overline{r}_i) = F_i$$

In each leaf, we evaluate the formula. To this end, we introduce a concept name A_θ for every subformula θ of φ. Then put:

$$L_n \sqcap T_i \sqsubseteq A_{x_i} \text{ for } 1 \leq i \leq n$$
$$L_n \sqcap F_i \sqsubseteq A_{\neg x_i} \text{ for } 1 \leq i \leq n$$
$$L_n \sqcap A_\chi \sqcap A_\theta \sqsubseteq A_{\chi \wedge \theta} \text{ for all subformulas } \chi \wedge \theta \text{ of } \phi$$
$$L_n \sqcap A_\chi \text{ and } L_n \sqcap A_\theta \sqsubseteq A_{\chi \vee \theta} \text{ for all subformulas } \chi \vee \theta \text{ of } \phi$$

To evaluate the QBF ψ, we proceed from the leafs to the root. Each level corresponds to a quantifier in ψ, and we distinguish the case of an existential quantifier from that of a universal quantifier:

$$L_i \sqcap \exists r_{i+1}.B \text{ and } L_i \sqcap \exists \overline{r}_{i+1}.B \sqsubseteq B \text{ for } 0 \leq i < n \text{ with } Q_{i+1} = \exists$$
$$L_i \sqcap \exists r_{i+1}.B \sqcap \exists \overline{r}_{i+1}.B \sqsubseteq B \text{ for } 0 \leq i < n \text{ with } Q_{i+1} = \forall$$
$$L_n \sqcap A_\varphi \sqsubseteq B$$

It is not hard to check that, as intended, ψ is valid iff $A \sqsubseteq_{\mathcal{T}} B$. It is easily verified that the right identities in the proof are acyclic in the sense of [12].

Theorem 2. *Subsumption in $\mathcal{EL}$ extended with range restrictions and acyclic right identities is* PSPACE*-hard.*

The best known upper bound for the considered version of $\mathcal{EL}$ is an EXPTIME one, and it only applies to acyclic right identities [13]. For the cyclic case, even decidability is unknown (when RIs are restricted to right identities).

6 Symmetric and Inverse Roles

$\mathcal{EL}^{++}$ provides for reflexive and transitive roles, but not for symmetric ones. The purpose of the current section is to explain why this is the case: adding symmetric roles leads to ExpTime-hardness already for $\mathcal{EL}$ with GCIs. This result is established in two steps. First, we consider $\mathcal{ELI}$, i.e., $\mathcal{EL}$ extended with inverse roles which come in the form of existential restrictions $\exists r^-.C$ interpreted as $\{d \in \Delta^{\mathcal{I}} \mid \exists e \in C^{\mathcal{I}} : (e,d) \in r^{\mathcal{I}}\}$. We show that, in $\mathcal{ELI}$ with GCIs, subsumption is ExpTime-hard. This fills a gap in [3], where only PSpace-hardness is established. In a second step, we reduce subsumption in $\mathcal{ELI}$ with GCIs to subsumption in $\mathcal{EL}$ with symmetric roles and GCIs.

The ExpTime-lower bound for $\mathcal{ELI}$ with GCIs is proved by a reduction of the word problem of polynomially space-bounded alternating Turing machines. Details are given in the appendix. A corresponding upper bound is derived from the DL $\mathcal{ALCI}$ [10].

Theorem 3. *In $\mathcal{ELI}$, subsumption w.r.t. GCIs is* ExpTime*-complete.*

Let $\mathcal{EL}^{\mathsf{sym}}$ be the extension of $\mathcal{EL}$ with symmetric roles, i.e., there is a countably infinite subset $\mathsf{N}_\mathsf{R}^{\mathsf{sym}} \subseteq \mathsf{N}_\mathsf{R}$ such that $r^{\mathcal{I}} = \{(y,x) \mid (x,y) \in r^{\mathcal{I}}\}$ for all $r \in \mathsf{N}_\mathsf{R}^{\mathsf{sym}}$. To show ExpTime-hardness of subsumption in $\mathcal{EL}^{\mathsf{sym}}$ with GCIs, we reduce subsumption in $\mathcal{ELI}$ with GCIs using a small trick due to Halpern and Moses [11]. Let $\mathcal{T}$ be an $\mathcal{ELI}$ TBox containing GCIs as the only kind of constraint and using role names $r_1, \ldots, r_k$. We introduce additional role names $s_1, \ldots, s_k$ and assume that $r_i, s_i \in \mathsf{N}_\mathsf{R}^{\mathsf{sym}}$, for all $i \in \{1, \ldots, k\}$. Then we modify $\mathcal{T}$ into a new TBox $\mathcal{T}'$ by replacing

- every concept $\exists r_i.C$ with $\exists r_i \exists s_i.C$ and
- every concept $\exists r_i^-.C$ with $\exists s_i \exists r_i.C$.

It is not hard to see that for any two concept names A and B, we have $A \sqsubseteq_{\mathcal{T}} B$ in $\mathcal{ELI}$ iff $A \sqsubseteq_{\mathcal{T}'} B$ in $\mathcal{EL}^{\mathsf{sym}}$. We have thus established ExpTime-hardness. A corresponding upper bound is obtained from the obvious reduction of subsumption in $\mathcal{EL}^{\mathsf{sym}}$ with GCIs to satisfiability in Converse-PDL [9].

Theorem 4. *In $\mathcal{EL}^{\mathsf{sym}}$, subsumption w.r.t. general TBoxes is* ExpTime*-complete.*

Acknowledgements. We are grateful to Meng Suntisrivaraporn for discussions and suggestions.

References

1. W3C working group on OWL. http://www.w3.org/2007/OWL/wiki/OWL_Working_Group.
2. W3C working group webpage on $\mathcal{EL}^{++}$. http://www.w3.org/2007/OWL/wiki/EL.
3. F. Baader, S. Brandt, and C. Lutz. Pushing the $\mathcal{EL}$ envelope. In *Proceedings of IJCAI'05*, pages 364–369. Professional Book Center, 2005.
4. F. Baader, S. Brandt, and C. Lutz. Pushing the $\mathcal{EL}$ envelope further, 2008. Available from http://lat.inf.tu-dresden.de/∼clu/papers/.

5. F. Baader, C. Lutz, and B. Suntisrivaraporn. Is tractable reasoning in extensions of the description logic $\mathcal{EL}$ useful in practice? In *Proceedings of M4M-05*, 2005.
6. F. Baader, C. Lutz, and B. Suntisrivaraporn. CEL—a polynomial-time reasoner for life science ontologies. In *Proceedings of IJCAR'06*, volume 4130 of *LNAI*, pages 287–291. Springer-Verlag, 2006.
7. F. Baader, C. Lutz, and B. Suntisrivaraporn. Efficient reasoning in $\mathcal{EL}^+$. In *Proceedings of DL2006*, number 189 in CEUR-WS (http://ceur-ws.org/), 2006.
8. A. K. Chandra, D. C. Kozen, and L. J. Stockmeyer. Alternation. *Journal of the ACM*, 28(1):114–133, 1981.
9. D. Giacomo and Massacci. Combining deduction and model checking into tableaux and algorithms for converse-PDL. *Information and Computation*, 162, 2000.
10. G. D. Giacomo and M. Lenzerini. Boosting the correspondence between description logics and propositional dynamic logics. In *Proceedings of the Twelfth National Conference on Artificial Intelligence (AAAI'94). Volume 1*, pages 205–212. AAAI Press, 1994.
11. J. Y. Halpern and Y. Moses. A guide to completeness and complexity for modal logics of knowledge and belief. *Artificial Intelligence*, 54(3):319–380, 1992.
12. I. Horrocks, O. Kutz, and U. Sattler. The even more irresistible SROIQ. In *Proceedings of KR06*, pages 57–67. AAAI Press, 2006.
13. I. Horrocks and U. Sattler. Decidabiltiy of $\int\langle\rangle$II with complex role inclusion axioms. *Artificial Intelligence*, 160(1–2):79–104, 2004.
14. J.E. Hopcroft and J.D. Ullman. *Introduction to Automata Theory, Languages and Computation*. Addison-Wesley, 1979.
15. Y. Kazakov. *Saturation-Based Decision Procedures for Extensions of the Guarded Fragment*. PhD thesis, Universität des Saarlandes, Saarbrücken, 2006.
16. A. Krisnadhi and C. Lutz. Data complexity in the $\mathcal{EL}$ family of description logics. In *Proceedings of LPAR2007*, volume 4790 of *LNAI*, pages 333–347. Springer-Verlag, 2007.
17. M. Krötzsch, S. Rudolph, and P. Hitzler. Conjunctive queries for a tractable fragment of OWL 1.1. In *Proceedings of ISWC 2007*, volume 4825 of *LNCS*, pages 310–323. Springer-Verlag, 2007.
18. C. H. Papadimitriou. *Computational Complexity*. Addison-Wesley, 1994.
19. A. Rector and I. Horrocks. Experience building a large, re-usable medical ontology using a description logic with transitivity and concept inclusions. In *Proceedings of the Workshop on Ontological Engineering, AAAI Spring Symposium (AAAI'97)*. AAAI Press, 1997.
20. S. Schulz, B. Suntisrivaraporn, and F. Baader. SNOMED CT's problem list: Ontologists' and logicians' therapy suggestions. In *Proceedings of The Medinfo 2007 Congress*, SHTI. IOS Press, 2007.
21. N. Sioutos, S. de Coronado, M. Haber, F. Hartel, W. Shaiu, and L. Wright. NCI thesaurus: a semantic model integrating cancer-related clinical and molecular information. *Journal of Biomedical Informatics*, 40(1):30–43, 2006.
22. K. Spackman. Managing clinical terminology hierarchies using algorithmic calculation of subsumption: Experience with SNOMED-RT. *Journal of the American Medical Informatics Association*, 2000. Fall Symposium Special Issue.
23. B. Suntisrivaraporn, F. Baader, S. Schulz, and K. Spackman. Replacing SEP-triplets in SNOMED CT using tractable description logic operators. In *Proceedings of AIME'07*, LNCS. Springer-Verlag, 2007.
24. The Gene Ontology Consortium. Gene Ontology: Tool for the unification of biology. *Nature Genetics*, 25:25–29, 2000.

A $\mathcal{ELI}$ with GCIs is ExpTime-hard

We prove ExpTime-hardness of subsumption in $\mathcal{ELI}$ with GCIs by reducing the word problem of polynomially space-bounded alternating Turing machines. An *Alternating Turing Machine (ATM)* is of the form $\mathcal{M} = (Q, \Sigma, q_0, \Delta)$. The set of *states* $Q = Q_\exists \uplus Q_\forall \uplus \{q_{\mathsf{acc}}\} \uplus \{q_{\mathsf{rej}}\}$ consists of *existential states* from $Q_\exists$, *universal states* from $Q_\forall$, an *accepting state* q_{acc}, and a *rejecting state* q_{rej}; Σ is the *alphabet* containing a *blank symbol* $\Box$; $q_0 \in Q_\exists \cup Q_\forall$ is the *starting* state; and the *transition relation* δ is of the form

$$\delta \subseteq Q \times \Sigma \times Q \times \Sigma \times \{L, R\}.$$

We write $\delta(q, a)$ for $\{(q', b, M) \mid (q, a, q', b, M) \in \delta\}$.

ATMs run on right-infinite tapes. A *configuration* of an ATM is a word wqw' with $w, w' \in \Sigma^*$ and $q \in Q$. The intended meaning is that the tape contains the word ww' (with only blanks before and behind it), the machine is in state q, and the head is on the leftmost symbol of w'. The *successor configurations* of a configuration wqw' are defined in the usual way in terms of the transition relation δ. A *halting configuration* is of the form wqw' with $q \in \{q_{\mathsf{acc}}, q_{\mathsf{rej}}\}$.

A *computation path* of an ATM $\mathcal{M}$ on a word w is a (finite or infinite) sequence of configurations $c_1, c_2, \dots$ such that $c_1 = q_0 w$ and c_{i+1} is a successor configuration of c_i for $i \geq 0$. The ATMs considered in this paper have only *finite* computation paths on any input. Since this case is simpler than the general one, we define acceptance for ATMs with finite computation paths, only, and refer to [8] for the full definition. Let $\mathcal{M}$ be such an ATM. A halting configuration is *accepting* iff it is of the form $wq_{\mathsf{acc}}w'$. For other configurations $c = wqw'$, the acceptance behaviour depends on q: if $q \in Q_\exists$, then c is accepting iff at least one successor configuration is accepting; if $q \in Q_\forall$, then c is accepting iff all successor configurations are accepting. Finally, the ATM $\mathcal{M}$ with starting state q_0 *accepts* the input w iff the *initial configuration* $q_0 w$ is accepting. We use $L(\mathcal{M})$ to denote the language accepted by $\mathcal{M}$, i.e., $L(\mathcal{M}) = \{w \in \Sigma^* \mid \mathcal{M} \text{ accepts } w\}$.

According to Theorem 3.4 of [8], there is a polynomially space-bounded ATM $\mathcal{M}$ whose word problem is ExpTime-hard. According to Theorem 2.6 of the same paper, we may w.l.o.g. assume that there exists a polynomial p such that the length of every computation path of $\mathcal{M}$ on $w \in \Sigma^n$ is bounded by $2^{p(n)}$, and all the configurations wqw' in such computation paths satisfy $|ww'| \leq p(n)$.

W.l.o.g., we assume that $\mathcal{M}$ never attempts to make a left move on the leftmost field of the tape. In fact, if $\mathcal{M}$ does not satisfy this condition, then it can easily be converted into an ATM $\mathcal{M}'$ that does satisfy it such that $L(\mathcal{M}) = L(\mathcal{M}')$ and the space consumption of $\mathcal{M}'$ is identical to the space consumption of $\mathcal{M}$.

Let $w = a_0 \cdots a_{n-1} \in \Sigma^*$ be an input to $\mathcal{M}$. In the following, we construct an $\mathcal{ELI}$ TBox $\mathcal{T}$ and concepts C and D such that $w \in L(\mathcal{M})$ iff $C \sqsubseteq_\mathcal{T} D$. Intuitively, in a model $\mathcal{I}$ of $\mathcal{T}$ each individual corresponds to a configuration and if two individuals are role successors in $\mathcal{I}$, then they correspond to successor configurations.

We assume that the triples in $\delta(q,a)$ are linearly ordered and that for $t \in \delta(q,a)$, n_t denotes the number of this triple in the ordering (starting with 0). Let

$$m := \max\{\#\delta(q,a) \mid q \in Q, a \in \Sigma\}.$$

In $\mathcal{T}$, we use role names $r_0, \dots, r_m$. Intuitively, a role name r_i connects two successor configurations c and c' such that c' is reachable from c by choosing triple number i from the set $\delta(q,a)$ that is relevant for c. We use the following concept names:

- S_q for $q \in Q$ to denote that the current state is q;
- H_i for $i < p(n)$ to denote that the head is currently on cell i;
- $C_a^{(i)}$ for $a \in \Sigma$ and $i < p(n)$ to denote that the current symbol in the i-th tape cell is a;
- Init to denote the initial configuration and A to denote accepting configurations;
- D to denote configurations that have a "defect" (to be explained below).
- Good is used as a marker for initial configurations that are either accepting or defect.

We can now start to assemble the TBox $\mathcal{T}$. It consists of the following GCIs:

- Set up the initial configuration:

$$\mathsf{Init} \sqsubseteq S_{q_0} \sqcap H_0 \sqcap \bigsqcap_{i<n} C_{a_i}^{(i)} \sqcap \bigsqcap_{n \le i < p(n)} C_{\square}^{(i)}$$

- Make a step to the right. For all $q \in Q$, $a \in \Sigma$, and $i < p(n) - 1$:

$$S_q \sqcap H_i \sqcap C_a^{(i)} \sqsubseteq \bigsqcap_{t=(q',a',R) \in \delta(q,a)} \exists r_{n_t}.(C_{a'}^{(i)} \sqcap S_{q'} \sqcap H_{i+1})$$

- Make a step to the left. For all $q \in Q$, $a \in \Sigma$, and $i \in \{1, \dots, p(n)\}$:

$$S_q \sqcap H_i \sqcap C_a^{(i)} \sqsubseteq \bigsqcap_{t=(q',a',L) \in \delta(q,a)} \exists r_{n_t}.(C_{a'}^{(i)} \sqcap S_{q'} \sqcap H_{i+1})$$

- Symbols that are not under the head do not change. For all $i, j < p(n)$ such that $i \neq j$ and $\ell < m$:

$$\exists r_m^-.(C_a^{(i)} \sqcap H_j) \sqsubseteq C_a^{(i)}$$

- Identify accepting configurations. For all $q \in Q_\forall$, $a \in \Sigma$, $i < p(n)$, and $q' \in Q_\exists$:

$$\begin{aligned} S_{q_{\mathsf{acc}}} &\sqsubseteq A \\ S_q \sqcap H_i \sqcap C_a^{(i)} \sqcap \bigsqcap_{i<\#\delta(q,a)} \exists r_i.A &\sqsubseteq A \\ S_{q'} \sqcap \exists r_i.A &\sqsubseteq A \quad \text{for all } i < m \end{aligned}$$

– Identify defects such as a tape cell labeled with more than one symbol:

$$H_i \sqcap H_j \sqsubseteq D \quad \text{for all } i, j < p(n) \text{ with } i \neq j$$
$$S_q \sqcap S'_q \sqsubseteq D \quad \text{for all } q, q' \in Q \text{ with } q \neq q'$$
$$C_a^{(i)} \sqcap C_{a'}^{(i)} \sqsubseteq D \quad \text{for all } i < p(n) \text{ and } a, a' \in \Sigma \text{ with } a \neq a'$$

– Propagate defects up to the initial configuration. For all $i < m$:

$$\exists r_i.D \sqsubseteq D$$

– The initial configuration is *good* if it is either accepting or defect:

$$\mathsf{Init} \sqcap A \sqsubseteq \mathsf{Good}$$
$$\mathsf{Init} \sqcap D \sqsubseteq \mathsf{Good}$$

Lemma 2. *$\mathcal{M}$ accepts w iff* $\mathsf{Init} \sqsubseteq_{\mathcal{T}} \mathsf{Good}$.

Proof. First assume that $w \notin L(\mathcal{M})$. Then we can construct an interpretation $\mathcal{I}$ by setting $\Delta^{\mathcal{I}}$ to the set of all configurations of $\mathcal{M}$ and interpreting all the concept names S_q, H_i, $C^{(i)}$ in the obvious way, Init as $\{q_0 w\}$, A as the set of accepting configurations, and D and Good as the empty set. The role names are interpreted as follows: $(c, c') \in r_i^{\mathcal{I}}$ iff $c = wqaw'$ and c' can be obtained from c by taking the i-th transition in $\delta(q, a)$. It is straightforward to check that this interpretation is a model of $\mathcal{T}$. Moreover, $q_0 w \in \mathsf{Init}^{\mathcal{I}} \setminus \mathsf{Good}^{\mathcal{I}}$ and thus $\mathsf{Init} \not\sqsubseteq_{\mathcal{T}} \mathsf{Good}$.

Conversely, assume that $\mathsf{Init} \not\sqsubseteq_{\mathcal{T}} \mathsf{Good}$. Then there is a model $\mathcal{I}$ of $\mathcal{T}$ such that there is a $d \in \mathsf{Init}^{\mathcal{I}} \setminus \mathsf{Good}^{\mathcal{I}}$. By construction of $\mathcal{T}$, d represents the initial configuration $q_0 w$ and we can follow the roles r_i to find successor configurations of $q_0 w$, successor configurations of $q_0 w$'s successor configurations, etc. Since $x \notin \mathsf{Good}^{\mathcal{I}}$, $x \notin D^{\mathcal{I}}$ which implies that none of the configurations that we reach when travelling roles r_i is defect. Also, $x \notin \mathsf{Good}^{\mathcal{I}}$ implies $x \notin A^{\mathcal{I}}$. This means that the initial configuration $q_0 w$ is not accepting, and thus $w \notin L(\mathcal{M})$. ❑

Since the size of $\mathcal{T}$ is polynomial in n, we have thus established the following.

Theorem 5. *In $\mathcal{EL}\mathcal{I}$, subsumption w.r.t. GCIs is* ExpTime*-hard.*