Locality and subsumption testing in $\mathcal{EL}$ and some of its extensions

Viorica Sofronie-Stokkermans

Max-Planck-Institut für Informatik, Campus E1 4, Saarbrücken, Germany

Abstract. We show that subsumption problems in $\mathcal{EL}$ and related description logics can be expressed as uniform word problems in classes of semilattices with monotone operators. We use possibilities of efficient local reasoning in such classes of algebras to obtain uniform PTIME decision procedures for CBox subsumption in $\mathcal{EL}$ and extensions thereof. These locality considerations allow us to present a new family of logics which extend $\mathcal{EL}$ and $\mathcal{EL}^+$ with n-ary roles and/or numerical domains.

1 Introduction

Description logics are logics for knowledge representation used in databases and ontologies. They provide a logical basis for modeling and reasoning about objects, classes (or concepts), and relationships (or links, or roles) between them. Recently, less expressive but tractable description logics such as $\mathcal{EL}$ [1] have attracted much interest. Although they have relatively restricted expressivity, this expressivity is sufficient for formalizing the type of knowledge used in widely used ontologies such as the medical ontology SNOMED [13,14]. Several papers were dedicated to studying the properties of $\mathcal{EL}$ and its extensions (e.g. $\mathcal{EL}^+$ [2]), especially to understanding the limits of tractability in extensions of $\mathcal{EL}$.

In this paper we show that subsumption problems in $\mathcal{EL}$ and $\mathcal{EL}^+$ can be expressed as uniform word problems in classes of semilattices with monotone operators. We use this, together with possibilities of efficient local reasoning in such classes of algebras to obtain, in a uniform way, PTIME decision procedures for $\mathcal{EL}$, $\mathcal{EL}^+$, and extensions thereof. These locality considerations allow us to present a new family of (many-sorted) logics which extend $\mathcal{EL}$ with n-ary roles. CBox subsumption is proved to be decidable in PTIME in the one-sorted case and in a two sorted case, in which numerical domains are allowed and numeric concepts (and relationships between then in aCBox) are interpreted in the ORD-Horn fragment of Allen's interval algebra.

Structure of the paper. In Sect. 2 we introduce the description logics $\mathcal{EL}$ and $\mathcal{EL}^+$ and show that CBox subsumption can be expressed as a uniform word problem in the class of semilattices with monotone operators satisfying certain composition axioms. In Sect. 3 we present general definitions and results on (stably) local equational theories. In Sect. 4 we show that the algebraic models of $\mathcal{EL}$ and $\mathcal{EL}^+$ have a stably local presentation, thus providing an alternative proof of the fact that CBox subsumption in $\mathcal{EL}^+$ is decidable in PTIME, and

Table 1. $\mathcal{EL}$ Constructors and their semantics

Constructor name	Syntax	Semantics
conjunction	$C_1 \sqcap C_2$	$C_1^{\mathcal{I}} \cap C_2^{\mathcal{I}}$
existential restriction	$\exists r.C$	$\{x \mid \exists y((x,y) \in r^{\mathcal{I}} \text{ and } y \in C^{\mathcal{I}})\}$

show that locality results can be used for defining extensions of $\mathcal{EL}$ and $\mathcal{EL}^+$ with a subsumption problem decidable in PTIME.

2 The description logics $\mathcal{EL}$ and $\mathcal{EL}^+$

The central notions in description logics are concepts and roles. In any description logic a set N_C of *concept names* and a set N_R of *roles* is assumed to be given. Complex concepts are defined starting with the concept names in N_C, with the help of a set of *concept constructors*. The available constructors determine the expressive power of a description logic. If we allow only intersection and existential restriction as concept constructors, we obtain the description logic $\mathcal{EL}$ [1], a logic used in terminological reasoning in medicine [13,14].

The semantics of description logics is defined in terms of interpretations $\mathcal{I} = (D^{\mathcal{I}}, \cdot^{\mathcal{I}})$, where $D^{\mathcal{I}}$ is a non-empty set, and the function $\cdot^{\mathcal{I}}$ maps each concept name $C \in N_C$ to a set $C^{\mathcal{I}} \subseteq D^{\mathcal{I}}$ and each role name $r \in N_R$ to a binary relation $r^{\mathcal{I}} \subseteq D^{\mathcal{I}} \times D^{\mathcal{I}}$. Table 1 shows the constructor names used in $\mathcal{EL}$ and their semantics. The extension of $\cdot^{\mathcal{I}}$ to concept descriptions is inductively defined using the semantics of the constructors.

Terminology; constraint box. A *terminology* (or TBox, for short) is a finite set consisting of *primitive concept definitions* of the form $C \equiv D$, where C is a concept name and D a concept description; and *general concept inclusions* (GCI) of the form $C \sqsubseteq D$, where C and D are concept descriptions. (Since definitions can be expressed as double inclusions, in what follows we only refer to GCIs.)

In [2], the extension $\mathcal{EL}^+$ of $\mathcal{EL}$ with role inclusion axioms is studied. Counterparts $\mathcal{EL}^d$ and $(\mathcal{EL}^+)^d$ of $\mathcal{EL}$ resp. $\mathcal{EL}^+$, having as constructors only union and universal restrictions and dual role inclusion axioms, can be defined. The relationships between concepts and roles are described using *constraint boxes*.

A *constraint box* (CBox) consists of a terminology $\mathcal{T}$ consisting of a set of GCI, and a set RI of role inclusions of the form $r_1 \circ \cdots \circ r_n \sqsubseteq s$.

Interpretation. An interpretation $\mathcal{I}$ is a model of the CBox $\mathcal{C} = \mathcal{T} \cup RI$ if it satisfies all general concept inclusions in $\mathcal{T}$ and all role inclusions in RI, i.e.:

- $C^{\mathcal{I}} \subseteq D^{\mathcal{I}}$ for all GCI $C \sqsubseteq D \in \mathcal{T}$.
- $r_1^{\mathcal{I}} \circ \ldots \circ r_n^{\mathcal{I}} \subseteq s^{\mathcal{I}}$ for all $r_1 \circ \ldots \circ r_n \subseteq s \in RI$.

Definition 1. *Let $\mathcal{C}$ be a CBox, and C_1, C_2 two concept descriptions. Then $C_1 \sqsubseteq_{\mathcal{C}} C_2$ if and only if $C_1^{\mathcal{I}} \subseteq C_2^{\mathcal{I}}$ for every model $\mathcal{I}$ of $\mathcal{C}$.*

In [2] it was shown that subsumption w.r.t. CBoxes in $\mathcal{EL}^+$ can be reduced in linear time to subsumption w.r.t. *normalized* CBoxes. In normalized CBoxes all GCIs have one of the forms: $C \sqsubseteq D, C_1 \sqcap C_2 \sqsubseteq D, C \sqsubseteq \exists r.D, \exists r.C \sqsubseteq D$, where C, C_1, C_2, D are concept names, and all role inclusions are of the form $r \subseteq s$ or $r_1 \circ r_2 \subseteq r$. Therefore, in what follows, we assume w.l.o.g. that CBoxes only contain role inclusions of the form $r \subseteq s$ and $r_1 \circ r_2 \subseteq r$.

An algebraic semantics for $\mathcal{EL}^+$. A translation of concept descriptions into terms in a signature associated with the set of constructors can be defined as follows. For every role name R, we introduce a unary function symbol $f_{\exists R}$. The translation is inductively defined by: $\overline{C} = C$ for every concept name C; $\overline{C_1 \sqcap C_2} = \overline{C}_1 \wedge \overline{C}_2$, and $\overline{\exists r.C} = f_{\exists r}(\overline{C})$.

There exists a one-to-one correspondence between interpretations of the description logics $\mathcal{E}L, \mathcal{E}L^+, \mathcal{I} = (D, \cdot^{\mathcal{I}})$ and (reducts of) Boolean algebras of sets of the form $(\mathcal{P}(D), \cap, \emptyset, D, \{f_{\exists r}\}_{r \in N_R})$, together with valuations $v : N_C \rightarrow \mathcal{P}(D)$, where $f_{\exists r}$ is defined, for every $U \subseteq D$, by $f_{\exists r}(U) = \{x \mid \exists y((x, y) \in r^{\mathcal{I}} \text{ and } y \in U)\}$. Let $v : N_C \rightarrow \mathcal{P}(D)$ with $v(A) = A^{\mathcal{I}}$ for all $A \in N_C$, and let $\overline{v}$ be the (unique) homomorphic extension of v to terms. Let C be a concept description and $\overline{C}$ be its associated term. Then $C^{\mathcal{I}} = \overline{v}(\overline{C})$ (denoted by $\overline{C}^{\mathcal{I}}$).

It is known that the TBox subsumption problem for the description logic $\mathcal{ALC}$ can be expressed as uniform word problem for Boolean algebras with suitable operators. In [9] we studied the link between TBox subsumption in the description logic $\mathcal{EL}$ and uniform word problems in corresponding classes of semilattices with monotone functions. We now show that these results extend in a natural way to the description logic $\mathcal{EL}^+$. Consider the following classes of algebras:

- $\mathsf{BAO}^{\exists}_{N_R}$ the class of all Boolean algebras with operators $\{f_{\exists r}\}_{r \in N_R}$ of the form $(B, \vee, \wedge, \neg, 0, 1, \{f_{\exists r}\}_{r \in N_R})$, where $f_{\exists r}$ is a join hemimorphism, i.e. $f_{\exists r}(x \vee y) = f_{\exists r}(x) \vee f_{\exists r}(y)$, $f_{\exists r}(0) = 0$;
- $\mathsf{DLO}^{\exists}_{N_R}$ the class of bounded distributive lattices with operators $\{f_{\exists r}\}_{r \in N_R}$, $(L, \vee, \wedge, 0, 1, \{f_{\exists r}\}_{r \in N_R})$, such that $f_{\exists r}$ is a join hemimorphism;
- $\mathsf{SLO}^{\exists}_{N_R}$ the class of all $\wedge$-semilattices with operators $\{f_{\exists r}\}_{r \in N_R}$, $(S, \wedge, 0, 1, \{f_{\exists R}\}_{R \in N_R})$, such that $f_{\exists R}$ is monotone.

Assume given a set RI of axioms of the form $r \subseteq s$ and $r_1 \circ r_2 \subseteq r$, with $r_1, r_2, r \in N_R$. We denote by $\mathsf{SLO}^{\exists}_{N_R}(RI)$ the subclass of $\mathsf{SLO}^{\exists}_{N_R}$ consisting of those algebras which satisfy the set of axioms:

$$RI_a = \{(f_{\exists r_2} \circ f_{\exists r_1})(x) \leq f_{\exists r}(x) \mid r_1 \circ r_2 \subseteq r \in RI\} \cup \{f_{\exists r}(x) \leq f_{\exists s}(x) \mid r \subseteq s \in RI\}.$$

Lemma 1 *Let $\mathcal{I} = (D, \cdot^{\mathcal{I}})$ be a model of an $\mathcal{EL}^+$ CBox, $\mathcal{C} = GCI \cup RI$. Then $(\mathcal{P}(D), \cap, \{f_{\exists r}\}_{r \in N_R}) \in \mathsf{SLO}^{\exists}_{N_R}(RI)$.*

Proof: Clearly, $(\mathcal{P}(D), \cap, \{f_{\exists r}\}_{r \in N_R}) \in \mathsf{SLO}^{\exists}_{N_R}$. Let $r_1, r_2, r \in N_R$ and $U \in \mathcal{P}(D)$.

$$\begin{aligned} f_{\exists r_1}(U) &= \{x \mid \exists y \in U \text{ s.t. } (x, y) \in r_1^{\mathcal{I}}\} \subseteq f_{\exists r}(U) \quad \text{if } r_1 \subseteq r \in RI \\ f_{\exists r_2}(f_{\exists r_1}(U)) &= \{x \mid \exists y((x, y) \in r_2^{\mathcal{I}} \text{ and } \exists z \in U, (y, z) \in r_1^{\mathcal{I}})\} \\ &= \{x \mid \exists z \in U \text{ s.t. } (x, z) \in (r_1 \circ r_2)^{\mathcal{I}}\} \subseteq f_{\exists r}(U) \text{ if } r_1 \circ r_2 \subseteq r \in RI. \end{aligned}$$

Lemma 2 *Every* $\mathbf{S} \in \mathsf{SLO}^{\exists}_{N_R}(RI)$ *embeds into a lattice in* $\mathsf{DLO}^{\exists}_{N_R}$ *satisfying the axioms in* RI. *Every lattice in* $\mathsf{DLO}^{\exists}_{N_R}$ *satisfying the axioms in* RI_a *embeds into a lattice in* $\mathsf{BAO}^{\exists}_{N_R}$ *satisfying the axioms in* RI_a.

Theorem 3 *Let* $\mathcal{T}$ *be an* $\mathcal{EL}$ *TBox consisting of the general concept inclusions* GCI. *Then* $C_1 \sqsubseteq_{\mathcal{T}} C_2$ *iff* $\mathsf{BAO}^{\exists}_{N_R} \models \left(\bigwedge_{C \sqsubseteq D \in \mathcal{T}} \overline{C} \leq \overline{D} \right) \rightarrow \overline{C_1} \leq \overline{C_2}$.

Proof: Follows from the fact that every algebra in $\mathsf{BAO}^{\exists}_{N_R}$ homomorphically embeds into a Boolean algebra of sets. □

Theorem 4 *If the only concept constructors are intersection and existential restriction, then for all concept descriptions* C_1, C_2 *and every* $\mathcal{EL}^+$ *CBox* $\mathcal{C} = GCI \cup RI$, $C_1 \sqsubseteq_{\mathcal{C}} C_2$ *iff* $\mathsf{SLO}^{\exists}_{N_R}(RI) \models \left(\bigwedge_{C \sqsubseteq D \in GCI} \overline{C} \leq \overline{D} \right) \rightarrow \overline{C_1} \leq \overline{C_2}$.

Proof: This is a direct consequence of Theorem 3 and Lemma 2. □

We will show that the word problem for the class of algebras $\mathsf{SLO}^{\exists}_{N_R}(RI)$ is decidable in PTIME. For this we prove that $\mathsf{SLO}^{\exists}_{N_R}(RI)$ has a "local" presentation. The locality definitions, and the methods for recognizing local presentations are given in Sect. 3. The application to $\mathsf{SLO}^{\exists}_{N_R}$ and $\mathsf{SLO}^{\exists}_{N_R}(RI)$ are given in Sect. 4.

3 Local equational theories

A set $\mathcal{K}$ of Horn clauses is *stably local* [5] if for every set G of ground clauses, if $\mathcal{K} \wedge G \models \perp$ then G can be refuted using the set $\mathcal{K}^{[G]}$ of all instances of $\mathcal{K}$ obtained by instantiating the variables with (ground) subterms of G, i.e. if

$$\mathcal{K} \wedge G \models \perp \text{ if and only if } \mathcal{K}^{[G]} \wedge G \models \perp .$$

The more general notion of Ψ-stably local theory (in which the instances to be considered are described by a closure operation Ψ) is introduced in [6]. Locality can be recognized by proving embeddability of partial into total models [10,12,6].

Partial and total models. A partial model is a model in which some function symbols may be partial. In this paper the models we consider are partially ordered algebraic structures, i.e. the only predicates are $\leq$ and $=$. If A is a partial structure and $\beta : X \rightarrow \mathcal{A}$ is a valuation we say that $(A, \beta) \models t_1 = t_2$ iff at least one of the following conditions holds:

(a) $\beta(t_1), \beta(t_2)$ are defined and $\beta(t_1) = \beta(t_2)$, or
(b) $\beta(t_1)$ and $\beta(t_2)$ are undefined, or
(c) $\beta(t_1)$ is defined, $t_2 = f(s_1, \ldots, s_n)$ and $\beta(s_i)$ is undefined for some i, or
(d) if $\beta(t_1)$ is defined, $t_2 = f(s_1, \ldots, s_n)$ and $\beta(s_i)$ is defined for all i then $\beta(t_2)$ has to be defined and $\beta(t_1) = \beta(t_2)$.

$(A, \beta) \models t_1 \leq t_2$ is defined similarly, replacing "$=$" with "$\leq$" in (a)–(d).

(A, β) *satisfies a clause* C (notation: $(A, \beta) \models C$) if it satisfies at least one literal in C. A is an *(Evans) partial model* of a set of clauses $\mathcal{K}$ if $(A, \beta) \models C$ for every valuation β and every clause C in $\mathcal{K}$.

Locality. Let $\mathcal{K}$ be a set of clauses. Let $\Psi_{\mathcal{K}}$ be a function associating with any set T of ground terms a set $\Psi_{\mathcal{K}}(T)$ of ground terms such that

(i) all ground subterms in $\mathcal{K}$ and T are in $\Psi_{\mathcal{K}}(T)$;
(ii) for all sets of ground terms T, T' if $T \subseteq T'$ then $\Psi_{\mathcal{K}}(T) \subseteq \Psi_{\mathcal{K}}(T')$;
(iii) for all sets of ground terms T, $\Psi_{\mathcal{K}}(\Psi_{\mathcal{K}}(T)) \subseteq \Psi_{\mathcal{K}}(T)$;
(iv) Ψ is compatible with any map h between constants, i.e. for any map $h : C \rightarrow C$, $\Psi_{\mathcal{K}}(\overline{h}(T)) = \overline{h}(\Psi_{\mathcal{K}}(T))$, where $\overline{h}$ is the unique extension of h to terms.

Let $\mathcal{K}^{[\Psi_{\mathcal{K}}(G)]}$ be the set of instances of $\mathcal{K}$ where the variables are instantiated with terms in $\Psi_{\mathcal{K}}(\mathsf{st}(\mathcal{K}, G))$, where $\mathsf{st}(\mathcal{K}, G)$ is the set of all ground terms occurring in $\mathcal{K}$ or G. We say that $\mathcal{K}$ is Ψ-stably local if it satisfies:

(SLoc^{Ψ}) for every finite set G of ground clauses, $\mathcal{K} \cup G \models \perp$ iff $\mathcal{K}^{[\Psi_{\mathcal{K}}(G)]} \cup G$ has no partial model in which all terms in $\Psi_{\mathcal{K}}(G)$ are defined.

In the particular case that $\Psi_{\mathcal{K}}(G) = \mathsf{st}(\mathcal{K}, G)$ we refer to *stable locality* of the extension. The corresponding condition is denoted SLoc.
If a set $\mathcal{K}$ of Horn clauses satisfies (SLoc^{Ψ}) then satisfiability of any set G of Horn clauses w.r.t. $\mathcal{K}$ is decidable in polynomial time in the size of $\Psi_{\mathcal{K}}(G)$. This follows from the fact that $\mathcal{K}^{[\Psi_{\mathcal{K}}(G)]} \cup G$ is a set of ground Horn clauses of size polynomial in the size of $\Psi_{\mathcal{K}}(G)$, and satisfiability of sets of ground Horn clauses (e.g. in a relational encoding, taking into account only suitable instances of the congruence axioms - which are again Horn and have a size quadratic in $|\Psi_{\mathcal{K}}(G)|$) can be checked in linear time ([4], see also [5]).

Locality and embeddability. Theories satisfying (SLoc^{Ψ}) can be recognized by showing that Evans partial models of $\mathcal{T}_1$ embed into total models.

Theorem 5 *Let $\mathcal{K}$ be a set of clauses. If $\Psi_{\mathcal{K}}$ satisfies conditions (i)–(iv) above, and every Evans partial model of $\mathcal{K}$ – in which the set of defined terms is closed under $\Psi_{\mathcal{K}}$ – weakly embeds into a total model of $\mathcal{K}$, then $\mathcal{K}$ satisfies* SLoc^{Ψ}.

4 Locality of $\mathcal{EL}^+$, $\mathcal{EL}$ and their extensions

We now show that the classes of algebraic models of $\mathcal{EL}^+$ and of $\mathcal{EL}$ have presentations which satisfy certain locality properties. This gives an alternative, algebraic explanation of the fact that TBOX subsumption in these logics is decidable in PTIME and allows for several generalizations.

4.1 Locality and $\mathcal{EL}^+$

We prove that the class $\mathsf{SLO}_{\Sigma}(RI)$ of semilattices with monotone operators in a set Σ satisfying a family RI_a of axioms of the form $\forall x \, (f_1 \circ \cdots \circ f_n)(x) \leq f(x)$

has a local presentation, and therefore the uniform word problem w.r.t. this class can be decided in PTIME. (We here restrict, w.l.o.g., to axioms as above with $n \in \{1, 2\}$.) It is known that the theory of lattices allows a local Horn axiomatization (cf. e.g. [8,3]). Let SL be such an axiomatization for the theory of lattices. We denote by $\mathsf{Mon}(\Sigma)$ the set $\{\mathsf{Mon}(f) \mid f \in \Sigma\}$, where

$$\mathsf{Mon}(f) \qquad \forall x, y(x \leq y \rightarrow f(x) \leq f(y)).$$

Theorem 6 *The set of Horn clauses $SL \cup \mathsf{Mon}(\Sigma) \cup RI_a$ has the property that every Evans partial model A with the properties:*

(i) for every $f \in \Sigma$, f_A is a partial function with finite definiton domain;
(ii) for each axiom in RI_a of the form $(f_1 \circ f_2)(x) \leq f(x)$, and every $a \in A$, if $f(a)$ is defined then $f_2(a)$ is defined in A;
(iii) $A \models SL \cup \mathsf{Mon}(\Sigma) \cup RI_a$;

weakly embeds into a total model of $SL \cup \mathsf{Mon}(\Sigma) \cup RI_a$.

Proof: Let A be an Evans partial model of $SL \cup \mathsf{Mon}(\Sigma) \cup RI_a$ with properties (i)–(iii). Since A is a poset, it embeds into a complete (semi)lattice S such that the meets that exist in A are preserved. For every $f \in \Sigma$ we define $\overline{f} : S \rightarrow S$ by

$$\overline{f}(a) = \bigwedge \{f(c) \mid a \leq c, c \in A, f_A(c) \text{ is defined}\}.$$

Then $\overline{f}$ is monotone [12]. It can be shown that the axioms in RI_a are satisfied.

Corollary 7 *The following are equivalent:*

(1) $SL \cup \mathsf{Mon}(\Sigma) \cup RI_a \models \forall \overline{x} \bigwedge_{i=1}^n s_i(\overline{x}) \leq s'_i(\overline{x}) \rightarrow s(\overline{x}) \leq s'(\overline{x})$;
(2) $SL \cup \mathsf{Mon}(\Sigma) \cup RI_a \wedge G \models \perp$, where $G = \bigwedge_{i=1}^n s_i(\overline{c}) \leq s'_i(\overline{c}) \wedge s(\overline{c}) \not\leq s'(\overline{c})$;
(3) $(SL \cup \mathsf{Mon}(\Sigma) \cup RI_a)^{[\Psi_{RI}(G)]} \wedge G \models \perp$ where $\Psi_{RI}(G) = \bigcup_{i \geq 0} \Psi^i_{RI}$, with $\Psi^0_{RI} = \mathsf{st}(G)$, and $\Psi^{i+1}_{RI} = \{f_2(d) \mid f(d) \in \Psi^i_{RI}, (f_1 \circ f_2)(x) \leq f(x) \in RI_a\}$.

Here $\mathsf{st}(G)$ is the set of all (ground) subterms occurring in G. Note that $\Psi_{RI}(G)$ can have at most $|\mathsf{st}(G)| \cdot |N_R|$ elements. Thus, its size is polynomial in the size of G. On the other hand, the number of clauses in $(SL \cup \mathsf{Mon}(\Sigma) \cup RI_a)^{[\Psi_{RI}(G)]}$ is polynomial in $|\Psi_{RI}(G)|$, and satisfiability of any set of ground clauses can be tested in polynomial time. This shows that the uniform word problem for the class $\mathsf{SLO}_\Sigma(RI)$ (and thus also for $\mathsf{SLO}^\exists_{NR}(RI)$) is decidable in polynomial time.

4.2 Locality and $\mathcal{EL}$

In [9] we proved that the algebraic counterpart of the description logic $\mathcal{EL}$, namely the class of semilattices with monotone operators – axiomatized by $SL \cup \mathsf{Mon}(\Sigma)$ – has a stronger locality property: for every set G of ground clauses

$$SL \cup \mathsf{Mon}(\Sigma) \wedge G \models \perp \quad \text{if and only if} \quad (SL \cup \mathsf{Mon}(\Sigma))[G] \wedge G \models \perp$$

where $\mathcal{K}[G]$ is the set of instances of $\mathcal{K}$ containing only ground terms occurring in G. In fact, we showed that the extension of the theory SL of semilattices with monotone functions is local in the sense defined in [10] (cf. e.g. [12]).

Theorem 8 *For any set G of ground clauses the following are equivalent:*

(1) $SL \cup \mathsf{Mon}(\Sigma) \wedge G \models \perp$.
(2) $SL \cup \mathsf{Mon}(\Sigma)[G] \wedge G$ *has no partial model A such that its $\{\wedge\}$-reduct is a (total) semilattice and the functions in Σ are partially defined, their domain of definition is finite and all terms in G are defined in A.*

Let $\mathsf{Mon}(\Sigma)[G]_0 \wedge G_0 \wedge \mathsf{Def}$ *be obtained from* $\mathsf{Mon}(\Sigma)[G] \wedge G$ *by purification, i.e. by replacing, in a bottom-up manner, all subterms $f(g)$ with $f \in \Sigma$, with newly introduced constants $c_{f(g)}$ and adding the definitions $f(g) = c_t$ to the set* Def. *The following are equivalent (and equivalent to (1) and (2)):*

(3) $\mathsf{Mon}(\Sigma)[G]_0 \wedge G_0 \wedge \mathsf{Def}$ *has no partial model $(A, \wedge, \{f_A\}_{f \in \Sigma})$ such that $(A, \wedge)$ is a semilattice and for all $f \in \Sigma$, f_A is partially defined, its domain of definition is finite and all terms in* Def *are defined in A;*
(4) $\mathsf{Mon}(\Sigma)[G]_0 \wedge G_0$ *is unsatisfiable in SL.*
(Note that in the presence of $\mathsf{Mon}(\Sigma)$ *the instances* $\mathsf{Con}[G]_0$ *of the congruence axioms for the functions in Σ are not necessary.)*
$\mathsf{Con}[G]_0 = \{g{=}g' \rightarrow c_{f(g)}{=}c_{f(g')} \mid f(g){=}c_{f(g)}, f(g'){=}c_{f(g')} \in \mathsf{Def}\}$.

This equivalence allows us to hierarchically reduce, in polynomial time, proof tasks in $SL \cup \mathsf{Mon}(\Sigma)$ to proof tasks in SL (cf. e.g. [12]) which can then be solved in polynomial time.

The results described above can easily be generalized to semilattices with n-ary monotone functions satisfying composition axioms. This allows us to define natural generalizations of $\mathcal{EL}$ and $\mathcal{EL}^+$. We start by presenting a generalization of $\mathcal{EL}$ in which n-ary roles are allowed. We then sketch possible extensions in which role inclusions are also taken into account.

4.3 Extensions of $\mathcal{EL}$

We consider extensions of $\mathcal{EL}$ with n-ary roles. The semantics is defined in terms of interpretations $\mathcal{I} = (D^{\mathcal{I}}, \cdot^{\mathcal{I}})$, where $D^{\mathcal{I}}$ is a non-empty set, concepts are interpreted as usual, and each n-ary role $R \in N_R$ is interpreted as an n-ary relation $R^{\mathcal{I}} \subseteq (D^{\mathcal{I}})^n$. The constructors are conjunction (interpreted as intersection) and existential restriction (is interpretation extends naturally that for $n = 2$):

$$(\exists R.(C_1, \dots C_n))^{\mathcal{I}} = \{x \mid \exists y_1, \dots, y_n \, (x, y_1, \dots, y_n) \in R^{\mathcal{I}} \text{ and } y_i \in C_i^{\mathcal{I}}\}.$$

A further extension is obtained by allowing for certain concrete sorts in the interpretation of concepts. These have the same support in all interpretations. We may additionally assume that there exist specific concrete concepts which have a fixed semantics in all interpretations, or that all concrete concepts have interpretations with additional properties.

Example 1. Consider a description logic having a usual (concept) sort and a 'concrete' sort num with fixed domain $\mathbb{N}$. We may be interested in general concrete concepts of sort num (interpreted as subsets of $\mathbb{R}$) or in special concepts of sort num such as $\uparrow n$, $\downarrow n$, or $[n, m]$ for $m, n \in \mathbb{R}$. For any interpretation $\mathcal{I}$, $\uparrow n^{\mathcal{I}} = \{x \in \mathbb{R} \mid x \geq n\}$, $\downarrow n^{\mathcal{I}} = \{x \in \mathbb{R} \mid x \leq n\}$, and $[n, m]^{\mathcal{I}} = \{x \in \mathbb{R} \mid n \leq x \leq m\}$. We will denote the arities of roles using a many-sorted framework. Let $(D, \mathbb{R}, \cdot^{\mathcal{I}})$ be an interpretation with two sorts concept and num. A role with arity $(s_1, \dots, s_n)$ is interpreted as a subset of $D_{s_1} \times \dots \times D_{s_n}$, where $D_{\mathsf{concept}} = D$ and $D_{\mathsf{num}} = \mathbb{R}$.

1. Let price be a binary role or arity (concept, num), which associates with every element of sort concept its possible prices. The concept
$$\exists \mathsf{price}.\uparrow n = \{x \mid \exists k \geq n : \mathsf{price}(x, k)\}$$
represents the class of all individuals with some price greater than n.
2. Let has-weight-price be a role of arity (concept, num, num). The concept $\exists$ has-weight-price.$(\uparrow \mathsf{y}, \downarrow \mathsf{p}) = \{x \mid \exists y' \geq \mathsf{y}, \exists p' \leq \mathsf{p}$ and has-weight-price$(x, y', p')\}$ denotes the family of individuals for which a weight above y and a price below p exist.

The example below can be generalized by allowing a set of concrete sorts. We can prove that semilattices with monotone n-ary operators define the algebraic semantics of this extension of $\mathcal{EL}$. The proof is analogous to that of Theorem 4.

Let $\mathsf{SLO}^{\exists}_{N_R,S}$ denote the class of all structures $(S, \mathcal{P}(A_1), \dots, \mathcal{P}(A_n), \{f_{\exists r} \mid r \in N_R\})$, where S is a semilattice, $A_1, \dots, A_n$ are concrete domains, and $\{f_{\exists r} \mid r \in N_R\}$ are n-ary monotone operators. In addition we may allow constants of concrete sort, interpreted as sets in $\mathcal{P}(A_i)$. The classes $\mathsf{DLO}^{\exists}_{N_R,S}$ and $\mathsf{BAO}^{\exists}_{N_R,S}$ of all distributive lattices resp. Boolean algebras with concrete supports and n-ary join hemimorphisms $\{f_{\exists r} \mid r \in N_R\}$ are defined similarly.

Theorem 9 *Assume that the only concept constructors are intersection and existential restriction. For all concept descriptions C_1, C_2, and every TBox $\mathcal{T}$ consisting of general concept inclusions GCI, $C_1 \sqsubseteq_{\mathcal{T}} C_2$ if and only if $\mathsf{SLO}^{\exists}_{N_R,S} \models \left(\bigwedge_{C \sqsubseteq D \in GCI} \overline{C} \leq \overline{D}\right) \rightarrow \overline{C_1} \leq \overline{C_2}$.*

Let SL_S be the class of all structures $\mathcal{A} = (A, \mathcal{P}(A_1), \dots, \mathcal{P}(A_n))$, with signature $\Pi = (S, \{\wedge\} \cup \Sigma, \mathsf{Pred})$ with $S = \{\mathsf{concept}, \mathsf{s_1}, \dots, \mathsf{s_n}\}$, $\mathsf{Pred} = \{\leq\} \cup \{\subseteq_i \mid 1 \leq i \leq n\}$, where $A \in SL$, the support of sort concept of $\mathcal{A}$ is A, and for all i the support sort s_i of $\mathcal{A}$ is $\mathcal{P}(A_i)$.

Theorem 10 ([12]) *Every structure $(A, \mathcal{P}(A_1), \dots, \mathcal{P}(A_n), \{f_A\}_{f \in \Sigma})$, where*

(i) $(A, \mathcal{P}(A_1), \dots, \mathcal{P}(A_n)) \in SL_S$, and
(ii) for every $f \in \Sigma$ of arity $s_1 \dots s_n \rightarrow s$, f_A is a partial function from $\prod_{i=1}^n U_{s_i}$ to U_s with a finite definition domain on which it is monotone,

weakly embeds into a total model of $SLO_{\Sigma,S}$ (axiomatized by $SL_S \cup \mathsf{Mon}(\Sigma)$).

Corollary 11 *Let $G = \bigwedge_{i=1}^n s_i(\overline{c}) \leq s'_i(\overline{c}) \wedge s(\overline{c}) \not\leq s'(\overline{c})$ be a set of ground clauses in the extension Π^c of Π with new constants. The following are equivalent:*

(1) $SL_S \cup \mathsf{Mon}(\Sigma) \wedge G \models \perp$;
(2) $SL_S \cup \mathsf{Mon}(\Sigma)[G] \wedge G$ *has no partial model with a total* $\{\wedge_{SL}\}$*-reduct in which all terms in* G *are defined.*

Let $\bigcup_{i=0}^{n} \mathsf{Mon}(\Sigma)[G]_i \wedge G_i \wedge \mathsf{Def}$ *be obtained from* $\mathsf{Mon}(\Sigma)[G] \wedge G$ *by purification, i.e. by replacing, in a bottom-up manner, all subterms* $f(g)$ *of sort* s *with* $f \in \Sigma$*, with newly introduced constants* $c_{f(g)}$ *of sort* s *and adding the definitions* $f(g) = c_t$ *to the set* Def*. We thus separate* $\mathsf{Mon}(\Sigma)[G] \wedge G$ *into a conjunction of constraints* $\Gamma_i = \mathsf{Mon}(\Sigma)[G]_i \wedge G_i$*, where* Γ_0 *is a constraint of sort* semilattice *and for* $1 \leq i \leq n$*,* Γ_i *is a set of constraints over terms of sort* i *(*i *being the concrete sort with fixed support* $\mathcal{P}(A_i)$*). Then the following are equivalent (and are also equivalent to (1) and (2)):*

(3) $\bigcup_{i=0}^{n} \mathsf{Mon}(\Sigma)[G]_i \wedge G_i \wedge \mathsf{Def}$ *has no partial model with a total* $\{\wedge_{SL}\}$*-reduct in which all terms in* Def *are defined.*
(4) $\bigcup_{i=0}^{n} \mathsf{Mon}(\Sigma)[G]_i \wedge G_i$ *is unsatisfiable in the many-sorted disjoint combination of* SL *and the concrete theories of* $\mathcal{P}(A_i)$*,* $1 \leq i \leq n$*.*

The complexity of the uniform word problem of $SL_S \cup \mathsf{Mon}(\Sigma)$ depends on the complexity of the problem of testing the satisfiability w.r.t. the many-sorted disjoint combination of SL with the concrete theories of $\mathcal{P}(A_i)$, $1 \leq i \leq n$ – of sets $C_{\mathsf{concept}} \cup \bigcup_{i=1}^{n} C_i \cup \mathsf{Mon}$, where C_{concept} and C_i are unit clauses of sort concept resp. s_i, and Mon consists of possibly mixed ground Horn clauses.

Specific extensions of the logic $\mathcal{EL}$ can be obtained by imposing restrictions on the interpretation of the "concrete"-type concepts within $\mathcal{P}(A_i)$. (We can e.g. require that numerical concepts are interpreted as intervals, as in Example 1.)

Theorem 12 *CBox subsumption is decidable in PTIME for the following extensions of* $\mathcal{EL}$ *with* n*-ary roles:*

(1) The one-sorted extension of $\mathcal{EL}$ *with* n*-ary roles.*
(2) The extension of $\mathcal{EL}$ *with sorts* {concept, num}*, such that (i) the semantics of the sort* concept *is the usual one, (ii) the concepts of sort* num *are interpreted as elements in the ORD-Horn, convex fragment of Allen's interval algebra [7], and (iii) any CBox can contain many-sorted GCI's over concepts, and constraints over the numerical data expressible in the ORD-Horn fragment.*

Example 2. Consider the special case described in Example 1. Assume that the concepts of sort num used in any TBox are of the form $\uparrow n, \downarrow m$ and $[n, m]$. Consider the TBox $\mathcal{T}$ consisting of the following GCIs:

$$\{\exists \mathsf{price}(\downarrow n_1) \sqsubseteq \mathsf{affordable},\ \exists \mathsf{weight}(\uparrow m_1) \sqcap \mathsf{car} \sqsubseteq \mathsf{truck},\ \downarrow n_1 \sqsubseteq \downarrow n,\ \uparrow m_1 \sqsubseteq \uparrow m,\ C \sqsubseteq \mathsf{car},$$
$$\mathsf{has\text{-}weight\text{-}price}(\uparrow m, \downarrow n) \sqsubseteq \exists \mathsf{price}(\downarrow n) \sqcap \exists \mathsf{weight}(\uparrow m),\ C \sqsubseteq \exists \mathsf{has\text{-}weight\text{-}price}(\uparrow m, \downarrow n)\}$$

In order to prove that $C \sqsubseteq_{\mathcal{T}} \mathsf{affordable} \sqcap \mathsf{truck}$ we proceed as follows. We refute $\bigwedge_{D \sqsubseteq D' \in \mathcal{T}} \overline{D} \leq \overline{D}' \wedge \overline{C} \not\leq \mathsf{affordable} \wedge \mathsf{truck}$. We purify the problem introducing definitions for the terms starting with existential restrictions and obtain the following set of constraints:

Def	C_{num}	C_{concept}	Mon
$f_{\mathsf{price}}(\downarrow n_1) = c_1$	$\downarrow n \leq \downarrow n_1$	$d_1 \leq \mathsf{affordable}$	$\downarrow n_1 \leq \downarrow n \rightarrow c_1 \leq c$
$f_{\mathsf{price}}(\downarrow n) = c$	$\uparrow m \leq \uparrow m_1$	$d_1 \wedge \mathsf{car} \leq \mathsf{truck}$	$\downarrow n_1 \geq \downarrow n \rightarrow c_1 \geq c$
$f_{\mathsf{weight}}(\uparrow m_1) = d_1$		$e \leq c \wedge d$	$\uparrow m_1 \leq \uparrow m \leq d_1 \leq d$
$f_{\mathsf{weight}}(\uparrow m) = d$		$C \leq \mathsf{car}$	$\uparrow m_1 \geq \uparrow m \leq d_1 \geq d$
$f_{\text{h-w-p}}(\uparrow m, \downarrow n) = e$		$C \leq e$	
		$C \not\leq \mathsf{affordable} \wedge \mathsf{truck}$	

The task of proving $C \sqsubseteq_{\mathcal{T}} \mathsf{affordable} \sqcap \mathsf{truck}$ can therefore be reduced to checking if $C_{\mathsf{num}} \wedge C_{\mathsf{concept}} \wedge \mathsf{Mon}$ is satisfiable w.r.t. the combination of SL (sort concept) with $\mathcal{P}(\mathbb{Q})$ (sort num). For this, we note that C_{num} entails the premises of the first, second, and fourth monotonicity rules. Thus, we can add $c \leq c_1$ and $d \leq d_1$ to C_{concept}. Thus, we deduce that $C \leq e \wedge \mathsf{car} \leq (c \wedge d) \wedge \mathsf{car} \leq c_1 \wedge (d_1 \wedge \mathsf{car}) \leq \mathsf{affordable} \wedge \mathsf{truck}$, which contradicts the last clause in C_{concept}.

4.4 Extensions of $\mathcal{EL}^+$

For roles with arbitrary arity we also consider role inclusions of the form $r_1 \circ r_2 \sqsubseteq r$ with the semantics that for every interpretation $\mathcal{I}$, if $(x_1, \dots, x_n) \in r_1^{\mathcal{I}}$ and $(x_n, \dots, x_{n+k}) \in r_2^{\mathcal{I}}$ then the tuple $(x_1, \dots, x_{n-1}, x_{n+1}, \dots, x_{n+k}) \in r^{\mathcal{I}}$. Monotone functions $f_{\exists r}, f_{\exists r_2}$ associated with r_1, r_2 can be constructed by defining:

$$f_{\exists r_i}(U_1, \dots, U_{m_i}) = \{x \mid \exists y_1, \dots, y_{m_i}, (x, y_1, \dots, y_{m_i}) \in r_i^{\mathcal{I}} \text{ and } y_j \in U_j\}.$$

The composition rules which correspond to the role inclusion constraints are:

$$\begin{aligned}
&f_{\exists r_1}(U_2, \dots, U_{n-1}, f_{\exists r_2}(U_{n+1}, \dots, U_{n+k})) = \\
&\quad = \{y_1 \mid \exists y_i \in U_i, 2 \leq i \leq n-1, \exists y_n \in f_{\exists r_2}(U_{n+1}, \dots, U_{n+k}), (y_1, y_2, \dots, y_n) \in r_1\} = \\
&\quad = \{y_1 \mid \exists y_i \in U_i, 2 \leq i \leq n-1, \exists y_i \in U_i, n+1 \leq i \leq n+k, \\
&\qquad\quad (y_n, y_{n+1}, \dots, y_{n+k}) \in r_2^{\mathcal{I}} \text{ and } (y_1, y_2, \dots, y_n) \in r_1^{\mathcal{I}}\} = \\
&\quad = \{y_1 \mid \exists y_i \in U_i, \text{ for } 2 \leq i \leq n+k, i \neq n, (y_1, y_2, \dots, y_{n-1}, y_{n+1}, \dots, y_{n+k}) \in r_2^{\mathcal{I}} \circ r_1^{\mathcal{I}}\} \\
&\quad \subseteq \{y_1 \mid \exists y_i \in U_i, \text{ for } 2 \leq i \leq n+k, i \neq n, (y_1, y_2, \dots, y_{n-1}, y_{n+1}, \dots, y_{n+k}) \in r^{\mathcal{I}}\} \\
&\quad = f_{\exists r}(U_2, \dots, U_{n-1}, U_{n+1}, \dots, U_{n+k}) \quad \text{ if } r_2 \circ r_1 \subseteq r.
\end{aligned}$$

Locality results similar to those in Thm. 6 can be obtained also in this case.

5 Conclusions

In this paper we have shown that subsumption problems in $\mathcal{EL}$ can be expressed as uniform word problems in classes of semilattices with monotone operators, and that subsumption problems in $\mathcal{EL}^+$ can be expressed as uniform word problems in classes of semilattices with monotone operators satisfying certain composition laws. This allowed us to obtain, in a uniform way, PTIME decision procedures for $\mathcal{EL}$, $\mathcal{EL}^+$, and extensions thereof. These locality considerations allow us to present a new family of PTIME (many-sorted) logics which extend $\mathcal{EL}$.

The results in [11] show that the class of semilattices with monotone operations allows ground (equational) interpolation. We plan to use the results presented in this paper for studying interpolation properties in extensions of $\mathcal{EL}$ and for analyzing possibilities of efficient (modular) reasoning in combinations of ontologies based on extensions of $\mathcal{EL}$.

Acknowledgments. This work was partly supported by the German Research Council (DFG) as part of the Transregional Collaborative Research Center "Automatic Verification and Analysis of Complex Systems" (SFB/TR 14 AVACS). See `www.avacs.org` for more information.

References

1. F. Baader. Terminological Cycles in a Description Logic with Existential Restrictions. In: G. Gottlob and T. Walsh, editors, *Proc. of the 18th International Joint Conference in Artificial Intelligence*, pages 325–330, Morgan Kaufmann, 2003.
2. F. Baader, C. Lutz, B. Suntisrivaraporn. Is Tractable Reasoning in Extensions of the Description Logic EL Useful in Practice? *Journal of Logic, Language and Information* (M4M special issue). To appear.
3. S. Burris. Polynomial time uniform word problems. *Mathematical Logic Quarterly*, 41:173–182, 1995.
4. W.F. Dowling and J.H. Gallier. Linear time algorithms for testing the satisfiability of propositional Horn formulae. *J. Logic Programming* 77, 192–217, 1984.
5. H. Ganzinger. Relating semantic and proof-theoretic concepts for polynomial time decidability of uniform word problems. In *Proc. 16th IEEE Symposium on Logic in Computer Science (LICS'01)*, pages 81–92. IEEE Computer Society Press, 2001.
6. C. Ihlemann, S. Jacobs, and V. Sofronie-Stokkermans. On local reasoning in verification. In *Proc. TACAS 2008*, LNCS 4963, pages 265–281, Springer 2008.
7. B. Nebel and H.-J. Bürckert. Reasoning about temporal relations: A maximal tractable subclass of Allen's interval algebra. *Journal of the ACM*, 42(1): 43–66, 1995.
8. T. Skolem. (1920). Logisch-kombinatorische Untersuchungen über die Erfüllbarkeit und Beweisbarkeit mathematischer Sätze nebst einem Theorem über dichte Mengen. *Skrifter utgit av Videnskabsselskapet i Kristiania, I. Matematisk-naturvidenskabelig klasse, 4*, pages 1–36.
9. V. Sofronie-Stokkermans. Automated theorem proving by resolution in non-classical logics. *Annals of Mathematics and Artificial Intelligence* (Special issue "Knowledge Discovery and Discrete Mathematics: Dedicated to the Memory of Peter L. Hammer"), 49 (1-4): 221-252, 2007.
10. V. Sofronie-Stokkermans. Hierarchic reasoning in local theory extensions. In *Proc. CADE-20*, LNAI 3632, pages 219–234. Springer, 2005.
11. V. Sofronie-Stokkermans. Interpolation in local theory extensions. In *Proc. IJCAR'2006*, LNAI 4130, pages 235–250. Springer, 2006.
12. V. Sofronie-Stokkermans and C. Ihlemann. Automated reasoning in some local extensions of ordered structures. *Journal of Multiple-Valued Logics and Soft Computing* (Special issue dedicated to ISMVL'07), 13 (4-6), 397-414, 2007.
13. K.A. Spackman, K.E. Campbell, R.A. Cote. SNOMED RT: A reference terminology for health care. *Journal of the Americal Medical Informatics Association*, pages 640-644, 1997. Fall Symposium Supplement.
14. K.A. Spackman. Normal forms for description logic expression of clinical concepts in SNOMED RT. *Journal of the Americal Medical Informatics Association*, pages 627–631, 2001. Symposium Supplement.