Query Answering over DL ABoxes: How to Pick the Relevant Symbols

Franz Baader[1], Meghyn Bienvenu[2], Carsten Lutz[3], and Frank Wolter[4]

[1] TU Dresden, Germany, baader@inf.tu-dresden.de
[2] Universität Bremen, Germany, meghyn@informatik.uni-bremen.de
[3] Universität Bremen, Germany, clu@informatik.uni-bremen.de
[4] University of Liverpool, UK, wolter@liverpool.ac.uk

1 Introduction

One of the main applications of description logics (DLs) is ontology-based data access: a conceptual model of a domain is formalized using a DL TBox, and this formalization is exploited to obtain complete answers when querying data stored in an ABox. The current availability of professional and comprehensive ontologies for the bio-medical domain such as SNOMED CT, NCI, and Galen allows an easy and inexpensive adoption of this approach in bio-medical applications such as querying electronic medical records [1]. In such applications, it is typical that an "off-the-shelf" ontology such as SNOMED CT is used together with ABoxes that derive from the actual application. Since ontologies such as SNOMED CT are huge, containing more than 400.000 concept names and embracing various areas such as anatomy, diseases, medication, and even social context and geographic location, it is usually the case that many symbols (concept or role names) defined in the ontology are excluded from the signature Σ used to formulate ABoxes in the given application. Such an excluded symbol S may be linked to the symbols in Σ via the TBox and thus still be relevant for querying Σ-ABoxes, but it may also be completely unrelated to Σ and thus never contribute to deriving certain answers to queries posed against Σ-ABoxes. Clearly, symbols of the latter kind are not relevant for formulating queries in the considered application.

The aim of this paper is (i) to propose a notion of *ABox relevance* of a symbol that describes when a symbol S is relevant for ABoxes formulated in a given signature Σ, with a given background TBox $\mathcal{T}$ in place; and (ii) to study the computational complexity of deciding ABox relevance. This decision problem is of interest for a variety of reasons. First, knowing which symbols are relevant for ABox querying is useful for the construction of meaningful queries because non-relevant symbols can be discarded. When working with TBoxes that have more than 400.000 concept names such as SNOMED CT, support of this type is clearly indispensable. Second, the set of relevant symbols can be used to guide module extraction [2–4]. Recall that module extraction is the problem of extracting a subset M from a TBox T so that M can be used instead of T in a particular application. In most cases, the extraction of M is guided by a signature Σ that is of interest for the application and about which the module should "say the same" as the original TBox. If the targeted application is query answering, it is natural to use as the signature Σ the set of symbols that are relevant for ABoxes formulated in the desired ABox signature. With the right notion of 'module' at hand, the extracted module can then be used instead of the original TBox for query answering. Note that our notion of relevance is based on an ABox signature instead of on a concrete ABox. The rationale behind this is that, in typical applications, the ABox changes frequently which makes it unrealistic to assume that the set of relevant symbols is re-computed after every ABox modification, not to speak of the rather costly module extraction.

The notion of ABox relevance depends on the query language used. In this paper, we study instance queries as the simplest kind of query commonly used, and conjunctive queries due to their recent popularity in the DL community [5–13]. After introducing preliminaries in Section 2, we present our notion of ABox relevance along with some basic observations in Section 3. We then analyze the complexity of deciding relevance in the $\mathcal{EL}$ family of DLs in Section 4, showing that it ranges from polynomial to ExpTime-complete. Results on the $\mathcal{ALC}$ family of DLs are given in Section 5, showing in particular that ABox relevance is decidable in $\mathcal{ALC}$ and $\mathcal{ALCI}$, but relevance regarding instance queries is undecidable in $\mathcal{ALCF}$ and relevance regarding conjunctive queries is undecidable in $\mathcal{ALCFI}$. Some proofs are deferred to the full version [14].

2 Preliminaries

We consider various DLs throughout the paper and use standard notation for syntax, semantics, and DL names, see [15]. In particular, we use $\mathsf{N_C}$ and $\mathsf{N_R}$ to denote the sets of concept names and role names, C, D to denote (potentially) composite concepts, A, B for concept names, r, s for role names, and a, b for individual names. When we speak of a *TBox*, we mean a set of *concept inclusions* (CIs) $C \sqsubseteq D$. An *ABox* is a set of *concept assertions* $A(a)$ and $\neg A(a)$ and *role assertions* $r(a, b)$. To distinguish this kind of ABox from ABoxes that admit composite concepts in concept assertions, we sometimes use the term *literal ABox.* We use $\mathsf{Ind}(\mathcal{A})$ to denote the set of individual names used in the ABox $\mathcal{A}$. As usual in the context of query answering, we adopt the unique name assumption (UNA).

We study two query languages: (i) the set $\mathcal{IQ}$ of *instance queries*, which take the form $A(v)$; and (ii) the set $\mathcal{CQ}$ of *conjunctive queries (CQs)*, which take the form $\exists \boldsymbol{v}.\varphi(\boldsymbol{v}, \boldsymbol{u})$ where φ is a conjunction of atoms of the form $A(t)$ and $r(t, t')$ with t, t' *terms*, i.e., variables or individual names. Note that we disallow composite concepts in instance queries and conjunctive queries, which is a realistic assumption for many applications. Also note that instance queries can only be used to query concept names, but not role names. This is the traditional definition, which is due to the fact that role assertions in an ABox can only be implied by an ABox if they are explicitly contained in it (and thus querying is trivial). Given a TBox $\mathcal{T}$, an ABox $\mathcal{A}$, and a (conjunctive or instance) query q with k answer variables $v_1, \ldots, v_k$, we write $\mathcal{T}, \mathcal{A} \models q[a_1, \ldots, a_k]$ if the tuple $(a_1, \ldots, a_k)$ of individual names is a *certain answer* to q w.r.t. $\mathcal{A}$ and $\mathcal{T}$ (defined in the usual way). We use $\mathsf{cert}_{\mathcal{T},\mathcal{A}}(q)$ to denote the set of all certain answers to q w.r.t. $\mathcal{A}$ and $\mathcal{T}$.

We use the term *symbol* to refer to a concept name or role name, *signature* to refer to a set of symbols, and $\mathsf{sig}(q)$ to denote the set of symbols used in the query q. Given a signature Σ, a *Σ-ABox* (resp. *Σ-concept*) is an ABox (resp. concept) using symbols from Σ only.

3 The ABox Relevance Problem

The following definition describes the set of symbols $\Sigma^{\mathcal{L}}_{\mathcal{T}}$ that can meaningfully be used in a query posed against ABoxes that are formulated in the signature Σ, with the TBox $\mathcal{T}$ in the background.

Definition 1. *Let $\mathcal{T}$ be a TBox, Σ a signature, and $\mathcal{L} \in \{\mathcal{IQ}, \mathcal{CQ}\}$ a query language. A symbol S is* $\mathcal{L}$-relevant for Σ given $\mathcal{T}$ *if there exists a Σ-ABox and $\mathcal{L}$-query q such that $\mathcal{A}$ is consistent w.r.t. $\mathcal{T}$, $S \in \mathsf{sig}(q)$, and $\mathsf{cert}_{\mathcal{T},\mathcal{A}}(q) \neq \emptyset$. The* $\mathcal{L}$-extension *of Σ given $\mathcal{T}$ is the following signature:*

$$\Sigma^{\mathcal{L}}_{\mathcal{T}} := \Sigma \cup \{S \in \mathsf{N_C} \cup \mathsf{N_R} \mid S \textit{ is } \mathcal{L}\textit{-relevant for } \mathcal{T} \textit{ and } \Sigma\}.$$

For example, the concept name A is both $\mathcal{IQ}$- and $\mathcal{CQ}$-relevant for $\Sigma = \{r\}$ given $\mathcal{T} = \{\exists r.\top \sqsubseteq A\}$, as witnessed by the query $q = A(v)$ and Σ-ABox $\{r(a,b)\}$ since $\mathsf{cert}_{\mathcal{T},\mathcal{A}}(q) = \{a\}$. Note that $\Sigma^{\mathcal{IQ}}_{\mathcal{T}}$ can never include any role names since role names cannot occur in an instance query. We are interested in *deciding $\mathcal{L}$-relevance* for $\mathcal{L} \in \{\mathcal{IQ}, \mathcal{CQ}\}$: given a TBox $\mathcal{T}$, a signature Σ and a symbol S, decide whether $S \in \Sigma^{\mathcal{L}}_{\mathcal{T}}$. Clearly, this problem can be used to compute the signature $\Sigma^{\mathcal{L}}_{\mathcal{T}}$.

It should not be surprising that, in general, we need not have $\Sigma^{\mathcal{IQ}}_{\mathcal{T}} = \Sigma^{\mathcal{CQ}}_{\mathcal{T}}$. For example, if $\mathcal{T} = \{A \sqsubseteq \exists r.B\}$ and $\Sigma = \{A\}$, then $B \notin \Sigma^{\mathcal{IQ}}_{\mathcal{T}}$, but $B \in \Sigma^{\mathcal{CQ}}_{\mathcal{T}}$. For the former, it suffices to note that $\mathsf{cert}_{\mathcal{T},\mathcal{A}}(B(v)) = \emptyset$ for all Σ-ABoxes $\mathcal{A}$. For the latter, note that $\mathsf{cert}_{\mathcal{T},\mathcal{A}}(\exists v.B(v)) = \{()\}$ when $\mathcal{A} = \{A(a)\}$ (and where $()$ is the empty tuple representing a positive answer to the Boolean query). The following lemma, which is independent of the DL in which TBoxes are formulated, shows that we can always concentrate on CQs of such a simple form. It is an easy consequence of the fact that, since composite concepts are disallowed, CQs are purely positive, existential, and conjunctive.

Lemma 1. *$A \in \mathsf{N_C}$ (resp. $r \in \mathsf{N_R}$) is $\mathcal{CQ}$-relevant for Σ given $\mathcal{T}$ iff there is an ABox $\mathcal{A}$ with $\mathsf{cert}_{\mathcal{T},\mathcal{A}}(\exists v.A(v)) \neq \emptyset$ (resp. $\mathsf{cert}_{\mathcal{T},\mathcal{A}}(\exists v, v'.r(v,v')) \neq \emptyset$).*

Lemma 1 allows us to consider only queries of the form $\exists v.A(v)$ and $\exists v, v'.r(v,v')$ when dealing with $\mathcal{CQ}$-relevance. From now on, we do this without further notice.

Answering conjunctive queries is typically more difficult than answering instance queries, both regarding the computational complexity and the required algorithms [7, 9]. Thus, it may be a little surprising that, as stated by the following result, $\mathcal{CQ}$-relevance can be polynomially reduced to $\mathcal{IQ}$-relevance. The converse is, in general, not known. In Section 4, we will see that it holds in the $\mathcal{EL}$ family of DLs.

Theorem 1. *In any DL with (qualified) existential restrictions, $\mathcal{CQ}$-relevance can be polynomially reduced to $\mathcal{IQ}$-relevance.*

Proof (sketch). Let $\mathcal{T}$ be a TBox, Σ a signature, B a concept name that does not occur in $\mathcal{T}$ and Σ, and s a role name that does not occur in $\mathcal{T}$ and Σ. Then

1. A is $\mathcal{CQ}$-relevant for Σ given $\mathcal{T}$ iff B is $\mathcal{IQ}$-relevant for $\Sigma \cup \{s\}$ given the TBox $\mathcal{T}' = \mathcal{T} \cup \mathcal{T}_B \cup \{A \sqsubseteq B\}$, where $\mathcal{T}_B = \{\exists r.B \sqsubseteq B \mid r = s$ or r occurs in $\mathcal{T}\}$;
2. r is $\mathcal{CQ}$-relevant for Σ given $\mathcal{T}$ iff B is $\mathcal{IQ}$-relevant for $\Sigma \cup \{s\}$ given the TBox $\mathcal{T}' = \mathcal{T} \cup \mathcal{T}_B \cup \{\exists r.\top \sqsubseteq B\}$, where $\mathcal{T}_B$ is as above.

The proofs of Points 1 and 2 are similar and we concentrate on Point 1. First suppose that A is $\mathcal{CQ}$-relevant for Σ given $\mathcal{T}$. Then there is a Σ-ABox $\mathcal{A}$ such that $\mathcal{T}, \mathcal{A} \models \exists v.A(v)$. Choose an $a_0 \in \mathsf{Ind}(\mathcal{A})$ and set $\mathcal{A}' := \mathcal{A} \cup \{s(a_0, b) \mid b \in \mathsf{Ind}(\mathcal{A})\}$. Using the fact that $\mathcal{T}, \mathcal{A} \models \exists v.A(v)$ and the definition of $\mathcal{A}'$ and $\mathcal{T}'$, it can be shown that $\mathcal{T}', \mathcal{A}' \models B(a_0)$. For the converse direction, suppose that B is $\mathcal{IQ}$-relevant for $\Sigma \cup \{s\}$ given $\mathcal{T}'$. Then there is a $\Sigma \cup \{s\}$-ABox $\mathcal{A}'$ such that $\mathcal{T}', \mathcal{A}' \models B(a)$ for some $a \in \mathsf{Ind}(\mathcal{A}')$. Let $\mathcal{A}$ be obtained from $\mathcal{A}'$ by removing all assertions $s(a, b)$. Using the fact that $\mathcal{T}', \mathcal{A}' \models B(a)$ and the definition of $\mathcal{A}'$ and $\mathcal{T}'$, it can be shown that $\mathcal{T}, \mathcal{A} \models \exists v.A(v)$. □

In some proofs, it will be convenient to drop the UNA. The following lemma states that this can be done w.l.o.g. in $\mathcal{ALCI}$ (and all its fragments such as $\mathcal{EL}$ and $\mathcal{ALC}$) because the certain answers and thus also the notion of $\mathcal{L}$-relevance does not change. The lemma is easily proved using the fact that, in $\mathcal{ALCI}$, we can easily convert a model $\mathcal{I}$ for an ABox and a TBox that violates the UNA into a model $\mathcal{I}'$ that satisfies the UNA by "duplicating points" and such that $\mathcal{I}$ and $\mathcal{I}'$ are bisimilar.

Lemma 2. *Let $\mathcal{T}$ be an $\mathcal{ALCI}$-TBox, $\mathcal{A}$ an ABox, and $q \in \mathcal{L}$. Then $\mathsf{cert}_{\mathcal{T},\mathcal{A}}(q)$ is identical with and without UNA.*

An analogous statement fails, e.g., for $\mathcal{ALCF}$. To see this, take $\mathcal{T} = \{\top \sqsubseteq (\leq 1\ r) \sqcup A\}$ and $\Sigma = \{r\}$. Then A is $\mathcal{IQ}$- and $\mathcal{CQ}$-relevant with UNA due to the ABox $\{r(a, b), r(a, b')\}$, but it is not relevant without UNA.

4 The $\mathcal{EL}$ Family

We study ABox relevance in the $\mathcal{EL}$ family of lightweight DLs [16]. In particular, we show that ABox relevance in plain $\mathcal{EL}$ can be decided in polynomial time, whereas it is ExpTime-complete in $\mathcal{ELI}$ and $\mathcal{EL}_\bot$. It is interesting to contrast these results with the complexity of subsumption and instance checking, which can be decided in polynomial time in the case of $\mathcal{EL}$ and $\mathcal{EL}_\bot$ and are ExpTime-complete in $\mathcal{ELI}$.

Throughout this section, we assume that the UNA is not imposed. This can be done w.l.o.g. due to Lemma 2. Since DLs of the $\mathcal{EL}$ family do not offer negation, it may be deemed unnatural to define ABox relevance based on literal ABoxes, which admit negation. However, as the following lemma demonstrates, there is actually no difference between defining ABox relevance based on literal ABoxes and *positive* ABoxes, in which all concept assertions are of the form $A(a)$ with A a concept name. This holds for both $\mathcal{IQ}$- and $\mathcal{CQ}$-relevance. The proof is via canonical models.

Lemma 3. *For every $\mathcal{ELI}_\bot$ TBox $\mathcal{T}$, literal ABox $\mathcal{A}$ consistent w.r.t. $\mathcal{T}$, and conjunctive query q, we have $\mathsf{cert}_{\mathcal{T},\mathcal{A}}(q) = \mathsf{cert}_{\mathcal{T},\mathcal{A}^-}(q)$, where $\mathcal{A}^-$ is the restriction of $\mathcal{A}$ to assertions of the form $A(a)$ and $r(a,b)$.*

We now state the announced converse of Theorem 1. The proof proceeds by showing that A is $\mathcal{IQ}$-relevant for Σ given $\mathcal{T}$ iff B is $\mathcal{CQ}$-relevant for $\Sigma \cup \{X\}$ given the TBox $\mathcal{T}' = \mathcal{T} \cup \{A \sqcap X \sqsubseteq B\}$, where B and X are concept names that do not occur in $\mathcal{T}$.

Theorem 2. *In $\mathcal{ELI}_\bot$, $\mathcal{IQ}$-relevance can be polynomially reduced to $\mathcal{CQ}$-relevance.*

Theorem 2 allow us to choose freely between $\mathcal{IQ}$ and $\mathcal{CQ}$ when proving lower and upper bounds for relevance in the $\mathcal{EL}$ family of DLs. Note that, by the example given in Section 2, these two notions do not coincide even in $\mathcal{EL}$.

Theorem 3. *In $\mathcal{EL}$, $\mathcal{IQ}$-relevance and $\mathcal{CQ}$-relevance can be decided in* PTime.

Proof. We consider $\mathcal{IQ}$-relevance. Let $\mathcal{T}$ be an $\mathcal{EL}$-TBox and Σ a signature. Define the *total Σ-ABox* as $\mathcal{A}_\Sigma := \{A(a_\Sigma) \mid A \in \Sigma\} \cup \{r(a_\Sigma, a_\Sigma) \mid r \in \Sigma\}$.

Claim. For all concept names A, A is $\mathcal{IQ}$-relevant for Σ given $\mathcal{T}$ iff $\mathcal{T}, \mathcal{A}_\Sigma \models A(a_\Sigma)$;

Since the instance problem can be solved in polynomial time in $\mathcal{EL}$ [16], Theorem 3 is an immediate consequence of the claim.

The “if” direction of the above claim is trivial. For the “only if” direction, let A be $\mathcal{IQ}$-relevant for Σ given $\mathcal{T}$. By Lemma 3, there is a positive Σ-ABox $\mathcal{A}$ such that $\mathcal{T}, \mathcal{A} \models A(a_0)$ for some $a_0 \in \mathsf{Ind}(\mathcal{A})$. Let $\mathcal{I}$ be a model of $\mathcal{T}$ and $\mathcal{A}_\Sigma$. We have to show that $a_0^\mathcal{I} \in A^\mathcal{I}$. Modify $\mathcal{I}$ by setting $b^\mathcal{I} := a_\Sigma^\mathcal{I}$ for all individual names b. It is easy to verify that $\mathcal{I}$ is a model of the positive ABox $\mathcal{A}$ and of $\mathcal{T}$. Since $\mathcal{T}, \mathcal{A} \models A(a_0)$, we have $a_0^\mathcal{I} \in A^\mathcal{I}$ as required. □

Note that we need very little for the proof of Theorem 3 to go through: it suffices that $\mathcal{A}_\Sigma$ is consistent with every TBox and that the DL in question is monotone. It follows that for all DLs of this sort, deciding $\mathcal{IQ}$- and $\mathcal{CQ}$-relevance has the same complexity as subsumption/instance checking (whose complexity coincides for almost every DL). The upper bound is obtained as in the proof of Theorem 3, based on instance checking. For the lower bound, note that C is subsumed by D w.r.t. $\mathcal{T}$ iff B is $\mathcal{IQ}$-/$\mathcal{CQ}$-relevant for $\mathcal{T} \cup \{A \sqsubseteq C, D \sqsubseteq B\}$ and the signature $\{A\}$, where $A, B \notin \mathsf{sig}(C, D, \mathcal{T})$. We thus obtain the following result for the DL $\mathcal{ELI}$, in which subsumption and instance checking are ExpTime-complete [17].

Theorem 4. *In $\mathcal{ELI}$, $\mathcal{IQ}$-relevance and $\mathcal{CQ}$-relevance are* ExpTime*-complete.*

The simplest extension of $\mathcal{EL}$ in which the total ABox $\mathcal{A}_\Sigma$ is not consistent w.r.t. every TBox is $\mathcal{EL}_\bot$. Here, deciding relevance is significantly harder than deciding subsumption/instance checking (which can be decided in polynomial time). We start by proving an NP lower bound for a very simple fragment of $\mathcal{EL}_\bot$: let $\mathcal{L}$ be the DL that admits only CIs of the form $A \sqcap A' \sqsubseteq B$ and $A \sqcap B \sqsubseteq \bot$, with A, A', and B concept names. This is a fragment of $\mathcal{EL}_\bot$, but also of those variants of DL-Lite that admit conjunction on the left-hand side of CIs [8].

Theorem 5. *In $\mathcal{L}$, $\mathcal{IQ}$-relevance and $\mathcal{CQ}$-relevance are* NP*-hard.*

Proof. Reduction from SAT. Let φ be a propositional formula in NNF using variables $v_0, \dots, v_n$ and $\mathsf{sub}(\varphi)$ the set of subformulas of φ. Define a TBox $\mathcal{T}$ as the union of the following:

- $A_{v_i} \sqcap A_{\neg v_i} \sqsubseteq \bot$ for all $i \leq n$;
- $A_\vartheta \sqcap A_\chi \sqsubseteq A_\psi$ for all $\psi = \vartheta \wedge \chi \in \mathsf{sub}(\varphi)$;
- $A_\vartheta \sqsubseteq A_\psi,\ A_\chi \sqsubseteq A_\psi$ for all $\psi = \vartheta \vee \chi \in \mathsf{sub}(\varphi)$.

Let $\Sigma = \{A_{v_i}, A_{\neg v_i} \mid i \leq n\}$. It can be verified that A_φ is $\mathcal{IQ}$-relevant for Σ given $\mathcal{T}$ iff φ is satisfiable. □

For full $\mathcal{EL}_\bot$, Theorem 5 can be improved to an ExpTime lower bound. The idea is to make use of an existing ExpTime lower bound for deciding conservative extensions in $\mathcal{EL}/\mathcal{EL}_\bot$ established in [18]. To implement this, we first establish a technical proposition. Its proof is similar to Lemma 22 (i) in [18] and given in the full paper.

Proposition 1. *If a concept name B is $\mathcal{IQ}$-relevant for a signature Σ given an $\mathcal{EL}_\bot$-TBox $\mathcal{T}$, then there is a Σ-concept C such that C is satisfiable w.r.t. $\mathcal{T}$ and $\mathcal{T} \models C \sqsubseteq B$.*

We now prove the lower bound.

Theorem 6. *In $\mathcal{EL}_\bot$, $\mathcal{IQ}$-relevance and $\mathcal{CQ}$-relevance are* ExpTime*-hard.*

Proof. We consider $\mathcal{IQ}$-relevance. The following result can be established by carefully analyzing the reduction underlying Theorem 36 in [18]: given an $\mathcal{EL}_\bot$-TBox $\mathcal{T}$, a signature Σ, and a concept name B, it is ExpTime-hard to decide if there exist a Σ-concept C such that C is satisfiable w.r.t. $\mathcal{T}$ and $\mathcal{T} \models C \sqsubseteq B$. Thus it suffices to show that the following conditions are equivalent, for any $\mathcal{EL}_\bot$-TBox $\mathcal{T}$, signature Σ, and concept name B:

1. there exists a Σ-concept C such that C is satisfiable w.r.t. $\mathcal{T}$ and $\mathcal{T} \models C \sqsubseteq B$;
2. there exists a Σ-ABox $\mathcal{A}$ such that $(\mathcal{T}, \mathcal{A})$ is consistent and $(\mathcal{T}, \mathcal{A}) \models B(a)$ for some $a \in \mathsf{Ind}(\mathcal{A})$.

The implication from Point 1 to Point 2 is trivial and the reverse direction is established by Proposition 1. □

To prove a matching upper bound for Theorem 6, we first establish a proposition that constrains the shape of ABoxes to be considered when deciding relevance in $\mathcal{EL}_\bot$. Here and in what follows, an ABox $\mathcal{A}$ is *tree-shaped* if

1. the directed graph $(\mathsf{Ind}(\mathcal{A}), \{(a, b) \mid r(a, b) \in \mathcal{A}$ for some $r \in \mathsf{N_R}\})$ is a tree and
2. for all $a, b \in \mathsf{Ind}(\mathcal{A})$, there is at most one role name r such that $r(a, b) \in \mathcal{A}$.

The following is a simple consequence of Proposition 1.

Proposition 2. *A concept name A is $\mathcal{IQ}$-relevant for a signature Σ given an $\mathcal{EL}_\bot$-TBox $\mathcal{T}$ iff there is a tree-shaped ABox $\mathcal{A}$ such that $\mathcal{A}$ is consistent w.r.t. $\mathcal{T}$ and $\mathcal{T}, \mathcal{A} \models A(a_0)$, with a_0 the root of $\mathcal{A}$.*

For the upper bound, we use non-deterministic bottom-up automata on finite, ranked trees. Such an automaton is a tuple $\mathfrak{A} = (Q, \mathcal{F}, Q_f, \Theta)$, where Q is a finite set of *states*, $\mathcal{F}$ is a *ranked alphabet*, $Q_f \subseteq Q$ is a set of *final states*, and Θ is a set of *transition rules* of the form $f(q_1, \dots, q_n) \to q$, where $n \geq 0$, $f \in \mathcal{F}$ is of rank n, and $q_1, \dots, q_n, q \in Q$. Note that transition rules for symbols of rank 0 replace initial states.

Automata work on finite, node-labeled, ordered trees $T = (V, E, \ell)$, where V is a finite set of nodes, $E \subseteq V \times V$ is a set of edges, and ℓ is a node-labeling function the maps each node $v \in V$ with i successors to a symbol $\ell(v) \in \mathcal{F}$ of rank i. We assume an implicit total order on the successors of each node. A *run* of the automaton $\mathfrak{A}$ on T is a map $\rho : V \to Q$ such that

- $\rho(\varepsilon) \in Q_f$, with $\varepsilon \in V$ the root of T;
- for all $v \in V$ with $\ell(v) = f$ and where v has (ordered) successors $v_1, \dots, v_n$, $n \geq 0$, we have that $f(\rho(v_1), \dots, \rho(v_n)) \to \rho(v)$ is a rule in Δ.

An automaton $\mathfrak{A}$ *accepts* a tree T if there is a run of $\mathfrak{A}$ on T. We use $L(\mathfrak{A})$ to denote the set of all trees accepted by $\mathfrak{A}$. It can be computed in polynomial time whether $L(\mathfrak{A}) = \emptyset$.

Theorem 7. *In $\mathcal{EL}_\bot$, $\mathcal{IQ}$-relevance and $\mathcal{CQ}$-relevance are* ExpTime*-complete.*

Proof. Let $\mathcal{T}$ be an $\mathcal{EL}_\bot$-TBox, Σ a signature, and A_0 a concept name such that it is to be decided whether A_0 is $\mathcal{IQ}$-relevant for Σ given $\mathcal{T}$. W.l.o.g., we may assume that A_0 occurs in $\mathcal{T}$. We use $\mathsf{sub}(\mathcal{T})$ to denote the set of all subconcepts of concepts occurring in $\mathcal{T}$ and set $\Gamma := \Sigma \cup \mathsf{sub}(\mathcal{T})$. A *$\Sigma$-type* is a finite set t of concept names that occur in Σ and such that $\bigsqcap t$ is satisfiable w.r.t. $\mathcal{T}$. A *Γ-type* is a subset t of Γ such that $\bigsqcap t$ is satisfiable w.r.t. $\mathcal{T}$. Given a Γ-type t, we use $\mathsf{cl}_\mathcal{T}(t)$ to denote the set $\{C \in \Gamma \mid \mathcal{T} \models \bigsqcap t \sqsubseteq C\}$. We use $\mathsf{ex}(\mathcal{T})$ to denote the number of concepts of the form $\exists r.C$ that occur in $\mathcal{T}$ (possibly as a subconcept). Define an automaton $\mathfrak{A} = (Q, \mathcal{F}, Q_f, \Delta)$ as follows:

- $\mathcal{F} = \{\langle t, r_1, \dots, r_n \rangle \mid t \text{ a } \Sigma\text{-type}, i < \mathsf{ex}(\mathcal{T})\}$ with $\langle t, r_1, \dots, r_n \rangle$ of rank n;
- Q is the set of Γ-types;
- $Q_f = \{q \in Q \mid A_0 \in q\}$;
- Δ consists of all rules $f(q_1, \dots, q_n) \to q$ with $f = \langle t, r_1, \dots, r_n \rangle$ such that

$$q = \mathsf{cl}_\mathcal{T}(t \cup \{\exists r.C \in \mathsf{sub}(\mathcal{T}) \mid r = r_i \text{ and } C \in q_i \text{ for some } i \text{ with } 1 \leq i \leq n\}.$$

In the full version of this paper, we show that $L(\mathfrak{A}) \neq \emptyset$ iff A_0 is $\mathcal{IQ}$-relevant for Σ given $\mathcal{T}$. Since $\mathfrak{A}$ is single-exponentially large in $|\mathcal{T}|$ and the emptiness problem can be decided in polynomial time in the size of the automaton, we obtain a single-exponential-time procedure for deciding relevance in $\mathcal{EL}_\bot$. □

5 Expressive DLs

We establish some first results for ABox relevance in $\mathcal{ALC}$ and its extensions. For $\mathcal{ALCI}$, we prove decidability of $\mathcal{IQ}$- (and thus also $\mathcal{CQ}$-) relevance, and a NExpTime$^{\text{NP}}$ upper bound; for $\mathcal{ALCF}$, we prove undecidability of $\mathcal{IQ}$-relevance.

5.1 ABox Relevance in $\mathcal{ALC}$ and $\mathcal{ALCI}$

The NExpTime$^{\text{NP}}$ upper bound is based on the following theorem, which places an upper bound on the size of ABoxes that we need to consider.

Theorem 8. *Let $\mathcal{T}$ be an $\mathcal{ALCI}$-TBox. If $A \in \mathsf{N_C}$ is $\mathcal{IQ}$-relevant for $\mathcal{T}$ w.r.t. propositional ABoxes, then there is a literal Σ-ABox $\mathcal{A}$ such that $\mathcal{A}$ is consistent w.r.t. $\mathcal{T}$, $\mathcal{T}, \mathcal{A} \models A(a)$ for some $a \in \mathsf{Ind}(\mathcal{A})$, and $|\mathsf{Ind}(\mathcal{A})| \leq 2^{|\mathcal{T}|+|\Sigma|}$.*

Proof. We do not make the UNA. We consider only the case $A \in \mathsf{N_C}$, as the case $r \in \mathsf{N_R}$ is analogous. Assume that A is $\mathcal{IQ}$-relevant for Σ given $\mathcal{T}$. Then there is a literal Σ-ABox $\mathcal{A}$ such that $\mathcal{A}$ is consistent w.r.t. $\mathcal{T}$ and $\mathcal{T}, \mathcal{A} \models A(a_0)$ for some $a_0 \in \mathsf{Ind}(\mathcal{A})$. Let $\mathcal{I}$ be a model of $\mathcal{A}$ and $\mathcal{T}$, and let $\mathcal{J}$ be the filtration of $\mathcal{I}$ w.r.t. $\Gamma = \mathsf{cl}(\mathcal{T}) \cup \{A, \neg A \mid A \in \Sigma\}$, i.e., define an equivalence relation $\sim\, \subseteq \Delta^{\mathcal{I}} \times \Delta^{\mathcal{I}}$ by setting $d \sim e$ iff

$$\{C \mid C \in \Gamma \wedge d \in C^{\mathcal{I}}\} = \{C \mid C \in \Gamma \wedge e \in C^{\mathcal{I}}\}$$

and set

$$\begin{aligned}
\Delta^{\mathcal{J}} &:= \{[d] \mid d \in \Delta^{\mathcal{I}}\} \\
A^{\mathcal{J}} &:= \{[d] \mid d \in A^{\mathcal{I}}\} \\
r^{\mathcal{J}} &:= \{([d],[e]) \mid \exists d' \in [d], e' \in [e] : (d',e') \in r^{\mathcal{I}}\} \\
a^{\mathcal{J}} &:= [a^{\mathcal{I}}]
\end{aligned}$$

Clearly, $|\Delta^{\mathcal{J}}| \leq 2^{|\mathcal{T}|+|\Sigma|}$. It is routine to prove that $\mathcal{J}$ is a model of $\mathcal{T}$. Define an ABox

$$\begin{aligned}
\mathcal{A}_{\mathcal{J}} = \; &\{A(a_{[d]}) \mid A \in \Sigma \wedge [d] \in A^{\mathcal{J}}\} \cup \\
&\{\neg A(a_{[d]}) \mid A \in \Sigma \wedge [d] \in (\neg A)^{\mathcal{J}}\} \cup \\
&\{r(a_{[d]}, a_{[e]}) \mid r \in \Sigma \wedge ([d],[e]) \in r^{\mathcal{J}}\}.
\end{aligned}$$

Clearly, $\mathcal{J}$ is a model of $\mathcal{A}_{\mathcal{J}}$. Thus, $\mathcal{A}_{\mathcal{J}}$ is consistent w.r.t. $\mathcal{T}$. It remains to show that $\mathcal{T}, \mathcal{A}_{\mathcal{J}} \models A(a_{[a_0^{\mathcal{I}}]})$. Let $\mathcal{J}'$ be a model of $\mathcal{A}_{\mathcal{J}}$ and $\mathcal{T}$. Define a model $\mathcal{I}'$ from $\mathcal{J}'$ by setting $a^{\mathcal{I}'} = (a_{[a^{\mathcal{I}}]})^{\mathcal{J}'}$ for all $a \in \mathsf{Ind}(\mathcal{A})$. It is readily checked that $\mathcal{I}'$ is a model of $\mathcal{A}$ and $\mathcal{T}$, and thus $a_0 \in A^{\mathcal{I}'}$, implying $a^{\mathcal{J}'}_{[a_0^{\mathcal{I}}]} \in A^{\mathcal{J}'}$ as required. □

It is interesting to note that the bound from Theorem 8 is tight. To see this, let

$$\begin{aligned}
\mathcal{T} := \{ \qquad\qquad\qquad\qquad\qquad\qquad\qquad A &\sqsubseteq \neg P_0 \sqcap \cdots \sqcap \neg P_{n-1} \\
\exists r.(P_0 \sqcap \cdots \sqcap P_i) &\sqsubseteq \neg P_i \\
\exists r.(P_0 \sqcap \cdots \sqcap P_{i-1} \sqcap \neg P_i) &\sqsubseteq P_i \\
\exists r.((\neg P_0 \sqcup \cdots \sqcup \neg P_{i-1}) \sqcap P_i) &\sqsubseteq P_i \\
\exists r.((\neg P_0 \sqcup \cdots \sqcup \neg P_{i-1}) \sqcap \neg P_i) &\sqsubseteq \neg P_i \\
P_0 \sqcap \cdots \sqcap P_{n-1} &\sqsubseteq X \;\}
\end{aligned}$$

and $\Sigma = \{A, r\}$. Then X is relevant for $\mathcal{T}$ and Σ, but the smallest witness ABox is an r-chain of length 2^n whose last element is an instance of A. Note that an ABox that has the form of a cycle of length $< 2^n$ is inconsistent w.r.t. $\mathcal{T}$. We now use Theorem 8 to prove membership in NExpTime$^{\text{NP}}$.

Theorem 9. *In $\mathcal{ALCI}$, $\mathcal{IQ}$-relevance and $\mathcal{CQ}$-relevance are in* NExpTime$^{\text{NP}}$.

Proof. We show the result for $\mathcal{IQ}$-relevance; the upper bound for $\mathcal{CQ}$-relevance follows by Theorem 2. Consider the following nondeterministic algorithm:

Step 1: Guess a Σ-ABox $\mathcal{A}$ such that $|\mathsf{Ind}(\mathcal{A})| = 2^{|\mathcal{T}|+|\Sigma|}$.

Step 2: Use an oracle to verify that $\mathcal{A}$ is consistent with $\mathcal{T}$. Reject if not.

Step 3: For each $a \in \mathsf{Ind}(\mathcal{A})$, use an oracle to check whether $\mathcal{A} \cup \{\neg A(a)\}$ is consistent with $\mathcal{T}$. Accept if for some $a \in \mathsf{Ind}(\mathcal{A})$ the ABox $\mathcal{A} \cup \{\neg A(a)\}$ is inconsistent with $\mathcal{T}$. Otherwise reject.

If the algorithm accepts, then we have found a Σ-ABox $\mathcal{A}$ consistent with $\mathcal{T}$ which implies some assertion $A(a)$, i.e. A is $\mathcal{IQ}$-relevant for Σ given $\mathcal{T}$. Conversely, if A is $\mathcal{IQ}$-relevant, then by Theorem 8, there must be some Σ-ABox $\mathcal{A}$ with at most $2^{|\mathcal{T}|+|\Sigma|}$ individuals which is consistent with $\mathcal{T}$ and such that $\mathcal{T}, \mathcal{A} \models A(a)$ for some $a \in \mathsf{Ind}(\mathcal{A})$. We create a new Σ-ABox from $\mathcal{A}$ as follows: $\mathcal{A}' = \mathcal{A} \cup \{\top(b_i) \mid 1 \leq i \leq 2^{|\mathcal{T}|+|\Sigma|} - |\mathsf{Ind}(\mathcal{A})|\}$. By construction, $\mathcal{A}'$ has precisely $2^{|\mathcal{T}|+|\Sigma|}$ individuals, is consistent with $\mathcal{T}$, and is such that $\mathcal{A}', \mathcal{T} \models A(a)$. If $\mathcal{A}'$ is guessed in Step 1, the algorithm accepts.

We remark that in Steps 2 and 3 of the algorithm, we test the consistency of literal ABoxes that are exponentially larger than the TBox $\mathcal{T}$. Because of this, the standard precompletion approach to deciding ABox consistency w.r.t. a TBox requires only nondeterministic polynomial time (rather than the usual deterministic single-exponential time). This means that we can use an NP-oracle in Steps 2 and 3, yielding membership in NExpTime$^{\text{NP}}$. □

We conjecture that $\mathcal{IQ}$- and $\mathcal{CQ}$-relevance are actually NExpTime$^{\text{NP}}$-complete, but leave the lower bound open for now.

5.2 Relevance in $\mathcal{ALCF}$ and $\mathcal{ALCFI}$

We show that the simple addition of functional roles to $\mathcal{ALC}$ leads to undecidability of $\mathcal{IQ}$-relevance, and that the further addition of inverse roles leads to undecidability of $\mathcal{CQ}$-relevance. Both proofs are by reduction of the tiling problem of finite (but unbounded) rectangles. An instance of this problem is given by a triple $(\mathfrak{T}, H, V)$ with $\mathfrak{T}$ a non-empty, finite set of *tile types* including an *initial tile* T_{init} to be placed on the lower left corner and a *final tile* T_{final} to be placed on the upper right corner, $H \subseteq \mathfrak{T} \times \mathfrak{T}$ a *horizontal matching relation*, and $V \subseteq \mathfrak{T} \times \mathfrak{T}$ a *vertical matching relation*. A *tiling* for $(\mathfrak{T}, H, V)$ is a map $f : \{0, \ldots, n\} \times \{0, \ldots, m\} \rightarrow \mathfrak{T}$ such that $n, m \geq 0$, $f(0,0) = T_{\mathsf{init}}$, $f(n,m) = T_{\mathsf{final}}$, $(f(i,j), f(i+1,j)) \in H$ for all $i < n$, and $(f(i,j), f(i,j+1)) \in v$ for all $i < m$. It is undecidable whether a tiling problem has a tiling.

For the reduction to $\mathcal{IQ}$-relevance in $\mathcal{ALCF}$, let $(\mathfrak{T}, H, V)$ be an instance of the tiling problem with $\mathfrak{T} = \{T_1, \ldots, T_p\}$. We construct a signature Σ and a TBox $\mathcal{T}$ such that $(\mathfrak{T}, H, V)$ has a solution iff a selected concept name A is $\mathcal{IQ}$-relevant for Σ given $\mathcal{T}$. More precisely, the ABox $\mathcal{A}$ witnessing $\mathcal{IQ}$-relevance has the form of an $n \times m$-rectangle together with a tiling for $(\mathfrak{T}, H, V)$. W.l.o.g., we concentrate on solutions where T_{final} occurs nowhere else than in the upper right corner. The ABox signature is

$$\Sigma = \{T_1, \ldots, T_p, x, y\}$$

where $T_1, \ldots, T_p$ are used as concept names and x and y are *functional* role names representing horizontal and vertical adjacency of points in the rectangle. In $\mathcal{T}$, we additionally use the concept names U, R, A, Y, Z, C, where U and R mark the upper and right border of the rectangle, A is the concept name used in the instance query, and Y, Z, and C are used for technical purposes explained below. More precisely, $\mathcal{T}$ is defined as the union of the following CIs, for all $(T_i, T_j) \in H$ and $(T_i, T_\ell) \in V$:

$$\begin{aligned}
T_{\mathsf{final}} &\sqsubseteq Y \sqcap U \sqcap R \\
\exists x.(T_j \sqcap Y \sqcap U) \sqcap T_i &\sqsubseteq U \sqcap Y \\
\exists y.(T_\ell \sqcap Y \sqcap R) \sqcap T_i &\sqsubseteq R \sqcap Y \\
(\exists x.\exists y.Z \sqcap \exists y.\exists x.Z) \sqcup (\exists x.\exists y.\neg Z \sqcap \exists y.\exists x.\neg Z) &\sqsubseteq C \\
\exists x.(T_j \sqcap Y \sqcap \exists y.Y) \sqcap \exists y.(T_\ell \sqcap Y \sqcap \exists x.Y) \sqcap C \sqcap T_i &\sqsubseteq Y \\
Y \sqcap T_{\mathsf{init}} &\sqsubseteq A
\end{aligned}$$

$$\begin{aligned}
U &\sqsubseteq \forall y.\bot \\
R &\sqsubseteq \forall x.\bot \\
\exists y.\neg R &\sqsubseteq \neg R \\
\exists x.\neg U &\sqsubseteq \neg U \\
\bigsqcup_{1 \leq s < t \leq p} T_s \sqcap T_t &\sqsubseteq \bot
\end{aligned}$$

Observe that the concept name A used in the instance query occurs only once in the TBox, on the right-hand side of a CI. Taken together, the upper part of $\mathcal{T}$ ensures the existence of a tiled $n \times m$-rectangle in a witness ABox. The concept name Y is entailed at every individual name in such an ABox that is part of the rectangle. Observe that the CIs for Y enforce the horizontal and vertical matching conditions. The CI for C enforces confluence, i.e., C is entailed at an individual name a if there is an individual b that is both an x-y-successor and a y-x-successor of a. This is so because, intuitively, Z is universally quantified: if confluence fails, we can interpret C in a way such that neither of the two disjuncts in the pre-condition of the CI for C is satisfied. The following lemma is proved in the full paper.

Lemma 4. *There is a tiling for $(\mathfrak{T}, H, V)$ iff there exists a Σ-ABox $\mathcal{A}$ that is consistent with $\mathcal{T}$ and such that $\mathcal{T}, \mathcal{A} \models A(a)$ for some a.*

Undecidability of $\mathcal{IQ}$-relevance now follows directly from Lemma 4.

Theorem 10. *In $\mathcal{ALCF}$, $\mathcal{IQ}$-relevance is undecidable.*

The reduction to $\mathcal{CQ}$-relevance in $\mathcal{ALCFI}$ is very similar to the previous one. We now assume that the roles x and y are functional *and inverse functional.*

The signature Σ is as in the previous proof, and also the TBox $\mathcal{T}$ is identical except that we replace the CI with C on the right-hand side with the following one, where $\mathcal{B}$ ranges over all Boolean combinations of the concept names Z_1, Z_2, i.e., over all concepts $L_1 \sqcap L_2$ where L_i is a literal over Z_i, for $i \in \{1, 2\}$:

$$\exists x.\exists y.\mathcal{B} \sqcap \exists y.\exists x.\mathcal{B} \sqsubseteq C$$

The following lemma is proved in the full paper.

Lemma 5. *There is a tiling for* $(\mathfrak{T}, H, V)$ *iff there exists a* Σ*-ABox* $\mathcal{A}$ *that is consistent with* $\mathcal{T}$ *and such that* $\mathcal{T}, \mathcal{A} \models \exists v.A(v)$.

We thus get the desired result.

Theorem 11. *In* $\mathcal{ALCFI}$, $\mathcal{CQ}$*-relevance is undecidable.*

6 Related Work

Several notions of relevance have been previously proposed in the philosophy and artificial intelligence literatures, but they are rather different in nature from the notion of relevance we study in this paper. For example, in the area of relevant logic [19], it is an inference, rather than a symbol, which is said to be relevant, and in the work of Levy et al. [20] it is a premise of a proof which may or may not be relevant to the deduction of a given formula. Relevance of a *signature* (hence symbol) can be found in Lakemeyer's study of relevance [21], in which he defines relevance of a signature to a formula given a theory as well as relevance of two signatures to each other given a theory. However, Lakemeyer's notions of relevance are defined only for propositional logic, and even in the case of propositional theories, do not appear to bear any relationship to ABox relevance as studied in this paper.

7 Conclusion

We have introduced a new notion of relevance that describes when a symbol can be used meaningfully in queries that are posed to ABoxes formulated in a given signature, with a given background TBox in place. We have established a relatively complete picture regarding the complexity of deciding $\mathcal{IQ}$- and $\mathcal{CQ}$-relevance in the $\mathcal{EL}$ family of lightweight DLs, and some first results for DLs of the $\mathcal{ALC}$ family. Some important open questions have been pointed out in the paper, most notably the exact complexity of relevance in $\mathcal{ALC}$ and $\mathcal{ALCI}$, and the decidability of $\mathcal{CQ}$-relevance in $\mathcal{ALCF}$. Another open issue is the formulation of a notion of relevance for queries that may contain composite concepts. This is not trivial due to the possibility of using tautological concepts in the query. Finally, we are currently investigating whether the set of relevant symbols as defined in this paper can be used to obtain more efficient algorithms for module extraction.

References

1. Patel, C., Cimino, J.J., Dolby, J., Fokoue, A., Kalyanpur, A., Kershenbaum, A., Ma, L., Schonberg, E., Srinivas, K.: Matching patient records to clinical trials using ontologies. In: Proceedings of the 6th International Semantic Web Conference and 2nd Asian Semantic Web Conference (ISWC/ASWC 2007). (2007) 816–829
2. Suntisrivaraporn, B.: Module extraction and incremental classification: A pragmatic approach for ontologies. In: Proceedings of the 5th European Semantic Web Conference (ESWC 2008). (2008) 230–244
3. Konev, B., Lutz, C., Walther, D., Wolter, F.: Semantic modularity and module extraction in description logics. In: Proceedings of the 18th European Conference on Artificial Intelligence (ECAI 2008). (2008) 55–59
4. Grau, B.C., Horrocks, I., Kazakov, Y., Sattler, U.: Modular reuse of ontologies: Theory and practice. Journal of Artificial Intelligence Research **31** (2008) 273–318
5. Krötzsch, M., Rudolph, S., Hitzler, P.: Conjunctive queries for a tractable fragment of OWL 1.1. In: Proceedings of the 6th International Semantic Web Conference (ISWC 2007). (2007) 310–323
6. Calvanese, D., Eiter, T., Ortiz, M.: Answering regular path queries in expressive description logics: An automata-theoretic approach. In: Proceedings of the 22nd AAAI Conference on Artificial Intelligence (AAAI 2007). (2007) 391–396
7. Glimm, B., Horrocks, I., Lutz, C., Sattler, U.: Conjunctive query answering for the description logic SHIQ. In: Proceedings of the 20th International Joint Conference on Artificial Intelligence (IJCAI 2007). (2007) 399–404
8. Calvanese, D., Giacomo, G.D., Lembo, D., Lenzerini, M., Rosati, R.: Tractable reasoning and efficient query answering in description logics: The DL-Lite family. Journal of Automated Reasoning **39**(3) (2007) 385–429
9. Lutz, C.: The complexity of conjunctive query answering in expressive description logics. In: Proceedings of the 4th International Joint Conference on Automated Reasoning (IJCAR 2008). (2008) 179–193
10. Eiter, T., Gottlob, G., Ortiz, M., Simkus, M.: Query answering in the description logic horn-SHIQ. In: Proceedings of the 11th European Conference on Logics in Artificial Intelligence (JELIA 2008). (2008) 166–179
11. Dolby, J., Fokoue, A., Kalyanpur, A., Ma, L., Schonberg, E., Srinivas, K., Sun, X.: Scalable grounded conjunctive query evaluation over large and expressive knowledge bases. In: Proceedings of the 7th International Semantic Web Conference (ISWC 2008). (2008) 403–418
12. Glimm, B., Horrocks, I., Sattler, U.: Unions of conjunctive queries in SHOQ. In: Proceedings of the 11th International Conference on the Principles of Knowledge Representation and Reasoning (KR 2008). (2008) 252–262
13. Lutz, C., Toman, D., Wolter, F.: Conjunctive query answering in the description logic $\mathcal{EL}$ using a relational database system. In: Proceedings of the 21st International Joint Conference on Artificial Intelligence (IJCAI 2009). (2009)
14. Baader, F., Bienvenu, M., Lutz, C., Wolter, F.: Query answering over DL aboxes: How to pick the relevant symbols (2009) Available from http://www.informatik.uni-bremen.de/∼clu/papers.
15. Baader, F., McGuiness, D.L., Nardi, D., Patel-Schneider, P., eds.: The Description Logic Handbook. Cambridge University Press (2003)
16. Baader, F., Brandt, S., Lutz, C.: Pushing the $\mathcal{EL}$ envelope. In: Proceedings of the 19th International Joint Conference on Artificial Intelligence (IJCAI 2005). (2005) 364–369
17. Baader, F., Lutz, C., Brandt, S.: Pushing the $\mathcal{EL}$ envelope further. In: Proceedings of the OWLED 2008 Workshop on OWL: Experiences and Directions. (2008)

18. Lutz, C., Wolter, F.: Deciding inseparability and conservative extensions in the description logic $\mathcal{EL}$. To appear in Journal of Symbolic Computation (2009)
19. Mares, E., Meyer, R.: Relevant Logic. In: The Blackwell Guide to Philosophical Logic. Blackwell (2001)
20. Levy, A.Y., Fikes, R., Sagiv, Y.: Speeding up inferences using relevance reasoning: A formalism and algorithms. Artificial Intelligence **97**(1-2) (1997) 83–136
21. Lakemeyer, G.: Relevance from an epistemic perspective. Artificial Intelligence **97**(1-2) (1997) 137–167