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 6, 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 longer proofs have been moved to the appendix.

2 Preliminaries

We consider various DLs throughout the paper and use standard notation for syntax, semantics, and DL names, see [14]. 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 $\mathcal{T}_B = \{\exists r.B \sqsubseteq B \mid r = s \text{ or } r \text{ occurs in } \mathcal{T}\}$;
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 $\mathcal{T}, \mathcal{A} \models \exists v.A(v)$. $\square$

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 [15]. 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. We start with a technical lemma that will be useful for several proofs later on. The lemma applies to ABoxes which are potentially infinite and *positive*, i.e. in which all concept assertions are of the form $A(a)$ with A a concept name.

Lemma 3. *For every $\mathcal{ELI}_\bot$-TBox $\mathcal{T}$ and positive, potentially infinite ABox $\mathcal{A}$ consistent w.r.t. $\mathcal{T}$, there is a model $\mathcal{I}_{\mathcal{T},\mathcal{A}}$ of $\mathcal{T}$ and $\mathcal{A}$ such that the following conditions are satisfied:*

1. *for any $a \in \mathsf{Ind}(\mathcal{A})$ and $\mathcal{ELI}_\bot$-concept C, $a^{\mathcal{I}_{\mathcal{T},\mathcal{A}}} \in C^{\mathcal{I}_{\mathcal{T},\mathcal{A}}}$ iff $\mathcal{T},\mathcal{A} \models C(a)$;*
2. *for any k-ary conjunctive query q and $(a_1,\dots,a_k) \in \mathsf{N}_\mathsf{I}^k$, $\mathcal{I}_{\mathcal{T},\mathcal{A}} \models q[a_1,\dots,a_k]$ iff $(a_1,\dots,a_k) \in \mathsf{cert}_{\mathcal{T},\mathcal{A}}(q)$.*

Proof. Let $\mathcal{T}$ be an $\mathcal{ELI}_\bot$-TBox and $\mathcal{A}$ an ABox such that $\mathcal{A}$ is consistent w.r.t. $\mathcal{T}$. For $a \in \mathsf{Ind}(\mathcal{A})$, a *path* for $\mathcal{A}$ and $\mathcal{T}$ is a finite sequence $a\, r_1\, C_1\, r_2 C_2 \cdots r_n\, C_n$, $n \geq 0$, where the C_i are concepts from $\mathcal{T}$ (probably occurring as a subconcept) and the r_i are roles such that the following conditions are satisfied:

- $a \in \mathsf{Ind}(\mathcal{A})^\mathcal{I}$,
- $\mathcal{T},\mathcal{A} \models \exists r_1.C_1(a)$ of $n \geq 1$
- for $1 \leq i < n$, $\mathcal{T} \models C_i \sqsubseteq \exists r_{i+1}.C_{i+1}$.

We use $\mathsf{paths}(\mathcal{T},\mathcal{A})$ to denote the set of all paths for $\mathcal{A}$ and $\mathcal{T}$. If $p \in \mathsf{paths}(\mathcal{T},\mathcal{A}) \setminus \mathsf{Ind}(\mathcal{A})$, then $\mathsf{tail}(p)$ denotes the last concept C_n in p. The *canonical model* $\mathcal{I}_{\mathcal{T},\mathcal{A}}$ of $\mathcal{T}$ and $\mathcal{A}$ is defined as follows:

$$
\begin{aligned}
\Delta^{\mathcal{I}_{\mathcal{T},\mathcal{A}}} &:= \mathsf{paths}(\mathcal{T},\mathcal{A})\\
A^{\mathcal{I}_{\mathcal{T},\mathcal{A}}} &:= \{a \in \mathsf{Ind}(\mathcal{A}) \mid \mathcal{T},\mathcal{A} \models A(a)\} \cup\\
&\quad\ \{p \in \mathsf{paths}(\mathcal{T},\mathcal{A}) \setminus \mathsf{Ind}(\mathcal{A}) \mid \mathcal{T} \models \mathsf{tail}(p) \sqsubseteq A\}\\
r^{\mathcal{I}_{\mathcal{T},\mathcal{A}}} &:= \{(p,q) \in \mathsf{paths}(\mathcal{T},\mathcal{A}) \times \mathsf{paths}(\mathcal{T},\mathcal{A}) \mid q = p \cdot r\, C \text{ for some concept } C\}\\
a^{\mathcal{I}_{\mathcal{T},\mathcal{A}}} &:= a \text{ for all } a \in \mathsf{Ind}(\mathcal{T},\mathcal{A})
\end{aligned}
$$

It is standard to verify that $\mathcal{I}_{\mathcal{T},\mathcal{A}}$ satisfies the stated properties. □

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. This holds for both $\mathcal{IQ}$- and $\mathcal{CQ}$-relevance and allows us henceforth to restrict our attention to positive ABoxes when working with $\mathcal{ELI}_\bot$ and its fragments.

Lemma 4. *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 $r(a,b)$.*

Proof. Since "$\supseteq$" is trivial, we concentrate on "$\subseteq$". Suppose $(a_1,\dots,a_k) \notin \mathsf{cert}_{\mathcal{T},\mathcal{A}^-}(q)$. Then there is a model $\mathcal{I}$ of $\mathcal{T}$ and $\mathcal{A}^-$ such that $\mathcal{I} \not\models q[a_1,\dots,a_k]$. By Point 2 of Lemma 3, $\mathcal{I}_{\mathcal{T},\mathcal{A}^-} \not\models q[a_1,\dots,a_k]$. To prove that $(a_1,\dots,a_k) \notin \mathsf{cert}_{\mathcal{T},\mathcal{A}^-}(q)$, it thus suffices to show that $\mathcal{I}_{\mathcal{T},\mathcal{A}}$ satisfies all negative concept assertions in $\mathcal{A}$. Let $\neg A(a) \in \mathcal{A}$. Since $\mathcal{A}$ is consistent w.r.t. $\mathcal{T}$, $\mathcal{T},\mathcal{A} \not\models A(a)$. By Point 1 of Lemma 3, $a^{\mathcal{I}_{\mathcal{T},\mathcal{A}^-}} \notin A^{\mathcal{I}_{\mathcal{T},\mathcal{A}^-}}$ and we are done. □

We now state the announced converse of Theorem 1.

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

Proof. We claim 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 $\mathcal{T}$.

For the "if" direction, assume that B is $\mathcal{CQ}$-relevant for $\Sigma \cup \{X\}$ given $\mathcal{T}'$ and let $\mathcal{A}'$ be a positive $\Sigma \cup \{X\}$-ABox such that $\mathcal{T}', \mathcal{A}' \models \exists v.B(v)$. By Point 2 of Lemma 3, $\mathcal{I}_{\mathcal{T}',\mathcal{A}'} \models \exists v.B(v)$. We want to show that there is an $a \in \mathsf{Ind}(\mathcal{A}')$ with $a^{\mathcal{I}_{\mathcal{T}',\mathcal{A}'}} \in B^{\mathcal{I}_{\mathcal{T}',\mathcal{A}'}}$. Assume to the contrary that there is no such a. Let $\mathcal{I}$ be obtained from $\mathcal{I}_{\mathcal{T}',\mathcal{A}'}$ by setting

$$\begin{aligned} X^{\mathcal{I}} &:= \{a^{\mathcal{I}_{\mathcal{T}',\mathcal{A}'}} \mid a \in \mathsf{Ind}(\mathcal{A}')\} \\ B^{\mathcal{I}} &:= B^{\mathcal{I}_{\mathcal{T}',\mathcal{A}'}} \cap X^{\mathcal{I}} \end{aligned}$$

It is easy to see that $\mathcal{I}$ is still a model of $\mathcal{T}'$ and $\mathcal{A}'$. By our assumption that there is no $a \in \mathsf{Ind}(\mathcal{A}')$ with $a^{\mathcal{I}_{\mathcal{T}',\mathcal{A}'}} \in B^{\mathcal{I}_{\mathcal{T}',\mathcal{A}'}}$, we have $B^{\mathcal{I}} = \emptyset$, in contradiction to $\mathcal{T}', \mathcal{A}' \models \exists v.B(v)$. Thus, the desired $a \in \mathsf{Ind}(\mathcal{A}')$ exists. By Point 1 of Lemma 3, $a^{\mathcal{I}_{\mathcal{T}',\mathcal{A}'}} \in B^{\mathcal{I}_{\mathcal{T}',\mathcal{A}'}}$ implies that $\mathcal{T}', \mathcal{A}' \models B(a)$. By definition of $\mathcal{T}'$, this implies $\mathcal{T}', \mathcal{A}' \models A(a)$. Again by definition of $\mathcal{T}'$, this clearly implies $\mathcal{T}, \mathcal{A} \models A(a)$, where $\mathcal{A}$ is obtained from $\mathcal{A}'$ by dropping all concept assertions of the form $X(b)$. Since $\mathcal{A}$ is a Σ-ABox and consistent w.r.t. $\mathcal{T}$ (since $\mathcal{A}'$ is consistent w.r.t. $\mathcal{T}$'), it witnesses that A is $\mathcal{IQ}$-relevant for Σ given $\mathcal{T}$.

For the "only if" direction, assume that A is $\mathcal{IQ}$-relevant for Σ given $\mathcal{T}$ and let $\mathcal{A}$ be a positive Σ-ABox such that $\mathcal{T}, \mathcal{A} \models A(a)$ for some $a \in \mathsf{Ind}(\mathcal{A})$. Set $\mathcal{A}' := \mathcal{A} \cup \{X(a)\}$. It is easy to see that $\mathcal{T}', \mathcal{A}' \models \exists v.B(v)$ and thus B is $\mathcal{CQ}$-relevant for $\Sigma \cup \{X\}$ given $\mathcal{T}'$. $\square$

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 $\mathcal{EL}$ [15], 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 4, 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. $\square$

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 [16].

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 $v_0, \ldots, v_n$ 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. $\square$

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 [17]. To implement this, we first establish a technical proposition. Its proof is similar to Lemma 22 (i) in [17] and given in Appendix A.

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 [17]: 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 $(\mathsf{Ind}(\mathcal{A}), \{(a, b) \mid r(a, b) \in \mathcal{A} \text{ for some } r \in \mathsf{N_R}\})$ 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 $f(q_1, \ldots, q_n) \rightarrow q$, where $n \geq 0$, $f \in \mathcal{F}$ is of rank n, and $q_1, \ldots, q_n, q \in Q$. 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 $\rho : V \rightarrow Q$ 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 $v_1, \ldots, v_n$, $n \geq 0$, we have that $f(\rho(v_1), \ldots, \rho(v_n)) \rightarrow \rho(v)$ 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 $L(\mathfrak{A}) = \emptyset$.

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:

– $\mathcal{F} = \{\langle t, r_1, \ldots, r_n \rangle \mid t \text{ a } \Sigma\text{-type}, i < \mathsf{ex}(\mathcal{T})\}$ with $\langle t, r_1, \ldots, r_n \rangle$ 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 $f(q_1, \ldots, q_n) \to q$ with $f = \langle t, r_1, \ldots, r_n \rangle$ such that

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

A proof of the following lemma is given in Appendix A.

Lemma 5. *$L(\mathfrak{A}) \neq \emptyset$ 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 $\mathcal{EL}_\bot$.

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

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; and for $\mathcal{ALCFI}$, we prove undecidability of $\mathcal{CQ}$-relevance.

5.1 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 $a_{[a_0^{\mathcal{I}}]}^{\mathcal{J}'} \in A^{\mathcal{J}'}$ as required. □

It is interesting to note that the bound from Theorem 9 is tight. To see this, let

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

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 9 to prove membership in $\text{NExpTime}^{\text{NP}}$.

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

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 9, 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 $\textsc{NExpTime}^{\text{NP}}$. $\square$

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

5.2 $\mathcal{IQ}$-relevance in $\mathcal{ALCF}$

We show that the simple addition of functional roles to $\mathcal{ALC}$ leads to undecidability of $\mathcal{IQ}$-relevance, using a reduction of the following tiling problem. An instance 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 $f : \{0, \dots, n\} \times \{0, \dots, m\} \to \mathfrak{T}$ 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, let $(\mathfrak{T}, H, V)$ be an instance of the tiling problem with $\mathfrak{T} = \{T_1, \dots, T_p\}$. 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

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

where $T_1, \dots, T_p$ 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 $(T_i, T_\ell) \in V$:

$$\begin{aligned}
T_{\mathsf{final}} &\sqsubseteq Y \sqcap U \sqcap R \\
\exists x.(T_j \sqcap Y \sqcap U) \sqcap T_i &\sqsubseteq U \sqcap Y \\
\exists y.(T_\ell \sqcap Y \sqcap R) \sqcap T_i &\sqsubseteq R \sqcap Y \\
(\exists x.\exists y.Z \sqcap \exists y.\exists x.Z) \sqcup (\exists x.\exists y.\neg Z \sqcap \exists y.\exists x.\neg Z) &\sqsubseteq C \\
\exists x.(T_j \sqcap Y \sqcap \exists y.Y) \sqcap \exists y.(T_\ell \sqcap Y \sqcap \exists x.Y) \sqcap C \sqcap T_i &\sqsubseteq Y \\
Y \sqcap T_{\mathsf{init}} &\sqsubseteq A \\
U &\sqsubseteq \forall y.\bot \\
R &\sqsubseteq \forall x.\bot \\
\exists y.\neg R &\sqsubseteq \neg R \\
\exists x.\neg U &\sqsubseteq \neg U \\
\bigsqcup_{1 \leq s < t \leq p} T_s \sqcap T_t &\sqsubseteq \bot
\end{aligned}$$

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 is proved in Appendix A.

Lemma 6. *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 follows directly from Lemma 6.

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

$\mathcal{CQ}$-relevance in $\mathcal{ALCFI}$

We prove undecidability by reducing the same tiling problem as in the previous section, using a very similar reduction. Let $(\mathfrak{T}, H, V)$ be an instance of the tiling problem with $\mathfrak{T} = \{T_1, \ldots, T_p\}$. As before, 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 $\mathcal{CQ}$-relevant for Σ given $\mathcal{T}$, i.e., if $\mathsf{cert}_{\mathcal{T},\mathcal{A}}(\exists v.A(v)) \neq \emptyset$ for some Σ-ABox $\mathcal{A}$. We now assume that the roles x and y are functional *and inverse functional*. The signature Σ is as in the previous section, and also the TBox 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 Appendix A.

Lemma 7. *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 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 [18], it is an inference, rather than a symbol, which is said to be relevant, and in the work of Levy et al. [19] 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 [20], 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., McGuiness, D.L., Nardi, D., Patel-Schneider, P., eds.: The Description Logic Handbook. Cambridge University Press (2003)
15. 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
16. 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)
17. Lutz, C., Wolter, F.: Deciding inseparability and conservative extensions in the description logic $\mathcal{EL}$. To appear in Journal of Symbolic Computation (2009)
18. Mares, E., Meyer, R.: Relevant Logic. In: The Blackwell Guide to Philosophical Logic. Blackwell (2001)
19. 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
20. Lakemeyer, G.: Relevance from an epistemic perspective. Artificial Intelligence **97**(1-2) (1997) 137–167

A Missing Proofs

Proof of Proposition 1. Let A be $\mathcal{IQ}$-relevant for Σ given $\mathcal{T}$. We start by establishing the following technical claim.

Claim. If Γ is an (infinite) set of $\mathcal{EL}_\bot$ concepts that is closed under conjunction and satisfiable w.r.t. a TBox $\mathcal{T}$, then there is a model $\mathcal{I}$ of $\mathcal{T}$ and a $d_\mathcal{I} \in \Delta^\mathcal{I}$ such that for all concepts D, $d_\mathcal{I} \in D^\mathcal{I}$ iff $\mathcal{T} \models C \sqsubseteq D$ for some $C \in \Gamma$.

To prove the claim, choose an individual name b and let $\mathcal{B}$ be the infinite ABox $\{D(b) \mid D \in \Gamma\}$. By Lemma 3, there is a model $\mathcal{I}_{\mathcal{T},\mathcal{B}}$ of $\mathcal{T}$ and $\mathcal{I}$ such that for any $\mathcal{EL}_\bot$ concept D, we have $b^{\mathcal{I}_{\mathcal{T},\mathcal{B}}} \in D^{\mathcal{I}_{\mathcal{T},\mathcal{B}}}$ iff $\mathcal{T},\mathcal{A} \models D(b)$. By definition of $\mathcal{B}$, closure of Γ under conjunction and compactness, we further have $\mathcal{T},\mathcal{A} \models D(b)$ iff $\mathcal{T} \models C \sqsubseteq D$ for some $C \in \Gamma$. This establishes the claim since we can choose $\mathcal{I} = \mathcal{I}_{\mathcal{T},\mathcal{B}}$ and $d_\mathcal{I} = b^{\mathcal{I}_{\mathcal{T},\mathcal{A}}}$.

Now choose an ABox $\mathcal{A}$ and individual name $a_0 \in \mathsf{Ind}(\mathcal{A})$ such that $\mathcal{T},\mathcal{A} \models B(a_0)$. Assume to the contrary of what is to be shown that there does not exist a Σ-concept C which is satisfiable w.r.t. $\mathcal{T}$ and such that $\mathcal{T} \models C \sqsubseteq B$. For each individual name a in $\mathcal{A}$, let t_a denote the set of Σ-concepts C such that $\mathcal{T},\mathcal{A} \models C(a)$ and let $\mathcal{I}_a$ be a model as in the above claim with $\Gamma = t_a$. By our assumption and the choice of $\mathcal{I}_a$, we have $d_{\mathcal{I}_{a_0}} \notin B^{\mathcal{I}_{a_0}}$.

We may assume that the $\mathcal{I}_a$ are mutually disjoint. Take the following union $\mathcal{I}$ of the models $\mathcal{I}_a$:

- $\Delta^\mathcal{I} = \bigcup_{a \in \mathsf{Ind}(\mathcal{A})} \Delta^{\mathcal{I}_a}$;
- $A^\mathcal{I} = \bigcup_{a \in \mathsf{Ind}(\mathcal{A})} A^{\mathcal{I}_a}$, for $A \in \mathsf{N_C}$;
- $r^\mathcal{I} = \bigcup_{a \in \mathsf{Ind}(\mathcal{A})} r^{\mathcal{I}_a} \cup \{(d_{\mathcal{I}_a}, d_{\mathcal{I}_b}) \mid r(a,b) \in \mathcal{A}\}$, for $r \in \mathsf{N_R}$;
- $a^\mathcal{I} = d_{\mathcal{I}_a}$, for $a \in \mathsf{Ind}(\mathcal{A})$.

For all $\mathcal{EL}$-concepts C, $a \in \mathsf{Ind}(\mathcal{A})$, and $d \in \Delta^{\mathcal{I}_a}$, we have

$(*)$ $d \in C^{\mathcal{I}_a}$ iff $d \in C^\mathcal{I}$.

The proof is by induction on the structure of C. The only interesting case is $C = \exists r.D$ and the direction from right to left. Assume $d \in C^\mathcal{I} \cap \Delta^{\mathcal{I}_a}$. For $d \neq d_{\mathcal{I}_a}$, $d \in C^{\mathcal{I}_a}$ follows immediately by IH. Assume $d = d_{\mathcal{I}_a}$. Take d' with $(d, d') \in r^\mathcal{I}$ and $d' \in D^\mathcal{I}$. Again, if $d' \in \Delta^{\mathcal{I}_a}$, then the claim follows immediately from the IH. Now assume $d' \notin \Delta^{\mathcal{I}_a}$. Then $d' = b$ for some b with $r(a,b) \in \mathcal{A}$. By IH, $d' \in D^{\mathcal{I}_b}$. Hence $\mathcal{T} \models E \sqsubseteq D$ for some $E \in t_b$. Then $\mathcal{T},\mathcal{A} \models E(b)$, Since $r(a,b) \in \mathcal{A}$, we obtain $\mathcal{T},\mathcal{A} \models \exists r.D(a)$. Therefore, $\exists r.D \in t_a$. Thus, the choice of $\mathcal{I}_a$ yields $d_{\mathcal{I}_a} \in C^{\mathcal{I}_a}$ as required.

By $(*)$ and since the $\mathcal{I}_a$ are models for $\mathcal{T}$, $\mathcal{I}$ is also a model of $\mathcal{T}$. Moreover, $\mathcal{I}$ by definition satisfies all role assertions in $\mathcal{A}$ and all concept assertions are satisfied by $(*)$ and since $A(a) \in \mathcal{A}$ implies $A \in t_a$ and thus $d_\mathcal{I} \in A^\mathcal{I}$. Since $d_{\mathcal{I}_{a_0}} \notin B^{\mathcal{I}_{a_0}}$, $(*)$ also yields $a_0^\mathcal{I} \notin B^\mathcal{I}$. Summing up, we have shown that $\mathcal{T},\mathcal{A} \not\models B(a_0)$, which is a contradiction. $\square$

Proof of Lemma 5. For the "if" direction, assume that A_0 is relevant for Σ given $\mathcal{T}$. By Proposition 2, there is a tree-shaped and consistent ABox $\mathcal{A}$ with root a_0 such that $\mathcal{T}, \mathcal{A} \models A_0(a_0)$. Define a mapping ρ that maps each $a \in \mathsf{Ind}(\mathcal{A})$ to the Γ-type $\rho(a) := \{C \in \Gamma \mid \mathcal{T}, \mathcal{A} \models C(a)\}$. We now inductively choose a subset $R \subseteq \mathsf{Ind}(\mathcal{A})$ as follows:

- Initially, set $R := \{a_0\}$;
- if $a \in R$ is not a leaf in $\mathcal{A}$ and no successors of a in $\mathcal{A}$ are yet in R, then choose for each $\exists r.C \in \rho(a)$ an $a' \in \mathsf{Ind}(\mathcal{A})$ with $r(a, a') \in \mathcal{A}$ and $C \in \rho(a')$ if such an a' exists, and add a' to R.

For each node in $\mathcal{A}$, fix a total order on the successors that are in R (call this an "R-successor"). For $a \in \mathsf{Ind}(\mathcal{A})$, we use $\sigma_{\mathcal{A}}(a)$ to denote the set $\{A \mid A(a) \in \mathcal{A}\}$. Define a tree $T = (V, E, \ell)$ as follows:

- $V = R$;
- $E = \{(a, b) \in V \times V \mid r(a, b) \in \mathcal{A}\}$ and the order of successor in T agrees with the chosen order on R-successors in $\mathcal{A}$;
- $\ell(a) = \langle \sigma_{\mathcal{A}}(a), r_1, \ldots, r_n \rangle$ where r_i is the (unique!) role such that $r_i(a, a_i) \in \mathcal{A}$, with a_i the i-th R-successor of a .

Let ρ' be the restriction of ρ to R. We show that r' is a run of $\mathfrak{A}$ on T. First, note that $\mathcal{T}, \mathcal{A} \models A_0(a_0)$ implies $A_0 \in \rho'(a_0)$, and thus $\rho'(a_0) \in Q_f$ (observe that a_0 is the root of T). It thus remains to show that for all $a \in V$ with successors $a_1, \ldots, a_n$ that are connected in $\mathcal{A}$ via $r_1, \ldots, r_n$, respectively, we have $\rho'(a) = \Omega$, where

$$\Omega = \mathsf{cl}_{\mathcal{T}}(\sigma_{\mathcal{A}}(a) \cup \{\exists r.C \in \mathsf{sub}(\mathcal{T}) \mid r = r_i \text{ and } C \in \rho'(a_i) \text{ for some } 1 \leq i \leq n\}).$$

The "$\supseteq$" direction is immediate by the definitions of ρ', $\sigma_{\mathcal{A}}$, and $\mathsf{cl}_{\mathcal{T}}$ and the semantics. For the "$\subseteq$" direction, assume that $C_0 \in \Gamma \setminus \Omega$. Let $b_1, \ldots, b_m$ be the successors of a in $\mathcal{A}$ (including those not in R) and for $1 \leq i \leq m$, let $\mathcal{A}_i$ be the restriction of $\mathcal{A}$ to b_i and all individual names that are reachable from it in $\mathcal{A}$. The proof of the following property is by a relatively straightforward model-theoretic construction based on canonical models. Details are left to the reader.

$(*)$ for all $C \in \Gamma$ and $b \in \mathsf{Ind}(\mathcal{A}_i)$, we have $\mathcal{T}, \mathcal{A} \models C(b)$ iff $\mathcal{T}, \mathcal{A}_i \models C(b)$.

To continue our argument, take the interpretations $\mathcal{I}_{\mathcal{T},\mathcal{A}}$ and $\mathcal{I}_{\mathcal{T},\mathcal{A}_1}, \ldots, \mathcal{I}_{\mathcal{T},\mathcal{A}_m}$. Assume w.l.o.g. that their domains are pairwise disjoint. Define a new interpretation $\mathcal{J}$ as follows:

- $\Delta^{\mathcal{J}} = \Delta^{\mathcal{I}_{\mathcal{T},\mathcal{A}}} \uplus \Delta^{\mathcal{I}_{\mathcal{T},\mathcal{A}_1}} \uplus \cdots \uplus \Delta^{\mathcal{I}_{\mathcal{T},\mathcal{A}_m}} \uplus \{a\}$;
- $A^{\mathcal{J}} = A^{\mathcal{I}_{\mathcal{T},\mathcal{A}}} \cup A^{\mathcal{I}_{\mathcal{T},\mathcal{A}_1}} \cup \cdots \cup A^{\mathcal{I}_{\mathcal{T},\mathcal{A}_m}} \cup S$, where $S = \{a\}$ if $A \in \Omega$ and $S = \emptyset$ otherwise;
- $r^{\mathcal{J}} = r^{\mathcal{I}_{\mathcal{T},\mathcal{A}}} \cup r^{\mathcal{I}_{\mathcal{T},\mathcal{A}_1}} \cup \cdots \cup r^{\mathcal{I}_{\mathcal{T},\mathcal{A}_m}} \cup \{(d, a) \mid (d, a^{\mathcal{I}_{\mathcal{T},\mathcal{A}}}) \in r^{\mathcal{I}_{\mathcal{T},\mathcal{A}}}\} \cup \{(a, b_i^{\mathcal{I}_{\mathcal{T},\mathcal{A}_i}}) \mid 1 \leq i \leq m \text{ and } r(a, b_i) \in \mathcal{A}\}$;
- $a^{\mathcal{J}} = a$;

– for all $b \in \mathsf{Ind}(\mathcal{A}) \setminus \{a\}$, $b^{\mathcal{J}} = b^{\mathcal{I}_{\mathcal{T},\mathcal{A}_i}}$ if $b \in \mathsf{Ind}(\mathcal{A}_i)$ and $b^{\mathcal{J}} = b^{\mathcal{I}_{\mathcal{T},\mathcal{A}}}$ if $b \notin \mathsf{Ind}(\mathcal{A}_1) \cup \cdots \cup \mathsf{Ind}(\mathcal{A}_m)$.

Observe that the interpretation of individual names is well defined since $\mathcal{A}$ is tree-shaped, which implies that the sets $\mathsf{Ind}(\mathcal{A}_i)$ are pairwise disjoint.

Claim. For all i with $1 \leq i \leq m$ and all $C \in \Gamma$, we have:

1. $d \in C^{\mathcal{J}}$ iff $d \in C^{\mathcal{I}_{\mathcal{T},\mathcal{A}_i}}$ for all $d \in \Delta^{\mathcal{I}_{\mathcal{T},\mathcal{A}_i}}$;
2. $a \in C^{\mathcal{J}}$ iff $C \in \Omega$;
3. $d \in C^{\mathcal{J}}$ iff $d \in C^{\mathcal{I}_{\mathcal{T},\mathcal{A}}}$ for all $d \in \Delta^{\mathcal{I}_{\mathcal{T},\mathcal{A}}}$.

All points are proved by induction on the structure of C. For Point 1, this is straightforward. For Point 2, the induction start is immediate by definition of $\mathcal{J}$. The induction step is trivial when C is a conjunction. Thus, let $C = \exists r.D$. If $a \in C^{\mathcal{J}}$, then there is a $d \in D^{\mathcal{J}}$ with $(a,d) \in r^{\mathcal{J}}$. By definition of $\mathcal{J}$, we have $d = b_i^{\mathcal{I}_{\mathcal{T},\mathcal{A}_i}}$ for some i with $1 \leq i \leq m$ and $r(a,b_i) \in \mathcal{A}$. By Point 1 of the claim and Point 1 of Lemma 3 and $(*)$, $\mathcal{T},\mathcal{A} \models D(b_i)$. By definition of R, there is a j with $1 \leq j \leq n$, $r_j = r$, and $\mathcal{T},\mathcal{A} \models D(a_j)$. It follows that $\exists r.D \in \Omega$ as required. For the converse direction, let $\exists r.D \in \Omega$. Then there is an i with $1 \leq i \leq n$, $r_i = r$, and $\mathcal{T},\mathcal{A} \models D(a_i)$. By $(*)$, Point 1 of Lemma 3, and Point 1 of the claim, we have $a_i^{\mathcal{J}} \in D^{\mathcal{J}}$. By definition of $r^{\mathcal{J}}$, we have $(a,a_i) \in r^{\mathcal{J}}$, thus $a \in (\exists r.D)^{\mathcal{J}}$ as required. In the induction proof of Point 3, only the "only if" direction of the case $C = \exists r.D$ is non-trivial. Let $(d,e) \in r^{\mathcal{J}}$ and $e \in D^{\mathcal{J}}$. If $e \neq a$, we only need to apply the induction hypothesis and use the semantics. Now assume that $e = a$. By Point 2, we get $D \in \Omega$, thus $\mathcal{T},\mathcal{A} \models D(a)$. It follows that $a^{\mathcal{I}_{\mathcal{T},\mathcal{A}}} \in D^{\mathcal{I}_{\mathcal{T},\mathcal{A}}}$. Since $(d,a) \in r^{\mathcal{J}}$, we have $(d, a^{\mathcal{I}_{\mathcal{T},\mathcal{A}}}) \in r^{\mathcal{I}_{\mathcal{T},\mathcal{A}}}$. Thus, $d \in (\exists r.D)^{\mathcal{I}_{\mathcal{T},\mathcal{A}}}$ by the semantics.

Since $\mathcal{I}_{\mathcal{T},\mathcal{A}}$ and the $\mathcal{I}_{\mathcal{T},\mathcal{A}_i}$ are models of $\mathcal{T}$ and Ω is closed under $\mathcal{T}$-consequence, the claim implies that $\mathcal{J}$ is a model of $\mathcal{T}$. Using the definition of $\mathcal{J}$, it is not hard to verify that $\mathcal{J}$ is also a model of $\mathcal{A}$. Since $C_0 \notin \Omega$, we have $a^{\mathcal{J}} \notin C_0^{\mathcal{J}}$ by Point 2 of the claim. It follows that $\mathcal{T},\mathcal{A} \not\models C_0(a)$. By definition of ρ', we obtain $C_0 \notin \rho'(a)$ as required.

For the "only if" direction, let $T = (V,E,\ell)$ be a tree accepted by $\mathcal{A}$, and ρ be a run of $\mathfrak{A}$ on T. Define an ABox

$$\begin{aligned}\mathcal{A} := \{&A(a_v) \mid v \in V \text{ and } \ell(v) = \langle t, r_1, \ldots, r_n\rangle \text{ with } A \in t\} \cup \\ &\{r(a_v, a_{v_i}) \mid v_i \text{ is } i\text{-th successor of } v \text{ and } \ell(v) = \langle t, r_1, \ldots, r_n\rangle \text{ with } r_i = r\}.\end{aligned}$$

We want to show that $\mathcal{A}$ witnesses the $\mathcal{IQ}$-relevance of A_0 given $\mathcal{T}$. We begin by proving the consistency of $\mathcal{A}$ with respect to $\mathcal{T}$. Let us define Ψ as the set of concepts C which are satisfiable w.r.t. $\mathcal{T}$ and such that $\exists r.C \in \Gamma$ for some role r. For each $C \in \Psi$, we let $\mathcal{J}_C$ be the canonical model of the ABox $\{B(b)\}$ and TBox $\mathcal{T} \cup \{B \equiv C\}$, and let x_C be the element in $\Delta^{\mathcal{J}_C}$ with $b^{\mathcal{J}_C} = x_C$. Suppose w.l.o.g. that the universes of the $\mathcal{J}_C$ are all disjoint. We use the interpretations

$\mathcal{J}_C$ to construct a new interpretation $\mathcal{I}$ as follows:

$$\Delta^{\mathcal{I}} = V \cup \bigcup_{C \in \Psi} \Delta^{\mathcal{J}_C}$$

$$A^{\mathcal{I}} = \{v \in V \mid A \in \rho(v)\} \cup \bigcup_{C \in \Psi} A^{\mathcal{J}_C}$$

$$r^{\mathcal{I}} = \{(v, w) \in E \mid w \text{ is } i\text{-th successor of } v \text{ and } \ell(v) = \langle t, r_1, ..., r_n \rangle \text{ where } r_i = r\}$$
$$\cup \{(v, x_C) \mid v \in V \text{ and } \exists r.C \in \rho(v)\} \cup \bigcup_{C \in \Psi} r^{\mathcal{J}_C}$$

$$a_v^{\mathcal{I}} = v$$

It is easy to see that $\mathcal{I}$ is a model of $\mathcal{A}$. In order to show that it is also a model of $\mathcal{T}$, we will require the following two properties:

1. $C \in \rho(v) \Rightarrow v \in C^{\mathcal{I}}$
2. $v \in C^{\mathcal{I}} \,\&\, C \in \Gamma \Rightarrow C \in \rho(v)$

If C is an atomic concept, then (1) follows directly from the definition of $A^{\mathcal{I}}$. If C is of the form $\exists r.D$, then we use the fact that v is connected via r to the individual x_D which belongs to $D^{\mathcal{J}_C}$, hence $D^{\mathcal{I}}$. If $C = C_1 \sqcap C_2$, then both $C_1 \in \rho(v)$ and $C_2 \in \rho(v)$ (by definition of the rule set Θ), so the statement follows by structural induction.

For statement (2), the proof is by induction on the co-depth of v. The base case is when v is a leaf node. The case where $C = A$ is trivial, and the case where C is a conjunction is easily dealt with by structural induction, so the only interesting case is when C is of the form $\exists r.D$. In this case, v must have some r-successor which is in $D^{\mathcal{I}}$, and since v has no successors in V, the r-successor must be x_E for some E such that $\exists r.E \in \rho(v)$. Now since $x_E \in D^{\mathcal{I}}$, it is easy to see that $x_E \in D^{\mathcal{J}_E}$, too. Using properties of canonical models (Lemma 3), we find that $D(b)$ is entailed by $\{B(b)\}$ and $\mathcal{T} \cup \{B \equiv E\}$, which means that $\mathcal{T} \models E \sqsubseteq D$. But in that case, we must have $\exists r.D \in \rho(v)$, as desired. Now let us consider the case where v is a non-leaf node with label $\langle t, r_1, ..., r_n \rangle$, and suppose that we have already shown statement (2) to hold for all of v's successors. Again, we restrict our attention to the interesting case where $C = \exists r.D$. If v's only r-successors satisfying D are outside V, then we can proceed as in the base case. Instead suppose that $r_i = r$, the i-th successor of v is w, and $w \in D^{\mathcal{I}}$. Then by the induction hypothesis, we must have $D \in \rho(w)$. It follows from the definition of the rule set Θ that $\exists r.D$ belongs to $\rho(v)$.

Now let us suppose that $C \sqcup D \in \mathcal{T}$ and $y \in C^{\mathcal{I}}$. The case where $y \in \Delta^{\mathcal{J}_E}$ for some $E \in \Psi$ is straightforward, so we concentrate on the case where $y \in V$. In this case, we know from statement (2) that $C \in \rho(y)$, which means that D must also belong to $\rho(y)$. It follows then from statement (1) that $y \in D^{\mathcal{I}}$, as desired. We have thus shown that $\mathcal{I}$ is a model of $\mathcal{A}$ and $\mathcal{T}$, so $\mathcal{A}$ is consistent with $\mathcal{T}$.

We now prove that some A_0 assertion is entailed by $\mathcal{A}$ and $\mathcal{T}$. We start by establishing the following claim.

Claim. For all $v \in V$ and $C \in \rho(v)$, we have $\mathcal{T}, \mathcal{A} \models C(a_v)$.

The proof is by induction on the co-depth of v. If v is a leaf and $C \in \rho(v)$, then the definition of Δ and $\mathcal{A}$ yields that $C \in \mathsf{cl}_{\mathcal{T}}(\sigma_{\mathcal{A}}(a_v))$. A straightforward semantic argument shows that this implies $\mathcal{T}, \mathcal{A} \models C(a_v)$. Now let v be a non-leaf with $\ell(v) = \langle t, r_1, \ldots, r_n \rangle$ and successors $v_1, \ldots, v_n$. Moreover, let $C \in \rho(v)$. Then

$$C \in \mathsf{cl}_{\mathcal{T}}(\sigma_{\mathcal{A}}(a_v) \cup \{\exists r.D \in \mathsf{sub}(\mathcal{T}) \mid r = r_i \text{ and } D \in r(a_{v_i}) \text{ for some } 1 \leq i \leq n\}).$$

By IH, we know that $D \in \rho(a_{v_i})$ implies $\mathcal{T}, \mathcal{A} \models D(a_{v_i})$. Thus, we can use the semantics to show that $\mathcal{T}, \mathcal{A} \models C(a_v)$. This finishes the proof of the claim.

By definition of Q_f and of runs, we have $A_0 \in r(\varepsilon)$ with ε the root of T. The claim thus yields that $\mathcal{T}, \mathcal{A} \models A_0(v_\varepsilon)$. □

Proof of Lemma 6.
Forward direction: Straightforward. Consider some $n \times m$ solution to the tiling problem. Create individuals $a_{i,j}$ for $0 \leq i \leq n-1$ and $0 \leq j \leq m-1$, and consider the ABox $\mathcal{A}$ composed of the following assertions:

- $x(a_{i,j}, a_{i+1,j})$ for $0 \leq i < n-1$ and $0 \leq j \leq m-1$
- $y(a_{i,j}, a_{i,j+1})$ for $0 \leq j < m-1$ and $0 \leq i \leq n-1$
- $T_k(a_{i,j})$ where T_k is the tile associated with the position (i, j)

It can be easily verified that $\mathcal{A}$ is consistent with $\mathcal{T}$ and satisfies $\mathcal{T}, \mathcal{A} \models A(a_{0,0})$.

Backward direction: Let $\mathcal{A}$ be a Σ-ABox consistent with $\mathcal{T}$ and such that $\mathcal{T}, \mathcal{A} \models A(a_A)$ for some $a_A \in \mathsf{Ind}(\mathcal{A})$. We exhibit a grid structure in $\mathcal{A}$ that gives rise to a tiling for $(\mathfrak{T}, H, V)$. We start by identifying a diagonal that starts at a_A and ends at an instance of T_{final}.

Claim 1. There is a set

$$\mathcal{G} := \{r_1(a_{i_0,j_0}, a_{i_1,j_1}), \ldots, r_{k-1}(a_{i_{k-1},j_{k-1}}, a_{i_k,j_k}), T_{\mathsf{final}}(a_{i_k,j_k})\} \subseteq \mathcal{A}$$

such that

- $i_0 = 0$, $j_0 = 0$, and $a_{0,0} = a_A$;
- for $1 \leq \ell < k$, we either have (i) $r_\ell = x$, $i_{\ell+1} = i_\ell + 1$, and $j_{\ell+1} = j_\ell$ or (ii) $r_\ell = y$, $j_{\ell+1} = j_\ell + 1$, and $i_{\ell+1} = i_\ell$.

Proof of claim. Assume there is no such sequence and let $\mathcal{I}$ be a model of $\mathcal{A}$ and $\mathcal{T}$. Since neither existential restrictions nor concept names T_i occur on the right-hand side of CIs in $\mathcal{T}$, we can w.l.o.g. assume that $\mathcal{I}$ is *minimal*, i.e., the following conditions are satisfied, for all $a \in \mathsf{Ind}(\mathcal{A})$, role names r, and $i \in \{1, \ldots, p\}$:

1. $\Delta^{\mathcal{I}} = \mathsf{Ind}(\mathcal{A})$;
2. $a^{\mathcal{I}} = a$;
3. $(a, a') \in r^{\mathcal{I}}$ implies $r(a, a') \in \mathcal{A}$;
4. $a \in T_h^{\mathcal{I}}$ implies $T_h(a) \in \mathcal{A}$.

We can convert $\mathcal{I}$ into a new model $\mathcal{J}$ of $\mathcal{A}$ and $\mathcal{T}$ that interprets Y as false at all points reachable from a_A in $\mathcal{I}$ (equivalently, in $\mathcal{A}$) and setting $A^{\mathcal{J}} = A^{\mathcal{I}} \setminus \{a_A\}$, which is a contradiction to $\mathcal{T}, \mathcal{A} \models A(a_A)$. (*end of proof of claim*).

Let n be the number of occurrences of the role x in the ABox $\mathcal{G}$ from Claim 1 and m the number of occurrences of y. We next show

Claim 2. We have that

(a) $\mathcal{T}, \mathcal{A} \models T_{\mathsf{init}}(a_{0,0})$;
(b) $\mathcal{T}, \mathcal{A} \models \neg R(a_{i,j})$ whenever $i < n$; otherwise, $\mathcal{T}, \mathcal{A} \models R(a_{i,j})$;
(c) $\mathcal{T}, \mathcal{A} \models \neg U(a_{i,j})$ whenever $j < m$; otherwise, $\mathcal{T}, \mathcal{A} \models R(a_{i,j})$;
(d) $\mathcal{T}, \mathcal{A} \models Y(a)$ for all $a \in \mathsf{Ind}(\mathcal{G})$;
(e) for all $a_{i,j} \in \mathsf{Ind}(\mathcal{G})$, there is a (unique) T_h with $\mathcal{T}, \mathcal{A} \models T_h(a_{i,j})$, henceforth denoted $T_{i,j}$;
(f) $(T_{i,j}, T_{i+1,j}) \in H$ for all $a_{i,j}, a_{i+1,j} \in \mathsf{Ind}(\mathcal{G})$ and $(T_{i,j}, T_{i,j+1}) \in V$ for all $a_{i,j}, a_{i,j+1} \in \mathsf{Ind}(\mathcal{G})$.

Proof of claim. Point (a) is an easy consequence of the fact that $a_{0,0} = a_A$ and $\mathcal{T}, \mathcal{A} \models A(a_A)$. For (b), first note that there is a unique $\ell \leq k$ such that $i_s = n$ for all $s \in \{\ell, \ldots, k\}$ and $i_p < n$ for all $s \in \{0, \ldots, \ell - 1\}$. We have $x(a_{i_{\ell-1},j_{\ell-1}}, a_{i_\ell,j_\ell}) \in \mathcal{G}$, thus the CI $R \sqsubseteq \forall x.\bot$ yields $\mathcal{T}, \mathcal{A} \models \neg R(a_{i_{\ell-1},j_{\ell-1}})$. To show that $\mathcal{T}, \mathcal{A} \models \neg R(a_{i_s,j_s})$ for all $s < \ell - 1$, it suffices to use the CIs $R \sqsubseteq \forall x.\bot$ and $\exists y.\neg R \sqsubseteq \neg R$.[1] To show that $\mathcal{T}, \mathcal{A} \models R(a_{i_s,j_s})$ for all $s \geq \ell - 1$, it suffices to note that (i) $T_{\mathsf{final}}(a_{i_k,j_k}) \in \mathcal{G}$ implies $\mathcal{T}, \mathcal{A} \models R(a_{i_k,j_k})$; (ii) that $y(a_{i_s,j_s}, a_{i_{s+1},j_{s+1}}) \in \mathcal{G}$ for $\ell \leq s < k$; and (iii) that we can apply the CI $\exists y.\neg R \sqsubseteq \neg R$. The proof of (c) is similar. We prove (d)-(f) together, showing by induction on ℓ that (d)-(f) are satisfied for all initial parts

$$\mathcal{G}_\ell := \{r_1(a_{i_0,j_0}, a_{i_1,j_1}), \ldots, r_{\ell-1}(a_{i_{\ell-1},j_{\ell-1}}, a_{i_\ell,j_\ell})\}$$

of $\mathcal{G}$, with $\ell \leq k$. For the base case, $a_{i_0,j_0} = a_A$ and $\mathcal{T}, \mathcal{A} \models A(a_0)$ clearly imply $\mathcal{T}, \mathcal{A} \models Y(a_{i_0,j_0})$, thus (d) is satisfied. Point (e) follows from (a) and the disjointness of tiles expressed in $\mathcal{T}$. Point (f) is vacuously true since there is only a single individual in $\mathcal{G}_0$. For the induction step, assume that $\mathcal{G}_{\ell-1}$ satisfies (d)-(f). By (b) and (c), we can distinguish four cases:

– $\mathcal{T}, \mathcal{A} \models \neg U(a_{i_{\ell-1},j_{\ell-1}})$ and $\mathcal{T}, \mathcal{A} \models \neg R(a_{i_{\ell-1},j_{\ell-1}})$.
  Since $\mathcal{G}_{\ell-1}$ satisfies (*d*), we have $\mathcal{T}, \mathcal{A} \models Y(a_{i_{\ell-1},j_{\ell-1}})$. Thus, by definition of $\mathcal{T}$ and since $\mathcal{T}, \mathcal{A} \models \neg U(a_{i_{\ell-1},j_{\ell-1}})$ and $\mathcal{T}, \mathcal{A} \models \neg R(a_{i_{\ell-1},j_{\ell-1}})$, we must have $\mathcal{T}, \mathcal{A} \models D(i_{\ell-1}, j_{\ell-1})$ with

$$D = \exists x.(T_j \sqcap Y \sqcap \exists y.Y) \sqcap \exists y.(T_\ell \sqcap Y \sqcap \exists x.Y) \sqcap C \sqcap T_i$$

  for some $(T_i, T_j) \in H$ and $(T_i, T_\ell) \in V$. Using the functionality of x and y, it is now easy to show that $\mathcal{G}_\ell$ satisfies (d)-(f).

[1] It is easy to work out a detailed, model-theoretic proof of this and similar claims below. We leave details to the reader.

– $\mathcal{T},\mathcal{A} \models \neg U(a_{i_{\ell-1},j_{\ell-1}})$ and $\mathcal{T},\mathcal{A} \models R(a_{i_{\ell-1},j_{\ell-1}})$.
Since $\mathcal{T},\mathcal{A} \models R(a_{i_{\ell-1},j_{\ell-1}})$, $\mathcal{T}$ ensures that there is no x-successor of $a_{i_{\ell-1},j_{\ell-1}}$ in $\mathcal{I}$. Moreover, $\mathcal{T},\mathcal{A} \models Y(a_{i_{\ell-1},j_{\ell-1}})$. Thus, by definition of $\mathcal{T}$ and since $\mathcal{T},\mathcal{A} \models \neg U(a_{i_{\ell-1},j_{\ell-1}})$ and $\mathcal{T},\mathcal{A} \models R(a_{i_{\ell-1},j_{\ell-1}})$, we must have $\mathcal{T},\mathcal{A} \models D(i_{\ell-1},j_{\ell-1})$ with
$$D = \exists y.(T_\ell \sqcap Y \sqcap R) \sqcap T_i$$
for some $(T_i,T_\ell) \in V$. We must have $i_\ell = i_{\ell-1}$, $j_\ell = j_{\ell-1}+1$, and $r_{\ell-1} = y$. Using the functionality of y, it is now easy to show that $\mathcal{G}_\ell$ satisfies (d)-(f).
– $\mathcal{T},\mathcal{A} \models U(a_{i_{\ell-1},j_{\ell-1}})$ and $\mathcal{T},\mathcal{A} \models \neg R(a_{i_{\ell-1},j_{\ell-1}})$.
Analogous to the previous case.
– $\mathcal{T},\mathcal{A} \models U(a_{i_{\ell-1},j_{\ell-1}})$ and $\mathcal{T},\mathcal{A} \models R(a_{i_{\ell-1},j_{\ell-1}})$.
Then the definition of $\mathcal{T}$ ensures that we have $x(a_{i_{\ell-1},j_{\ell-1}},a_{i_\ell,j_\ell}) \notin \mathcal{A}$ and $y(a_{i_{\ell-1},j_{\ell-1}},a_{i_\ell,j_\ell}) \notin \mathcal{A}$. It follows that $\ell-1=k$, in contradiction to $\ell \leq k$.

(*end of proof of claim*).

Next, we extend $\mathcal{G}$ to a full grid such that Conditions (a)-(e) from Claim 2 are still satisfied. Once this is achieved, it is trivial to read off a solution for the tiling problem. The construction of the grid consists of exhaustive application of the following two steps:

1. if $x(a_{i,j},a_{i+1,j}), y(a_{i+1,j},a_{i+1,j+1}) \in \mathcal{G}$ with $i<n$ and $a_{i,j+1} \notin \mathsf{Ind}(\mathcal{G})$, then identify an $a_{i,j+1} \in \mathsf{Ind}(\mathcal{A})$ such that $y(a_{i,j},a_{i,j+1}), x(a_{i,j+1},a_{i+1,j+1}) \in \mathcal{A}$ and add the latter two assertions to $\mathcal{G}$.
2. if $y(a_{i,j},a_{i,j+1}), x(a_{i,j+1},a_{i+1,j+1}) \in \mathcal{G}$ with $i<n$ and $a_{i+1,j} \notin \mathsf{Ind}(\mathcal{G})$, then identify an $a_{i+1,j} \in \mathsf{Ind}(\mathcal{A})$ such that $y(a_{i,j},a_{i+1,j}), x(a_{i+1,j},a_{i+1,j+1}) \in \mathcal{A}$ and add the latter two assertions to $\mathcal{G}$.

It is not hard to see that exhaustive application of these rules yields a full grid, i.e., for the final $\mathcal{G}$ we have (i) $\mathsf{Ind}(\mathcal{G}) = \{a_{i,j} \mid i \leq n, j \leq m\}$, (ii) $x(a_{i,j},a_{i',j'}) \in \mathcal{G}$ iff $i'=i+1$ and $j=j'$, and (iii) $y(a_{i,j},a_{i',j'}) \in \mathcal{G}$ iff $i=i'$ and $j'=j+1$. A solution to the tiling problem can be read off from this grid due to Conditions (e) and (f).

Since the two steps of the construction are completely analogous, we only deal with Case 1 in detail. Thus let $x(a_{i,j},a_{i+1,j}), y(a_{i+1,j},a_{i+1,j+1}) \in \mathcal{G}$ with $a_{i,j+1} \notin \mathsf{Ind}(\mathcal{G})$. Clearly, $i<n$ and $j<m$. By (b) and (c), we thus have $\mathcal{T},\mathcal{A} \models \neg R(a_{i,j})$ and $\mathcal{T},\mathcal{A} \models \neg U(a_{i,j})$. Since $\mathcal{T},\mathcal{A} \models Y(a_{i_{\ell-1},j_{\ell-1}})$ by (d), the definition of $\mathcal{T}$ yields $\mathcal{T},\mathcal{A} \models D(a_{i_{\ell-1},j_{\ell-1}})$ with
$$D = \exists x.(T_j \sqcap Y \sqcap \exists y.Y) \sqcap \exists y.(T_\ell \sqcap Y \sqcap \exists x.Y) \sqcap C \sqcap T_i$$
for some $(T_i,T_j) \in H$ and $(T_i,T_\ell) \in V$. Since there are no existential restrictions on the right-hand side of CIs in $\mathcal{T}$ and by the functionality of x and y, there are $a_{i,j+1}, b \in \mathsf{Ind}(\mathcal{A})$ such that $y(a_{i,j},a_{i,j+1}), x(a_{i,j+1},b) \in \mathcal{A}$, $a_{i,j+1}, b \in Y^{\mathcal{I}}$, and $T_{i,j+1} = T_\ell$. With this choice, (a) and (d)-(f) are clearly satisfied. To establish the properties stated in Step 1 above, we have to show that $b = a_{i+1,j+1}$. From this, the satisfaction of (b) and (c) before we apply the construction step, and the CIs
$$R \sqsubseteq \forall x.\bot \quad \exists y.\neg R \sqsubseteq \neg R \quad \exists x.\neg U \sqsubseteq \neg U \quad U \sqsubseteq \forall y.\bot$$

it then follows that (b) and (c) are still satisfied after the step.

Suppose to the contrary of what remains to be shown that $b \neq a_{i+1,j+1}$. Since $\mathcal{T}, \mathcal{A} \models \neg R(a_{i,j})$, $\mathcal{T}, \mathcal{A} \models \neg U(a_{i,j})$, and $\mathcal{T}, \mathcal{A} \models Y(a_{i_{\ell-1},j_{\ell-1}})$, the definition of $\mathcal{T}$ yields $\mathcal{T}, \mathcal{A} \models C(a_{i,j})$. Take a model $\mathcal{I}$ of $\mathcal{T}$ and $\mathcal{A}$, and set $Z^{\mathcal{I}} = (Z^{\mathcal{I}} \cup \{b^{\mathcal{I}}\}) \setminus \{a^{\mathcal{I}}_{i+1,j+1}\}$. Now interpret C, Y, U, and R minimally such that all axioms in $\mathcal{T}$ are still satified, i.e., set

$$
\begin{aligned}
C^{\mathcal{I}} &= (\exists x.\exists y.Z \sqcap \exists y.\exists x.Z)^{\mathcal{I}} \cup (\exists x.\exists y.\neg Z \sqcap \exists y.\exists x.\neg Z)^{\mathcal{I}} \\
Y^{\mathcal{I}} &= (\exists x.(T_j \sqcap Y \sqcap \exists y.Y) \sqcap \exists y.(T_\ell \sqcap Y \sqcap \exists x.Y) \sqcap C \sqcap T_i)^{\mathcal{I}} \cup \\
&\quad T^{\mathcal{I}}_{\mathsf{final}} \cup (\exists x.(T_j \sqcap Y \sqcap U) \sqcap T_i)^{\mathcal{I}} \cup (\exists y.(T_\ell \sqcap Y \sqcap R) \sqcap T_i)^{\mathcal{I}} \\
U^{\mathcal{I}} &= (T_{\mathsf{final}})^{\mathcal{I}} \cup (\exists x.(T_j \sqcap Y \sqcap U) \sqcap T_i)^{\mathcal{I}} \\
R^{\mathcal{I}} &= (T_{\mathsf{final}})^{\mathcal{I}} \cup (\exists y.(T_\ell \sqcap Y \sqcap R) \sqcap T_i)^{\mathcal{I}}
\end{aligned}
$$

It is not hard to verify that $\mathcal{I}$ is still a model of $\mathcal{T}$ and $\mathcal{A}$. By definition of $Z^{\mathcal{I}}$ and $C^{\mathcal{I}}$ and by functionality of x and y, we have $a_{i,j} \notin C^{\mathcal{I}}$. It follows that $\mathcal{T}, \mathcal{A} \not\models C(a_{i,j})$, which is a contradiction. $\square$

Proof of Lemma 7.

Forward direction: Straightforward, see previous proof.

Backward direction: Let $\mathcal{A}$ be a Σ-ABox consistent with $\mathcal{T}$ and such that $\mathcal{T}, \mathcal{A} \models \exists v.A(v)$. Call a model $\mathcal{I}$ of $\mathcal{A}$ and $\mathcal{T}$ *minimal* if the following conditions are satisfied, for all $a \in \mathsf{Ind}(\mathcal{A})$, role names r, and $i \in \{1, \ldots, p\}$:

1. $\Delta^{\mathcal{I}} = \mathsf{Ind}(\mathcal{A})$;
2. $a^{\mathcal{I}} = a$;
3. $(a, a') \in r^{\mathcal{I}}$ implies $r(a, a') \in \mathcal{A}$;
4. $a \in T^{\mathcal{I}}_h$ implies $T_h(a) \in \mathcal{A}$.

Since neither existential restrictions nor concept names T_i occur on the right-hand side of CIs in $\mathcal{T}$, it is not hard to verify that there is a minimal model $\mathcal{I}$ of $\mathcal{A}$ and $\mathcal{T}$. We additionally assume w.l.o.g., that

- $\mathcal{I}$ is Y, C-*minimal*: if $\mathcal{J}$ is obtained from $\mathcal{I}$ by deleting elements of $Y^{\mathcal{I}}$ and $C^{\mathcal{I}}$ while keeping the extension of all other symbols unchanged, then $\mathcal{J}$ is not a model of $\mathcal{A}$ and $\mathcal{T}$.
- $\mathcal{I}$ is A-*minimal*: there is no minimal model $\mathcal{J}$ of $\mathcal{T}$ and $\mathcal{A}$ such that $A^{\mathcal{J}} \subsetneq A^{\mathcal{I}}$.

Let $a_A \in A^{\mathcal{I}}$. We now exhibit a grid structure in $\mathcal{A}$ that gives rise to a tiling for $(\mathfrak{T}, H, V)$. We start by identifying a diagonal that starts at a_A and ends at an instance of T_{final}.

Claim 1. There is a set

$$\mathcal{G} := \{r_1(a_{i_0,j_0}, a_{i_1,j_1}), \ldots, r_{k-1}(a_{i_{k-1},j_{k-1}}, a_{i_k,j_k}), T_{\mathsf{final}}(a_{i_k,j_k})\} \subseteq \mathcal{A}$$

such that

– $i_0 = 0$, $j_0 = 0$, and $a_{0,0} = a_A$;
– for $1 \leq \ell < k$, we either have (i) $r_\ell = x$, $i_{\ell+1} = i_\ell + 1$, and $j_{\ell+1} = j_\ell$ or (ii) $r_\ell = y$, $j_{\ell+1} = j_\ell + 1$, and $i_{\ell+1} = i_\ell$.

Proof of claim. If there is no such sequence, we can convert $\mathcal{I}$ into a new model $\mathcal{J}$ of $\mathcal{A}$ and $\mathcal{T}$ by interpreting Y as false at all points reachable from a_A in $\mathcal{I}$ (equivalently: $\mathcal{A}$) and setting $A^{\mathcal{J}} = A^{\mathcal{I}} \setminus \{a_A\}$, which is a contradiction to A-minimality of $\mathcal{I}$. (*end of proof of claim*).

Let n be the number of occurrences of the role x in the ABox $\mathcal{G}$ from Claim 1 and m the number of occurrences of y. We next show

Claim 2. We have that

(a) $a_{0,0} \in T_{\mathsf{init}}^{\mathcal{I}}$.
(b) $a_{i,j} \in R^{\mathcal{I}}$ implies $i = n$;
(c) $a_{i,j} \in U^{\mathcal{I}}$ implies $j = m$;
(d) $a \in Y^{\mathcal{I}}$ for all $a \in \mathsf{Ind}(\mathcal{G})$;
(e) for all $a_{i,j} \in \mathsf{Ind}(\mathcal{G})$, there is a (unique) T_h with $a_{i,j} \in T_h^{\mathcal{I}}$, henceforth denoted $T_{i,j}$;
(f) $(T_{i,j}, T_{i+1,j}) \in H$ for all $a_{i,j}, a_{i+1,j} \in \mathsf{Ind}(\mathcal{G})$ and $(T_{i,j}, T_{i,j+1}) \in V$ for all $a_{i,j}, a_{i,j+1} \in \mathsf{Ind}(\mathcal{G})$.

Proof of claim. Point (a) is an easy consequence of the fact that $a_{0,0} = a_A$, $a_A \in A^{\mathcal{I}}$, and $\mathcal{I}$ is A-minimal. For (b), first note that there is a unique $\ell \leq k$ such that $i_s = n$ for all $s \in \{\ell, \ldots, k\}$ and $i_p < n$ for all $s \in \{0, \ldots, \ell - 1\}$. We have $y(a_{i_{\ell-1},j_{\ell-1}}, a_{i_\ell,j_\ell}) \in \mathcal{G}$, thus the CI $R \sqsubseteq \forall x.\bot$ yields $a_{i_{\ell-1},j_{\ell-1}} \notin R^{\mathcal{I}}$. To show that $a_{i_s,j_s} \notin R^{\mathcal{I}}$ for all $s < \ell - 1$, it suffices to use the CIs $\exists x.\neg R \sqsubseteq \neg R$ and $\exists y.\neg R \sqsubseteq \neg R$. The proof of (c) is similar. We prove (d)-(f) together, showing by induction on ℓ that (d)-(f) are satisfied for all initial parts

$$\mathcal{G}_\ell := \{r_1(a_{i_0,j_0}, a_{i_1,j_1}), \ldots, r_{\ell-1}(a_{i_{\ell-1},j_{\ell-1}}, a_{i_\ell,j_\ell})\}$$

of $\mathcal{G}$, with $\ell \leq k$. For the base case, $a_{i_0,j_0} = a_A \in A^{\mathcal{I}}$ and A-minimality of $\mathcal{I}$ clearly imply $a_{i_0,j_0} \in Y^{\mathcal{I}}$, thus (d) is satisfied. Point (e) follows from (a) and the disjointness of tiles expressed in $\mathcal{T}$. Point (f) is vacuously true since there is only a single individual in $\mathcal{G}_0$. For the induction step, assume that $\mathcal{G}_{\ell-1}$ satisfies (d)-(f). We distinguish four cases:

– $a_{i_{\ell-1},j_{\ell-1}} \in (\neg U \sqcap \neg R)^{\mathcal{I}}$.
Since $\mathcal{G}_{\ell-1}$ satisfies (d), we have $a_{i_{\ell-1},j_{\ell-1}} \in Y^{\mathcal{I}}$ and, by definition of $\mathcal{T}$, Y, C-minimality of $\mathcal{I}$ together with $a_{i_{\ell-1},j_{\ell-1}} \in (\neg U \sqcap \neg R)^{\mathcal{I}}$ ensure that

$$a_{i_{\ell-1},j_{\ell-1}} \in (\exists x.(T_j \sqcap Y \sqcap \exists y.Y) \sqcap \exists y.(T_\ell \sqcap Y \sqcap \exists x.Y) \sqcap C \sqcap T_i)^{\mathcal{I}}$$

for some $(T_i, T_j) \in H$ and $(T_i, T_\ell) \in V$. Using the functionality of x and y, it is now easy to show that $\mathcal{G}_\ell$ satisfies (d)-(f).

- $a_{i_{\ell-1},j_{\ell-1}} \in (\neg U \sqcap R)^{\mathcal{I}}$.
Since $a_{i_{\ell-1},j_{\ell-1}} \in R^{\mathcal{I}}$, $\mathcal{T}$ ensures that there is no x-successor of $a_{i_{\ell-1},j_{\ell-1}}$ in $\mathcal{I}$. Moreover, $a_{i_{\ell-1},j_{\ell-1}} \in Y^{\mathcal{I}}$. From Y, C-minimality of $\mathcal{I}$ and the definition of $\mathcal{T}$, we get
$$a_{i_{\ell-1},j_{\ell-1}} \in (\exists y.(T_\ell \sqcap Y \sqcap R) \sqcap T_i)^{\mathcal{I}}$$
for some $(T_i, T_\ell) \in V$. We must have $i_\ell = i_{\ell-1}$, $j_\ell = j_{\ell-1} + 1$, and $r_{\ell-1} = y$. Using the functionality of y, it is now easy to show that $\mathcal{G}_\ell$ satisfies (d)-(f).
- $a_{i_{\ell-1},j_{\ell-1}} \in (U \sqcap \neg R)^{\mathcal{I}}$.
Analogous to the previous case.
- $a_{i_{\ell-1},j_{\ell-1}} \in (U \sqcap R)^{\mathcal{I}}$.
Then there is neither an x-successor nor a y-successor of $a_{i_{\ell-1},j_{\ell-1}} \in (U \sqcap R)^{\mathcal{I}}$. It follows that $\ell - 1 = k$, in contradiction to $\ell \leq k$.

(*end of proof of claim*).

Next, we extend $\mathcal{G}$ to a full grid such that Conditions (a)-(e) from Claim 2 are still satisfied. Once this is achieved, it is trivial to read off a solution for the tiling problem. The construction of the grid consists of exhaustive application of the following two steps:

1. if $x(a_{i,j}, a_{i+1,j}), y(a_{i+1,j}, a_{i+1,j+1}) \in \mathcal{G}$ with $i < n$ and $a_{i,j+1} \notin \mathsf{Ind}(\mathcal{G})$, then identify an $a_{i,j+1} \in \mathsf{Ind}(\mathcal{A})$ such that $y(a_{i,j}, a_{i,j+1}), x(a_{i,j+1}, a_{i+1,j+1}) \in \mathcal{A}$ and add the latter two assertions to $\mathcal{G}$.
2. if $y(a_{i,j}, a_{i,j+1}), x(a_{i,j+1}, a_{i+1,j+1}) \in \mathcal{G}$ with $i < n$ and $a_{i+1,j} \notin \mathsf{Ind}(\mathcal{G})$, then identify an $a_{i+1,j} \in \mathsf{Ind}(\mathcal{A})$ such that $y(a_{i,j}, a_{i+1,j}), x(a_{i+1,j}, a_{i+1,j+1}) \in \mathcal{A}$ and add the latter two assertions to $\mathcal{G}$.

It is not hard to see that exhaustive application of these rules yields a full grid, i.e., for the final $\mathcal{G}$ we have (i) $\mathsf{Ind}(\mathcal{G}) = \{a_{i,j} \mid i \leq n, j \leq m\}$, (ii) $x(a_{i,j}, a_{i',j'}) \in \mathcal{G}$ iff $i' = i + 1$ and $j = j'$, and (iii) $y(a_{i,j}, a_{i',j'}) \in \mathcal{G}$ iff $i = i'$ and $j' = j + 1$. A solution to the tiling problem can be read off from this grid due to Conditions (e) and (f).

Since the two steps of the construction are completely analogous, we only deal with Case 1 in detail. Thus let $x(a_{i,j}, a_{i+1,j}), y(a_{i+1,j}, a_{i+1,j+1}) \in \mathcal{G}$ with $a_{i,j+1} \notin \mathsf{Ind}(\mathcal{G})$. Clearly, $i < n$ and $j < m$. By (b) and (c), we thus have $a_{i,j} \notin (R \sqcup U)^{\mathcal{I}}$. Since $a_{i,j} \in Y^{\mathcal{I}}$ by (d) and $\mathcal{I}$ is Y, C-minimal, we get that

$$a_{i,j} \in (\exists x.(T_j \sqcap Y \sqcap \exists y.Y) \sqcap \exists y.(T_\ell \sqcap Y \sqcap \exists x.Y) \sqcap C \sqcap T_i)^{\mathcal{I}}$$

for some $(T_i, T_j) \in H$ and $(T_i, T_\ell) \in V$. Thus and since $\mathcal{I}$ is minimal, we can choose $a_{i,j+1}, b \in \mathsf{Ind}(\mathcal{A})$ such that $y(a_{i,j}, a_{i,j+1}), x(a_{i,j+1}, b) \in \mathcal{A}$, $a_{i,j+1}, b \in Y^{\mathcal{I}}$, and $T_{i,j+1} = T_\ell$. With this choice, (a) and (d)-(f) are clearly satisfied. To establish the properties stated in Step 1 above, we have to show that $b = a_{i+1,j+1}$. From this, the satisfaction of (b) and (c) before the construction step, and the CIs

$$R \sqsubseteq \forall x.\bot \quad \exists y.\neg R \sqsubseteq \neg R \quad \exists x.\neg U \sqsubseteq \neg U \quad U \sqsubseteq \forall y.\bot$$

it follows that (b) and (c) are still satisfied after the step.

Suppose to the contrary of what remains to be shown that $b \neq a_{i+1,j+1}$. Since $a_{i,j} \in Y^{\mathcal{I}}$, we have $a_{i,j} \in C^{\mathcal{I}}$. Due to Y, C-minimality and by definition of $\mathcal{T}$ and functionality of x and y, we have $b, a_{i+1,j+1} \in \mathcal{B}^{\mathcal{I}}$ for some Boolean combination $\mathcal{B}$ of Z_1, Z_2. Let $\mathcal{B}'$, $\mathcal{B}''$, $\mathcal{B}'''$ be such combinations that are distinct from each other and from $\mathcal{B}$. Since $\mathcal{A}$ is finite and x and y are inverse functional, we find a unique maximal and finite sequence $b_0, \ldots, b_r \in \mathsf{Ind}(\mathcal{A})$ such that

(i) $b_0 = a_{i+1,j+1}$ for some $i \leq r$;
(ii) for all $i < r$, there is a $c \in \mathsf{Ind}(\mathcal{A})$ such that $(c, b_i) \in (y \circ x)^{\mathcal{I}}$ and $(c, b_{i+1}) \in (x \circ y)^{\mathcal{I}}$;
(iii) all b_i are distinct individuals.

We can uniquely extend the sequence to a new sequence $b_{-r'}, b_{-r'+1}, \ldots, b_0, \ldots, b_r$ such that the Conditions (i) to (iii) are still satisfied, with i in (ii) now ranging from $-r'$ to $r - 1$.

Define a new interpretation $\mathcal{J}$ starting with $\mathcal{I}$ by reinterpreting the concept names Z_1, Z_2 such that

- $b_i \in \mathcal{B}'$ if i is even and $-r' \leq i < r$;
- $b_i \in \mathcal{B}$ if i is odd and $-r' \leq i < r$;
- $b_r \in \mathcal{B}''$.

Moreover, we remove $a_{i,j}$ from $C^{\mathcal{I}}$ and shrink $Y^{\mathcal{I}}$ such that $a_{i,j} \notin Y^{\mathcal{J}}$, $x(a, b) \in \mathcal{A}$ with $b \notin Y^{\mathcal{J}}$ implies $a \notin Y^{\mathcal{J}}$, and $y(a, b) \in \mathcal{A}$ with $b \notin Y^{\mathcal{J}}$ implies $a \notin Y^{\mathcal{J}}$. In particular, this will result in $a_{0,0} = a_A \notin Y^{\mathcal{J}}$. Define $A^{\mathcal{J}} = A^{\mathcal{I}} \setminus \{a_A\}$. It can be verified that $\mathcal{J}$ is a model of $\mathcal{A}$ and $\mathcal{T}$, contradicting A-minimality of $\mathcal{I}$. $\square$