Query Inseparability by Games

E. Botoeva[1], R. Kontchakov[2], V. Ryzhikov[1], F. Wolter[3], and M. Zakharyaschev[2]

[1] KRDB Research Centre, Free University of Bozen-Bolzano, Italy
{botoeva,ryzhikov}@inf.unibz.it
[2] Department of Computer Science and Inf. Systems, Birkbeck, University of London, U.K.
{roman,michael}@dcs.bbk.ac.uk
[3] Department of Computer Science, University of Liverpool, U.K.
wolter@liverpool.ac.uk

Abstract. We investigate conjunctive query inseparability of description logic knowledge bases (KBs) with respect to a given signature, a fundamental problem for KB versioning, module extraction, forgetting and knowledge exchange. We develop a game-theoretic technique for checking query inseparability of KBs expressed in fragments of Horn-$\mathcal{ALCHI}$, and show a number of complexity results ranging from P to ExpTime and 2ExpTime. We also employ our results to resolve two major open problems for *OWL 2 QL* by showing that TBox query inseparability and the membership problem for universal UCQ-solutions in knowledge exchange are both ExpTime-complete for combined complexity.

Introduction

A description logic (DL) knowledge base (KB) consists of a terminological box (TBox), storing conceptual knowledge, and an assertion box (ABox), storing data. Typical applications of KBs involve answering queries over incomplete data sources (ABoxes) augmented by ontologies (TBoxes) that provide additional information about the domain of interest as well as a convenient vocabulary for user queries. The standard query language in such applications, which balances expressiveness and computational complexity, is the language of conjunctive queries (CQs).

With typically large data, often tangled ontologies, and the hard problem of answering CQs over ontologies, various transformation and comparison tasks are becoming indispensable for KB engineering and maintenance. For example, to make answering certain CQs more efficient, one may want to extract from a given KB a smaller module returning the same answers to those CQs as the original KB; to provide the user with a more convenient query vocabulary, one may want to reformulate the KB in a new language. These tasks are known as module extraction [19] and knowledge exchange [2]; other relevant tasks include versioning, revision and forgetting [10, 20, 15].

In this paper, we investigate the following relationship between KBs that is fundamental for all such tasks. Let Σ be a signature consisting of concept and role names. We call KBs $\mathcal{K}_1$ and $\mathcal{K}_2$ *Σ-query inseparable* and write $\mathcal{K}_1 \equiv_\Sigma \mathcal{K}_2$ if any CQ formulated in Σ has the same answers over $\mathcal{K}_1$ and $\mathcal{K}_2$.The relativisation to (smaller) signatures is crucial to support the tasks mentioned above:

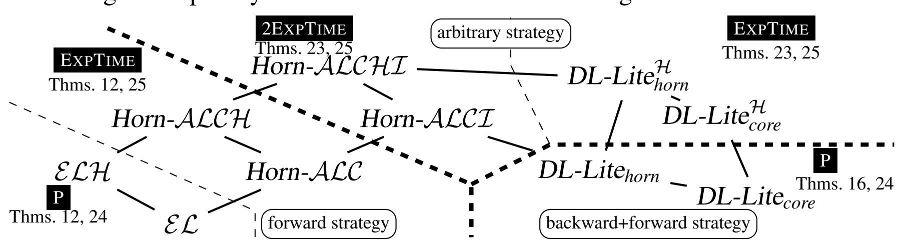
(versioning) When comparing two *versions* $\mathcal{K}_1$ and $\mathcal{K}_2$ of a KB with respect to their answers to CQs in a relevant signature Σ, the task is to check whether $\mathcal{K}_1 \equiv_\Sigma \mathcal{K}_2$.

(modularisation) A *Σ-module of* a KB $\mathcal{K}$ is a KB $\mathcal{K}' \subseteq \mathcal{K}$ such that $\mathcal{K}' \equiv_\Sigma \mathcal{K}$. If we are only interested in answering CQs in Σ over $\mathcal{K}$, then we can achieve our aim by querying any Σ-module of $\mathcal{K}$ instead of $\mathcal{K}$ itself.

(knowledge exchange) In knowledge exchange, we want to transform a KB $\mathcal{K}_1$ in a signature Σ_1 to a new KB $\mathcal{K}_2$ in a disjoint signature Σ_2 connected to Σ_1 via a declarative mapping specification given by a TBox $\mathcal{T}_{12}$. Thus, the target KB $\mathcal{K}_2$ should satisfy the condition $\mathcal{K}_1 \cup \mathcal{T}_{12} \equiv_{\Sigma_2} \mathcal{K}_2$, in which case it is called a *universal UCQ-solution* (CQ and UCQ inseparabilities coincide for Horn DLs).

(forgetting) A KB $\mathcal{K}'$ is the result of *forgetting* a signature Σ in a KB $\mathcal{K}$ if $\mathcal{K}'$ does not use Σ and gives the same answers to CQs without symbols in Σ as $\mathcal{K}$: that is, $\mathsf{sig}(\mathcal{K}') \subseteq \mathsf{sig}(\mathcal{K}) \setminus \Sigma$ and $\mathcal{K}' \equiv_{\mathsf{sig}(\mathcal{K}) \setminus \Sigma} \mathcal{K}$.

We study the data and combined complexity of deciding Σ-query inseparability for KBs expressed in various fragments of the DL *Horn-$\mathcal{ALCHI}$* [14], which include *DL-Lite$^{\mathcal{H}}_{core}$* [7] and $\mathcal{EL}$ [3] underlying the W3C profiles *OWL 2 QL* and *OWL 2 EL*. To establish upper complexity bounds, we develop a novel game-theoretic technique for checking finite Σ-homomorphic embeddability between (possibly infinite) materialisations of KBs. For all of the considered DLs, Σ-query inseparability turns out to be P-complete for data complexity, which matches the data complexity of CQ evaluation for all of our DLs lying outside the *DL-Lite* family. For combined complexity, the obtained tight complexity results are summarised in the diagram below.



Most interesting are ExpTime-completeness of *DL-Lite$^{\mathcal{H}}_{core}$* and 2ExpTime-completeness of *Horn-$\mathcal{ALCI}$*, which contrast with NP-completeness and ExpTime-completeness of CQ evaluation for those logics. For *DL-Lite* without role inclusions and $\mathcal{ELH}$, Σ-query inseparability is P-complete, while CQ evaluation is NP-complete. In general, it is the combined presence of inverse roles and qualified existential restrictions (or role inclusions) that makes Σ-query inseparability hard.

We apply our results to resolve two important open problems. First, we show that the membership problem for universal UCQ-solutions in knowledge exchange for KBs in *DL-Lite$^{\mathcal{H}}_{core}$* is ExpTime-complete for combined complexity, which settles an open question of Arenas *et al.* [1], where only PSpace-hardness was established. We also show that Σ-query inseparability of *DL-Lite$^{\mathcal{H}}_{core}$* TBoxes is ExpTime-complete, which closes the PSpace–ExpTime gap that was left open by Konev *et al.* [12].

Recall that TBoxes $\mathcal{T}_1$ and $\mathcal{T}_2$ are Σ-query inseparable if, for *all* Σ-ABoxes $\mathcal{A}$ (which only use concept and role names from Σ), the KBs $(\mathcal{T}_1, \mathcal{A})$ and $(\mathcal{T}_2, \mathcal{A})$ are Σ-query inseparable. TBox and KB inseparabilities have different applications. The former supports ontology engineering when data is not known or changes frequently: one can equivalently replace one TBox with another only if they return the same answers to queries for *every* Σ-ABox. In contrast, KB inseparability is useful in applications where data is stable—such as knowledge exchange or variants of module extraction and forgetting with fixed data—in order to use the KB in a new application or as a compilation step to make CQ answering more efficient. As we show below, TBox and KB Σ-query inseparabilities also have different computational properties.

All the omitted proofs can be found in the full version of the paper [6].

Preliminaries

All the DLs for which we investigate KB Σ-query inseparability are Horn fragments of $\mathcal{ALCHI}$. To define these DLs, we fix sequences of *individual names* a_i, *concept names* A_i, and *role names* P_i, where $i < \omega$. A *role* is either a role name P_i or an *inverse role* P_i^-; we assume that $(P_i^-)^- = P_i$. Concepts in the DLs $\mathcal{ALCI}$, $\mathcal{ALC}$, and $\mathcal{EL}$ are defined as usual. *DL-Lite*$_{horn}$*-concepts* are constructed from concept names using the constructors $\top$, $\bot$, $\sqcap$, and $\exists R.\top$ and *DL-Lite*$_{core}$-concepts are *DL-Lite*$_{horn}$-concepts without $\sqcap$; in other words, they are *basic concepts* of the form $\bot$, $\top$, A_i or $\exists R.\top$.

For a DL $\mathcal{L}$, an $\mathcal{L}$*-concept inclusion* (CI) takes the form $C \sqsubseteq D$, where C and D are $\mathcal{L}$-concepts. An $\mathcal{L}$*-TBox*, $\mathcal{T}$, is a finite set of $\mathcal{L}$-CIs. An $\mathcal{ALCHI}$, *DL-Lite*$^{\mathcal{H}}_{horn}$ and *DL-Lite*$^{\mathcal{H}}_{core}$ TBox can also contain a finite number of *role inclusions* (RIs) $R_1 \sqsubseteq R_2$, where the R_i are roles. In $\mathcal{ELH}$ TBoxes, RIs have no inverse roles. *DL-Lite* TBoxes may also contain *disjointness constraints* $B_1 \sqcap B_2 \sqsubseteq \bot$ and $R_1 \sqcap R_2 \sqsubseteq \bot$, for basic concepts B_i and roles R_i. To introduce the Horn fragments of these DLs, we require the standard recursive definition [9, 11] of positive occurrences of a concept. A TBox $\mathcal{T}$ is *Horn* if no concept of the form $C \sqcup D$ occurs positively in $\mathcal{T}$, and no concept of the form $\neg C$ or $\forall R.C$ occurs negatively in $\mathcal{T}$. In the DL *Horn-*$\mathcal{L}$ only *Horn-*$\mathcal{L}$-TBoxes are allowed. An *ABox*, $\mathcal{A}$, is a finite set of *assertions* of the form $A_k(a_i)$ or $P_k(a_i, a_j)$. An $\mathcal{L}$-TBox $\mathcal{T}$ and an ABox $\mathcal{A}$ together form an $\mathcal{L}$ *knowledge base* (KB) $\mathcal{K} = (\mathcal{T}, \mathcal{A})$. The set of individual names in $\mathcal{K}$ is denoted by $\mathsf{ind}(\mathcal{K})$.

The semantics for the DLs is defined in the usual way based on interpretations $\mathcal{I} = (\Delta^{\mathcal{I}}, \cdot^{\mathcal{I}})$ that comply with the *unique name assumption*: $a_i^{\mathcal{I}} \neq a_j^{\mathcal{I}}$ for $i \neq j$ [4]. We write $\mathcal{I} \models \alpha$ in case an inclusion or assertion α is true in $\mathcal{I}$. If $\mathcal{I} \models \alpha$, for all $\alpha \in \mathcal{T} \cup \mathcal{A}$, then $\mathcal{I}$ is a *model* of a KB $\mathcal{K} = (\mathcal{T}, \mathcal{A})$; in symbols: $\mathcal{I} \models \mathcal{K}$. $\mathcal{K}$ is *consistent* if it has a model. $\mathcal{K} \models \alpha$ means that $\mathcal{I} \models \alpha$ for all $\mathcal{I} \models \mathcal{K}$.

A *conjunctive query* (CQ) $\boldsymbol{q}(\boldsymbol{x})$ is a formula $\exists \boldsymbol{y}\, \varphi(\boldsymbol{x}, \boldsymbol{y})$, where φ is a conjunction of atoms of the form $A_k(z_1)$ or $P_k(z_1, z_2)$ with $z_i \in \boldsymbol{x} \cup \boldsymbol{y}$. A tuple $\boldsymbol{a}$ in $\mathsf{ind}(\mathcal{K})$ (of the same length as $\boldsymbol{x}$) is a *certain answer* to $\boldsymbol{q}(\boldsymbol{x})$ over $\mathcal{K}$ if $\mathcal{I} \models \boldsymbol{q}(\boldsymbol{a})$ for all $\mathcal{I} \models \mathcal{K}$; in this case we write $\mathcal{K} \models \boldsymbol{q}(\boldsymbol{a})$.

A *signature*, Σ, is a set of concept and role names. By a Σ-concept, Σ-role, Σ-CQ, etc. we understand any concept, role, CQ, etc. constructed using the names from Σ.

Definition 1. Let $\mathcal{K}_1$ and $\mathcal{K}_2$ be KBs and Σ a signature. Then $\mathcal{K}_1$ *Σ-query entails* $\mathcal{K}_2$ if $\mathcal{K}_2 \models \boldsymbol{q}(\boldsymbol{a})$ implies $\mathcal{K}_1 \models \boldsymbol{q}(\boldsymbol{a})$ for *all* Σ-CQs $\boldsymbol{q}(\boldsymbol{x})$ and all tuples $\boldsymbol{a}$ in $\mathsf{ind}(\mathcal{K}_2)$. And $\mathcal{K}_1$ and $\mathcal{K}_2$ are *Σ-query inseparable* if they Σ-query entail each other; in which case we write $\mathcal{K}_1 \equiv_\Sigma \mathcal{K}_2$.

Since Σ-query inseparability can be reduced to two Σ-query entailment checks, we can prove complexity upper bounds for entailment. Conversely, for most languages we have a semantically transparent reduction of Σ-query entailment to Σ-query inseparability:

Theorem 2. *For any of our DLs $\mathcal{L}$ containing $\mathcal{EL}$ or having role inclusions, Σ-query entailment for $\mathcal{L}$-KBs is* LOGSPACE*-reducible to Σ-query inseparability for $\mathcal{L}$-KBs.*

We now consider the relationship between inseparability and universal UCQ-solutions in knowledge exchange. Suppose $\mathcal{K}_1$ and $\mathcal{K}_2$ are KBs in disjoint signatures Σ_1 and Σ_2. Let $\mathcal{T}_{12}$ be a mapping consisting of inclusions of the form $S_1 \sqsubseteq S_2$, where the S_i are concept (or role) names in Σ_i. Then $\mathcal{K}_2$ is a universal UCQ-solution for $(\mathcal{K}_1, \mathcal{T}_{12}, \Sigma_2)$ if $\mathcal{K}_1 \cup \mathcal{T}_{12} \equiv_{\Sigma_2} \mathcal{K}_2$. Deciding the latter is called the *membership problem for universal UCQ-solutions*. For DLs $\mathcal{L}$ with role inclusions, the problem whether $\mathcal{K}_1 \cup \mathcal{T}_{12} \equiv_{\Sigma_2} \mathcal{K}_2$ is a Σ_2-query inseparability problem in $\mathcal{L}$. Conversely, we have:

Theorem 3. *Σ-query entailment for any of our DLs $\mathcal{L}$ is* LOGSPACE*-reducible to the membership problem for universal UCQ-solutions in $\mathcal{L}$.*

Semantic Characterisation

In this section, we give a semantic characterisation of Σ-query entailment based on an abstract notion of materialisation and finite homomorphisms between such structures.

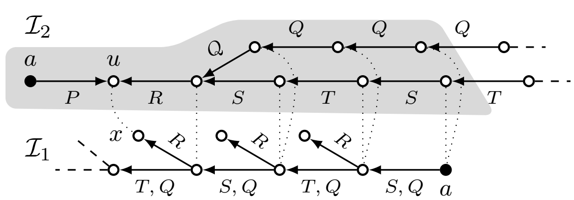
Let $\mathcal{K}$ be a KB. An interpretation $\mathcal{I}$ is called a *materialisation* of $\mathcal{K}$ if $\mathcal{K} \models \boldsymbol{q}(\boldsymbol{a})$ just in case $\mathcal{I} \models \boldsymbol{q}(\boldsymbol{a})$, for all CQs $\boldsymbol{q}(\boldsymbol{x})$ and tuples $\boldsymbol{a}$ in $\mathsf{ind}(\mathcal{K})$. We say that $\mathcal{K}$ is *materialisable* if it has a materialisation. Materialisations can be used to characterise KB Σ-query entailment by means of Σ-homomorphisms. For an interpretation $\mathcal{I}$ and a signature Σ, the *Σ-types* $\boldsymbol{t}_\Sigma^\mathcal{I}(x)$ and $\boldsymbol{r}_\Sigma^\mathcal{I}(x,y)$ of $x, y \in \Delta^\mathcal{I}$ are defined by taking

$$\boldsymbol{t}_\Sigma^\mathcal{I}(x) = \{\, \Sigma\text{-concept name } A \mid x \in A^\mathcal{I} \,\}, \quad \boldsymbol{r}_\Sigma^\mathcal{I}(x,y) = \{\, \Sigma\text{-role } R \mid (x,y) \in R^\mathcal{I} \,\}.$$

Suppose $\mathcal{I}_i$ is a materialisation of $\mathcal{K}_i$, $i = 1, 2$. A function $h\colon \Delta^{\mathcal{I}_2} \to \Delta^{\mathcal{I}_1}$ is a *Σ-homomorphism* from $\mathcal{I}_2$ to $\mathcal{I}_1$ if, for any $a \in \mathsf{ind}(\mathcal{K}_2)$ and any $x, y \in \Delta^{\mathcal{I}_2}$,

– $h(a^{\mathcal{I}_2}) = a^{\mathcal{I}_1}$ whenever $\boldsymbol{t}_\Sigma^{\mathcal{I}_2}(a) \neq \emptyset$ or $\boldsymbol{r}_\Sigma^{\mathcal{I}_2}(a,y) \neq \emptyset$ for some $y \in \Delta^{\mathcal{I}_2}$, and
– $\boldsymbol{t}_\Sigma^{\mathcal{I}_2}(x) \subseteq \boldsymbol{t}_\Sigma^{\mathcal{I}_1}(h(x))$, $\boldsymbol{r}_\Sigma^{\mathcal{I}_2}(x,y) \subseteq \boldsymbol{r}_\Sigma^{\mathcal{I}_1}(h(x),h(y))$.

As answers to Σ-CQs are preserved under Σ-homomorphisms, $\mathcal{K}_1$ Σ-query entails $\mathcal{K}_2$ if there is a Σ-homomorphism from $\mathcal{I}_2$ to $\mathcal{I}_1$. However, the converse does not hold.



Example 4. Suppose $\mathcal{I}_2$ and $\mathcal{I}_1$ on the right are materialisations of KBs $\mathcal{K}_2$ and $\mathcal{K}_1$, where a is the only ABox individual. Let $\Sigma = \{Q, R, S, T\}$. Then there is no Σ-homomorphism from $\mathcal{I}_2$ to $\mathcal{I}_1$ (as $\boldsymbol{r}_\Sigma^{\mathcal{I}_2}(a,u) = \emptyset$, we can map u to, say, x but then only the shaded part of $\mathcal{I}_2$ can be mapped Σ-homomorphically to $\mathcal{I}_1$). However, for any Σ-query $\boldsymbol{q}(\boldsymbol{x})$, $\mathcal{I}_2 \models \boldsymbol{q}(\boldsymbol{c})$ implies $\mathcal{I}_1 \models \boldsymbol{q}(\boldsymbol{c})$ as any finite subinterpretation of $\mathcal{I}_2$ can be Σ-homomorphically mapped to $\mathcal{I}_1$.

We say that $\mathcal{I}_2$ is *finitely Σ-homomorphically embeddable into* $\mathcal{I}_1$ if, for every *finite* subinterpretation $\mathcal{I}_2'$ of $\mathcal{I}_2$, there exists a Σ-homomorphism from $\mathcal{I}_2'$ to $\mathcal{I}_1$.

To prove the following theorem, one can regard any finite subinterpretation of $\mathcal{I}_2$ as a CQ whose variables are elements of $\Delta^{\mathcal{I}_2}$, with the answer variables being in $\mathsf{ind}(\mathcal{K}_2)$.

Theorem 5. *Suppose $\mathcal{K}_i$ is a consistent KB with a materialisation $\mathcal{I}_i$, $i = 1, 2$. Then $\mathcal{K}_1$ Σ-query entails $\mathcal{K}_2$ iff $\mathcal{I}_2$ is finitely Σ-homomorphically embeddable into $\mathcal{I}_1$.*

One problem with applying Theorem 5 is that materialisations are in general infinite for any of the DLs considered in this paper. We address this problem by introducing finite representations of materialisations. Let $\mathcal{K}$ be a KB and let $\mathcal{G} = (\Delta^{\mathcal{G}}, \cdot^{\mathcal{G}}, \rightsquigarrow)$ be a finite structure such that $\Delta^{\mathcal{G}} = \mathsf{ind}(\mathcal{K}) \cup \Omega$, for $\mathsf{ind}(\mathcal{K}) \cap \Omega = \emptyset$, $\cdot^{\mathcal{G}}$ is an interpretation function on $\Delta^{\mathcal{G}}$ with $A_i^{\mathcal{G}} \subseteq \Delta^{\mathcal{G}}$, $P_i^{\mathcal{G}} \subseteq \mathsf{ind}(\mathcal{K}) \times \mathsf{ind}(\mathcal{K})$, and $(\Delta^{\mathcal{G}}, \rightsquigarrow)$ is a directed graph (containing loops) with nodes $\Delta^{\mathcal{G}}$ and edges $\rightsquigarrow \subseteq \Delta^{\mathcal{G}} \times \Omega$, in which every edge $u \rightsquigarrow v$ is labelled with a set $(u, v)^{\mathcal{G}} \neq \emptyset$ of roles satisfying the condition: if $u_1 \rightsquigarrow v$ and $u_2 \rightsquigarrow v$, then $(u_1, v)^{\mathcal{G}} = (u_2, v)^{\mathcal{G}}$. We call $\mathcal{G}$ a *generating structure for* $\mathcal{K}$ if the interpretation $\mathcal{M}$ defined below is a materialisation of $\mathcal{K}$. A *path* in $\mathcal{G}$ is a sequence $\sigma = u_0 \dots u_n$ with $u_0 \in \mathsf{ind}(\mathcal{K})$ and $u_i \rightsquigarrow u_{i+1}$ for $i < n$. Let $\mathsf{tail}(\sigma) = u_n$ and let $\mathsf{path}(\mathcal{G})$ be the set of paths in $\mathcal{G}$. The materialisation $\mathcal{M}$ is given by: $\Delta^{\mathcal{M}} = \mathsf{path}(\mathcal{G})$,

$$a^{\mathcal{M}} = a, \text{ for } a \in \mathsf{ind}(\mathcal{K}), \quad A^{\mathcal{M}} = \{\sigma \mid \mathsf{tail}(\sigma) \in A^{\mathcal{G}}\},$$
$$P^{\mathcal{M}} = P^{\mathcal{G}} \cup \{(\sigma, \sigma u) \mid \mathsf{tail}(\sigma) \rightsquigarrow u, P \in (\mathsf{tail}(\sigma), u)^{\mathcal{G}}\}$$
$$\cup \{(\sigma u, \sigma) \mid \mathsf{tail}(\sigma) \rightsquigarrow u, P^- \in (\mathsf{tail}(\sigma), u)^{\mathcal{G}}\}.$$

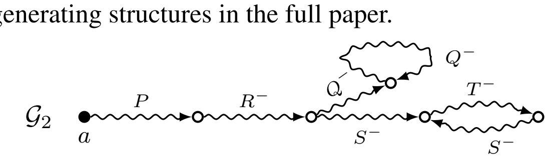
DL $\mathcal{L}$ has *finitely generated materialisations* if every $\mathcal{L}$-KB has a generating structure.

Theorem 6. *Horn-$\mathcal{ALCHI}$ and all of its fragments defined above have finitely generated materialisations. Moreover,*

- *for any $\mathcal{L} \in \{\mathcal{ALCHI}, \mathcal{ALCI}, \mathcal{ALCH}, \mathcal{ALC}\}$ and any Horn-$\mathcal{L}$ KB $(\mathcal{T}, \mathcal{A})$, a generating structure can be constructed in time $|\mathcal{A}| \cdot 2^{p(|\mathcal{T}|)}$, p a polynomial*;
- *for any $\mathcal{L}$ in the $\mathcal{EL}$ and DL-Lite families introduced above and any $\mathcal{L}$-KB $(\mathcal{T}, \mathcal{A})$, a generating structure can be constructed in time $|\mathcal{A}| \cdot p(|\mathcal{T}|)$, p a polynomial.*

Finite generating structures have been defined for $\mathcal{EL}$ [16], *DL-Lite* [13] and more expressive Horn DLs [8]. With the exception of *DL-Lite*, however, the relation $\rightsquigarrow$ guiding the construction of materialisations was implicit. We show how the existing constructions can be converted to generating structures in the full paper.

Example 7. The materialisation $\mathcal{I}_2$ from Example 4 can be generated by the structure $\mathcal{G}_2$ shown on the right.



For a generating structure $\mathcal{G}$ for $\mathcal{K}$ and a signature Σ, the *Σ-types* $\boldsymbol{t}_{\Sigma}^{\mathcal{G}}(u)$ and $\boldsymbol{r}_{\Sigma}^{\mathcal{G}}(u, v)$ of $u, v \in \Delta^{\mathcal{G}}$ are defined by taking $\boldsymbol{t}_{\Sigma}^{\mathcal{G}}(u) = \{\, \Sigma\text{-concept name } A \mid u \in A^{\mathcal{G}} \,\}$, $\boldsymbol{r}_{\Sigma}^{\mathcal{G}}(u, v)$ as $\{\, \Sigma\text{-role } R \mid (u, v) \in R^{\mathcal{G}} \,\}$ if $u, v \in \mathsf{ind}(\mathcal{K})$, as $\{\, \Sigma\text{-role } R \mid R \in (u, v)^{\mathcal{G}} \,\}$ if $u \rightsquigarrow v$, and $\emptyset$ otherwise, where $(P^-)^{\mathcal{G}}$ is the converse of $P^{\mathcal{G}}$. We also define $\bar{\boldsymbol{r}}_{\Sigma}^{\mathcal{G}}(u, v)$ to contain the inverses of the roles in $\boldsymbol{r}_{\Sigma}^{\mathcal{G}}(u, v)$; note that $\bar{\boldsymbol{r}}_{\Sigma}^{\mathcal{G}}(u, v)$ is not the same as $\boldsymbol{r}_{\Sigma}^{\mathcal{G}}(v, u)$. We write $u \rightsquigarrow^{\Sigma} v$ if $u \rightsquigarrow v$ and $\boldsymbol{r}_{\Sigma}^{\mathcal{G}}(u, v) \neq \emptyset$.

In the next section, we show that, for a DL $\mathcal{L}$ having finitely generated materialisations, Σ-query entailment for $\mathcal{L}$-KBs can be reduced to the problem of finding a winning strategy in a game played on the generating structures for these KBs.

Σ-Query Entailment by Games

Suppose a DL $\mathcal{L}$ has finitely generated materialisations, $\mathcal{K}_i$ is a consistent $\mathcal{L}$-KB, for $i = 1, 2$, and Σ a signature. Let $\mathcal{G}_i = (\Delta^{\mathcal{G}_i}, \cdot^{\mathcal{G}_i}, \rightsquigarrow_i)$ be a generating structure for $\mathcal{K}_i$ and $\mathcal{M}_i$ be its materialisation; $\mathcal{G}_i^\Sigma$ and $\mathcal{M}_i^\Sigma$ denote the restrictions of $\mathcal{G}_i$ and $\mathcal{M}_i$ to Σ.

We begin with a very simple game on the finite generating structure $\mathcal{G}_2^\Sigma$ and the possibly infinite materialisation $\mathcal{M}_1^\Sigma$.

Infinite game $G_\Sigma(\mathcal{G}_2, \mathcal{M}_1)$. This game is played by two players: player 2 and player 1. The *states* of the game are of the form $\mathfrak{s}_i = (u_i \mapsto \sigma_i)$, for $i \geq 0$, where $u_i \in \Delta^{\mathcal{G}_2}$ and $\sigma_i \in \Delta^{\mathcal{M}_1}$ satisfy the following condition:

(s$_1$) $\boldsymbol{t}_\Sigma^{\mathcal{G}_2}(u_i) \subseteq \boldsymbol{t}_\Sigma^{\mathcal{M}_1}(\sigma_i)$.

The game starts in a state $\mathfrak{s}_0 = (u_0 \mapsto \sigma_0)$ with $\sigma_0 = u_0$ in case $u_0 \in \mathsf{ind}(\mathcal{K}_2)$. In each round $i > 0$, player 2 challenges player 1 with some $u_i \in \Delta^{\mathcal{G}_2}$ such that $u_{i-1} \rightsquigarrow_2^\Sigma u_i$. Player 1 has to respond with a $\sigma_i \in \Delta^{\mathcal{M}_1}$ satisfying **(s$_1$)** and

(s$_2$) $\boldsymbol{r}_\Sigma^{\mathcal{G}_2}(u_{i-1}, u_i) \subseteq \boldsymbol{r}_\Sigma^{\mathcal{M}_1}(\sigma_{i-1}, \sigma_i)$.

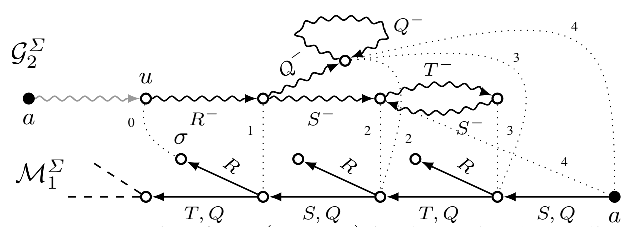
This gives the next state $\mathfrak{s}_i = (u_i \mapsto \sigma_i)$. Note that of all the u_i only u_0 may be an ABox individual; however, there is no such a restriction on the σ_i. A *play of length* $n \geq 0$ *starting from* $\mathfrak{s}_0$ is any sequence $\mathfrak{s}_0, \ldots, \mathfrak{s}_n$ of states obtained as described above. For an ordinal $\lambda \leq \omega$, we say that player 1 has a λ*-winning strategy in the game* $G_\Sigma(\mathcal{G}_2, \mathcal{M}_1)$ *starting from a state* $\mathfrak{s}_0$ if, for any play of length $i < \lambda$, which starts from $\mathfrak{s}_0$ and conforms with this strategy, and any challenge of player 2 in round $i+1$, player 1 has a response. The following theorem gives a game-theoretic flavour to the criterion of Theorem 5 (see the full paper for a proof).

Theorem 8. *$\mathcal{M}_2$ is finitely Σ-homomorphically embeddable into $\mathcal{M}_1$ iff*

(abox) $\boldsymbol{r}_\Sigma^{\mathcal{M}_2}(a, b) \subseteq \boldsymbol{r}_\Sigma^{\mathcal{M}_1}(a, b)$*, for any* $a, b \in \mathsf{ind}(\mathcal{K}_2)$*, and,*

(win) *for any* $u_0 \in \Delta^{\mathcal{G}_2}$ *and* $n < \omega$*, there exists* $\sigma_0 \in \Delta^{\mathcal{M}_1}$ *such that player 1 has an* n*-winning strategy in the game* $G_\Sigma(\mathcal{G}_2, \mathcal{M}_1)$ *starting from* $(u_0 \mapsto \sigma_0)$.

Example 9. Let $\Sigma = \{Q, R, S, T\}$. Consider $\mathcal{G}_2^\Sigma$ and $\mathcal{M}_1^\Sigma$ shown in the picture on the right. For any $n < \omega$ and $u \in \Delta^{\mathcal{G}_2}$, player 1 has an n-winning strategy in $G_\Sigma(\mathcal{G}_2, \mathcal{M}_1)$. A 4-winning strategy starting from $(u \mapsto \sigma)$ is shown by dotted lines (in round 2, player 2 has two possible challenges). For a larger n, a suitable σ can be chosen further away from the root a of $\mathcal{M}_1$.



The criterion of Theorem 8 does not seem to be a big improvement on Theorem 5 as we still have to deal with an infinite materialisation. Our aim now is to show that condition **(win)** in the infinite game $G_\Sigma(\mathcal{G}_2, \mathcal{M}_1)$ can be checked by analysing a more complex game on the *finite* generating structures $\mathcal{G}_2$ and $\mathcal{G}_1$. We consider four types of strategies in $G_\Sigma(\mathcal{G}_2, \mathcal{M}_1)$. For each strategy type, τ, we define a game $G^\tau_\Sigma(\mathcal{G}_2, \mathcal{G}_1)$ such that, for any $u_0 \in \Delta^{\mathcal{G}_2}$, the following conditions are equivalent:

($<\omega$) for every $n < \omega$, player 1 has an n-winning τ-strategy in $G_\Sigma(\mathcal{G}_2, \mathcal{M}_1)$ starting from some $(u_0 \mapsto \sigma_0^n)$;

(ω) player 1 has an ω-winning strategy in $G^\tau_\Sigma(\mathcal{G}_2, \mathcal{G}_1)$ starting from some state depending on u_0 and τ.

We begin with 'forward' winning strategies sufficient for DLs without inverse roles.

Forward strategy and game $G^f_\Sigma(\mathcal{G}_2, \mathcal{G}_1)$. We say that a λ-strategy ($\lambda \leq \omega$) for player 1 in the game $G_\Sigma(\mathcal{G}_2, \mathcal{M}_1)$ is *forward* if, for any play of length $i-1 < \lambda$, which conforms with this strategy, and any challenge $u_{i-1} \rightsquigarrow^\Sigma_2 u_i$ by player 2, the response σ_i of player 1 is such that either $\sigma_{i-1}, \sigma_i \in \mathsf{ind}(\mathcal{K}_1)$ or $\sigma_i = \sigma_{i-1}v$, for some $v \in \Delta^{\mathcal{G}_1}$.

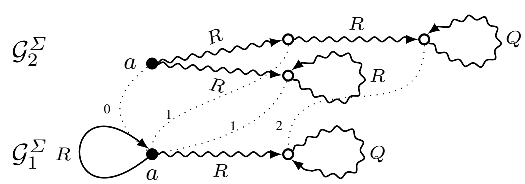
For example, if the $\mathcal{G}_i$, $i = 1, 2$, satisfy the condition

(f) the Σ-labels on $\rightsquigarrow_i$-edges contain no inverse roles,

then *every* strategy in $G_\Sigma(\mathcal{G}_2, \mathcal{M}_1)$ is forward. This is clearly the case for *Horn-$\mathcal{ALCH}$*, *Horn-$\mathcal{ALC}$*, $\mathcal{ELH}$ and $\mathcal{EL}$, which by definition do not have inverse roles.

The existence of a forward λ-winning strategy for player 1 in $G_\Sigma(\mathcal{G}_2, \mathcal{M}_1)$ is equivalent to the existence of an ω-winning strategy in the game $G^f_\Sigma(\mathcal{G}_2, \mathcal{G}_1)$, which is defined similarly to $G_\Sigma(\mathcal{G}_2, \mathcal{M}_1)$ but with two modifications: (1) it is played on $\mathcal{G}_2$ and $\mathcal{G}_1$; and (2) the response $x_i \in \Delta^{\mathcal{G}_1}$ of player 1 to a challenge $u_{i-1} \rightsquigarrow^\Sigma_2 u_i$ must be such that either $x_{i-1}, x_i \in \mathsf{ind}(\mathcal{K}_1)$ or $x_{i-1} \rightsquigarrow_1 x_i$, and **(s$_1$)**–**(s$_2$)** hold (with $\mathcal{G}_1$ and x_i in place of $\mathcal{M}_1$ and σ_i).

Example 10. Let $\mathcal{G}_2$ and $\mathcal{G}_1$ be as shown on the right. Then, for any $u \in \Delta^{\mathcal{G}_2}$, there is $x \in \Delta^{\mathcal{G}_1}$ such that player 1 has an ω-winning strategy in $G^f_\Sigma(\mathcal{G}_2, \mathcal{G}_1)$ starting from $(u \mapsto x)$.



The next theorem follows from König's Lemma:

Lemma 11. *For any $u_0 \in \Delta^{\mathcal{G}_2}$, condition* **($<\omega$)** *holds for forward strategies in $G_\Sigma(\mathcal{G}_2, \mathcal{M}_1)$ iff* **(ω)** *holds in $G^f_\Sigma(\mathcal{G}_2, \mathcal{G}_1)$ for some state $(u_0 \mapsto x_0)$.*

$G^f_\Sigma(\mathcal{G}_2, \mathcal{G}_1)$ is a standard simulation or reachability game on finite graphs, where the existence of ω-winning strategies for player 1 follows from the existence of n-winning strategies for $n = O(|\mathcal{G}_2| \times |\mathcal{G}_1|)$, which can be checked in polynomial time [18, 5]. By Theorem 6 and **(f)**, we obtain:

Theorem 12. *For combined complexity, checking Σ-query entailment is in* P *for $\mathcal{EL}$ and $\mathcal{ELH}$ KBs, and in* ExpTime *for Horn-$\mathcal{ALC}$ and Horn-$\mathcal{ALCH}$ KBs. For data complexity, it is in* P *for all these DLs.*

In comparison to forward strategies, the winning strategies used in Example 9 can be described as 'backward.'

Backward strategy and game $G^b_\Sigma(\mathcal{G}_2, \mathcal{G}_1)$. A λ-strategy for player 1 in $G_\Sigma(\mathcal{G}_2, \mathcal{M}_1)$ is *backward* if, for any play of length $i-1 < \lambda$, which conforms with this strategy, and any challenge $u_{i-1} \rightsquigarrow^\Sigma_2 u_i$ by player 2, the response σ_i of player 1 is the *immediate predecessor* of σ_{i-1} in $\mathcal{M}_1$ in the sense that $\sigma_{i-1} = \sigma_i v$, for some $v \in \Delta^{\mathcal{G}_1}$ (player 1 loses in case $\sigma_{i-1} \in \mathsf{ind}(\mathcal{K}_1)$). Note that, since $\mathcal{M}_1$ is tree-shaped, the response of player 1 to any different challenge $u_{i-1} \rightsquigarrow^\Sigma_2 u'_i$ must be the same σ_i; cf. Example 9.

That is why the states of the game $G^b_\Sigma(\mathcal{G}_2, \mathcal{G}_1)$ are of the form $(\Xi_i \mapsto x_i)$, where $\Xi_i \subseteq \Delta^{\mathcal{G}_2}$, $\Xi_i \neq \emptyset$, and $x_i \in \Delta^{\mathcal{G}_1}$ satisfy the following condition:

($\mathbf{s}'_1$) $\boldsymbol{t}^{\mathcal{G}_2}_\Sigma(u) \subseteq \boldsymbol{t}^{\mathcal{G}_1}_\Sigma(x_i)$, for all $u \in \Xi_i$.

The game starts in a state $(\Xi_0 \mapsto x_0)$ such that

($\mathbf{s}'_0$) if $u \in \Xi_0 \cap \mathsf{ind}(\mathcal{K}_2)$, then $x_0 = u \in \mathsf{ind}(\mathcal{K}_1)$.

For each $i > 0$, player 2 always challenges player 1 with the set $\Xi_i = \Xi^{\rightsquigarrow}_{i-1}$, where

$$\Xi^{\rightsquigarrow} = \{v \in \Delta^{\mathcal{G}_2} \mid u \rightsquigarrow^\Sigma_2 v, \text{ for some } u \in \Xi\},$$

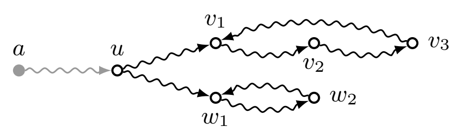
provided that it is not empty (otherwise, player 2 loses). Player 1 responds with $x_i \in \Delta^{\mathcal{G}_1}$ such that $x_i \rightsquigarrow_1 x_{i-1}$ and **($\mathbf{s}'_1$)** and the following condition hold:

($\mathbf{s}'_2$) $\boldsymbol{r}^{\mathcal{G}_2}_\Sigma(u, v) \subseteq \bar{\boldsymbol{r}}^{\mathcal{G}_1}_\Sigma(x_{i-1}, x_i)$, for all $u \in \Xi_{i-1}$, $v \in \Xi_i$.

Lemma 13. *For any $u_0 \in \Delta^{\mathcal{G}_2}$, condition* **($<\boldsymbol{\omega}$)** *holds for backward strategies in $G_\Sigma(\mathcal{G}_2, \mathcal{M}_1)$ iff* **($\boldsymbol{\omega}$)** *holds in $G^b_\Sigma(\mathcal{G}_2, \mathcal{G}_1)$ for some state $(\{u_0\} \mapsto x_0)$.*

Although Lemmas 11 and 13 look similar, the game $G^b_\Sigma(\mathcal{G}_2, \mathcal{G}_1)$ turns out to be more complex than $G^f_\Sigma(\mathcal{G}_2, \mathcal{G}_1)$.

$\mathcal{G}^\Sigma_2$



Example 14. To illustrate, consider $\mathcal{G}^\Sigma_2$ shown on the right (with concepts and roles omitted) and an arbitrary $\mathcal{G}_1$. A play in $G^b_\Sigma(\mathcal{G}_2, \mathcal{G}_1)$ may proceed as: $(\{u\} \mapsto x_0)$, $(\{v_1, w_1\} \mapsto x_1)$, $(\{v_2, w_2\} \mapsto x_2)$, $(\{v_3, w_1\} \mapsto x_3)$, etc. This gives at least 6 different sets Ξ_i. But if $\mathcal{G}_2$ contained k cycles of lengths $p_1, \dots, p_k$, where p_i is the ith prime number, then the number of states in $G^b_\Sigma(\mathcal{G}_2, \mathcal{G}_1)$ could be exponential ($p_1 \times \dots \times p_k$). In fact, we have the following:

Lemma 15. *Checking* **($\boldsymbol{\omega}$)** *in Lemma 13 is* coNP*-hard.*

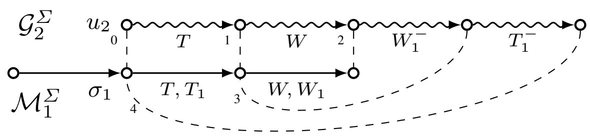
Observe that in the case of *DL-Lite$_{core}$* and *DL-Lite$_{horn}$*, (which have inverse roles but no RIs), generating structures $\mathcal{G} = (\Delta^\mathcal{G}, \cdot^\mathcal{G}, \rightsquigarrow)$ can be defined so that, for any $u \in \Delta^\mathcal{G}$ and R, there is *at most one* v with $u \rightsquigarrow v$ and $R \in \boldsymbol{r}^\mathcal{G}(u, v)$ [13]. As a result, any n-winning strategy starting from $(u_0 \mapsto \sigma_0)$ in $G_\Sigma(\mathcal{G}_2, \mathcal{M}_1)$ consists of a (possibly empty) backward part followed by a (possibly empty) forward part. Moreover, in the backward games for these DLs, the sets Ξ_i are always *singletons*. Thus, the number of states in the combined backward/forward games on the $\mathcal{G}_i$ is polynomial, and the existence of winning strategies can be checked in polynomial time.

Theorem 16. *Checking Σ-query entailment for DL-Lite$_{core}$ and DL-Lite$_{horn}$ KBs is in* P *for both combined and data complexity.*

An arbitrary strategy for player 1 in $G_\Sigma(\mathcal{G}_2, \mathcal{M}_1)$ is a combination of a backward strategy and a number of start-bounded strategies to be defined next.

Start-bounded strategy and game $G^s_\Sigma(\mathcal{G}_2, \mathcal{G}_1)$. A strategy for player 1 in the game $G_\Sigma(\mathcal{G}_2, \mathcal{M}_1)$ starting from a state $(u_0 \mapsto \sigma_0)$ is *start-bounded* if it never leads to $(u_i \mapsto \sigma_i)$ such that $\sigma_0 = \sigma_i v$, for some v and $i > 0$. In other words, player 1 cannot use those elements of $\mathcal{M}_1$ that are located closer to the ABox than σ_0; the ABox individuals in $\mathcal{M}_1$ can only be used if $\sigma_0 \in \mathsf{ind}(\mathcal{K}_1)$.

Example 17. The strategy starting from $(u_2 \mapsto \sigma_1)$ and shown on the right is start-bounded.



In the game $G^s_\Sigma(\mathcal{G}_2, \mathcal{G}_1)$, player 1 will have to guess *all* the points of $\mathcal{G}_2$ that are mapped to the same point of $\mathcal{M}_1$.

The *states* of $G^s_\Sigma(\mathcal{G}_2, \mathcal{G}_1)$ are of the form $(\varGamma_i, \varXi_i \mapsto x_i)$, $i \geq 0$, where $\varGamma_i, \varXi_i \subseteq \Delta^{\mathcal{G}_2}$, $\varXi_i \neq \emptyset$, $x_i \in \Delta^{\mathcal{G}_1}$ and **(s$'_1$)** holds. The initial state is of the form $(\emptyset, \varXi_0 \mapsto x_0)$ such that **(s$'_0$)** holds. In each round $i > 0$, player 2 challenges player 1 with some $u \rightsquigarrow_2^\Sigma v$ such that $u \in \varXi_{i-1}$ and

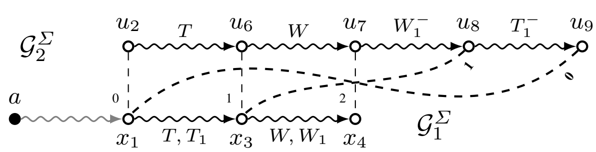
(nbk) if $v \in \varGamma_{i-1}$ then $\boldsymbol{r}^{\mathcal{G}_2}_\Sigma(u, v) \not\subseteq \bar{\boldsymbol{r}}^{\mathcal{G}_1}_\Sigma(x_{i-2}, x_{i-1})$.

Player 1 responds with either a state $(\varXi_{i-1}, \varXi_i \mapsto x_i)$ such that $x_{i-1} \rightsquigarrow_1 x_i$ (and so $x_i \notin \mathsf{ind}(\mathcal{K}_1)$) and **(s$''_2$)** holds, or a state $(\emptyset, \varXi_i \mapsto x_i)$ such that $x_{i-1}, x_i \in \mathsf{ind}(\mathcal{K}_1)$ and

(s$''_2$) $\boldsymbol{r}^{\mathcal{G}_2}_\Sigma(u, v) \subseteq \boldsymbol{r}^{\mathcal{G}_1}_\Sigma(x_{i-1}, x_i)$.

We make challenges $u \rightsquigarrow_2^\Sigma v$, for which $u \in \varXi_{i-1}$ and **(nbk)** does not hold, 'illegitimate' because x_{i-2} can always be used as a response. Because of this, player 1 always moves 'forward' in $\mathcal{G}_1$, but has to guess appropriate sets $\varXi_i$ in advance. Note that $\varGamma_i$ is always uniquely determined by x_{i-1}, x_i and $\varXi_{i-1}$ (and it is either $\varXi_{i-1}$ or empty).

Example 18. Let $\mathcal{G}_2^\Sigma$ and $\mathcal{G}_1^\Sigma$ be as shown on the right (cf. Example 17). We show that player 1 has an ω-winning strategy in

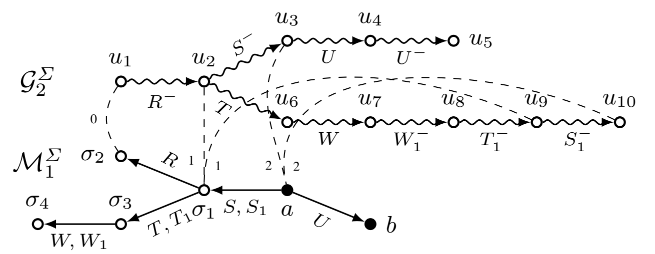


$G^s_\Sigma(\mathcal{G}_2, \mathcal{G}_1)$ starting from $(\emptyset, \{u_2, u_9\} \mapsto x_1)$. Player 2 challenges with $u_2 \rightsquigarrow_2^\Sigma u_6$, and player 1 responds with $(\{u_2, u_9\}, \{u_6, u_8\} \mapsto x_3)$. Then player 2 picks $u_6 \rightsquigarrow_2^\Sigma u_7$ and player 1 responds with $(\{u_6, u_8\}, \{u_7\} \mapsto x_4)$, where the game ends. Note the crucial guesses $\{u_2, u_9\} \mapsto x_1$ and $\{u_6, u_8\} \mapsto x_3$ made by player 1. If player 1 failed to guess that u_8 must also be mapped to x_3 and responded with $(\{u_2, u_9\}, \{u_6\} \mapsto x_3)$, then after the challenge $u_6 \rightsquigarrow_2^\Sigma u_7$ and response $(\{u_6\}, \{u_7\} \mapsto x_4)$), player 2 would pick $u_7 \rightsquigarrow_2^\Sigma u_8$, to which player 1 could not respond.

Lemma 19. *For any $u_0 \in \Delta^{\mathcal{G}_2}$, condition* **($<\omega$)** *holds for start-bounded strategies in $G_\Sigma(\mathcal{G}_2, \mathcal{M}_1)$ iff* **(ω)** *holds in $G^s_\Sigma(\mathcal{G}_2, \mathcal{G}_1)$ for some state $(\emptyset, \varXi_0 \mapsto x_0)$ with $u_0 \in \varXi_0$.*

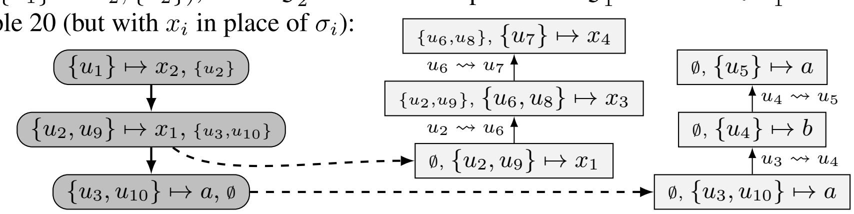
Arbitrary strategies and game $G^a_\Sigma(\mathcal{G}_2,\mathcal{G}_1)$. An arbitrary winning strategy in the game $G_\Sigma(\mathcal{G}_2,\mathcal{M}_1)$ can be composed of one backward and a number of start-bounded strategies.

Example 20. Consider $\mathcal{G}_2^\Sigma$ and $\mathcal{M}_1^\Sigma$ shown on the right. Starting from $(u_1 \mapsto \sigma_2)$, player 1 can respond to the challenges $u_1 \rightsquigarrow_2^\Sigma u_2 \rightsquigarrow_2^\Sigma u_3$ according to the backward strategy; the challenges $u_2 \rightsquigarrow_2^\Sigma u_6 \rightsquigarrow_2^\Sigma u_7 \rightsquigarrow_2^\Sigma u_8 \rightsquigarrow_2^\Sigma u_9$ according to the start-bounded strategy as in Example 17; the challenges $u_3 \rightsquigarrow_2^\Sigma u_4 \rightsquigarrow_2^\Sigma u_5$ also according to the obvious start-bounded strategy; finally, the challenge $u_9 \rightsquigarrow_2^\Sigma u_{10}$ needs a response according to the backward strategy. We will combine the two backward strategies into a single one, but keep the start-bounded ones separate.



The game $G^a_\Sigma(\mathcal{G}_2,\mathcal{G}_1)$ begins as the game $G^b_\Sigma(\mathcal{G}_2,\mathcal{G}_1)$, but with states of the form $(\Xi_i \mapsto x_i, \Psi_i)$, $i \geq 0$, where $\Xi_i \subseteq \Delta^{\mathcal{G}_2}$ and $x_i \in \Delta^{\mathcal{G}_1}$ satisfy (**s**$'_1$) and Ψ_i is a (possibly empty) subset of $\Xi_i^{\rightsquigarrow}$, which indicates initial challenges in start-bounded games. The initial state satisfies (**s**$'_0$). In each round $i > 0$, if $x_{i-1} \in \mathsf{ind}(\mathcal{K}_1)$ then player 2 launches the start-bounded game $G^s_\Sigma(\mathcal{G}_2,\mathcal{G}_1)$ with the initial state $(\emptyset, \Xi_{i-1} \mapsto x_{i-1})$. Otherwise, if $x_{i-1} \notin \mathsf{ind}(\mathcal{K}_1)$, player 2 has two options. First, he can challenge player 1 with the set Ψ_{i-1} (that is, similar to the backward game but with a possibly smaller Ψ_{i-1} in place of $\Xi_{i-1}^{\rightsquigarrow}$); player 1 responds to this challenge with a state $(\Xi_i \mapsto x_i, \Psi_i)$ such that $\Psi_{i-1} \subseteq \Xi_i$, $x_i \rightsquigarrow_1 x_{i-1}$ and (**s**$'_2$) holds. Second, player 2 can launch the start-bounded game $G^s_\Sigma(\mathcal{G}_2,\mathcal{G}_1)$ with the initial state $(\emptyset, \Xi_{i-1} \mapsto x_{i-1})$, where the first challenge of player 2 must be picked from $\Phi_{i-1} = \Xi_{i-1}^{\rightsquigarrow} \setminus \Psi_{i-1}$.

Example 21. We illustrate the ω-winning strategy for player 1 in $G^a_\Sigma(\mathcal{G}_2,\mathcal{G}_1)$ starting from $(\{u_1\} \mapsto x_2, \{u_2\})$, where $\mathcal{G}_2^\Sigma$ is from Example 20 and $\mathcal{G}_1^\Sigma$ looks like $\mathcal{M}_1^\Sigma$ from Example 20 (but with x_i in place of σ_i):



Lemma 22. *For any $u_0 \in \Delta^{\mathcal{G}_2}$, condition* (**$<\omega$**) *holds for arbitrary strategies in the game $G_\Sigma(\mathcal{G}_2,\mathcal{M}_1)$ iff* (**ω**) *holds in $G^a_\Sigma(\mathcal{G}_2,\mathcal{G}_1)$ for some $(\Xi_0 \mapsto x_0, \Psi_0)$ with $u_0 \in \Xi_0$.*

Condition (**ω**) in Lemma 22 is checked in time $O(|\mathsf{ind}(\mathcal{K}_2)| \times 2^{|\Delta^{\mathcal{G}_2} \setminus \mathsf{ind}(\mathcal{K}_2)|} \times |\Delta^{\mathcal{G}_1}|)$, which can be readily seen by analysing the full game graph for $G^a_\Sigma(\mathcal{G}_2,\mathcal{G}_1)$ (similar to that in Example 21). By Theorem 6, we then obtain:

Theorem 23. *For combined complexity, the Σ-query entailment problem is in* 2EXPTIME *for Horn-$\mathcal{ALCHI}$ and Horn-$\mathcal{ALCI}$ KBs and in* EXPTIME *for DL-Lite$^{\mathcal{H}}_{horn}$ and DL-Lite$^{\mathcal{H}}_{core}$ KBs. For data complexity, these problems are all in* P.

Discussion

We have shown that, for all of our DLs, Σ-query entailment and inseparability are in P for data complexity. The next theorem establishes a matching lower bound:

Theorem 24. *For data complexity, Σ-query entailment and inseparability are* P*-hard for DL-Lite$_{core}$ and $\mathcal{EL}$ KBs.*

For combined complexity, ExpTime-hardness of Σ-query inseparability for Horn-$\mathcal{ALC}$ can be proved by reduction of the subsumption problem: we have $\mathcal{T} \models A \sqsubseteq B$ iff $(\mathcal{T}, \{A(a)\})$ and $(\mathcal{T} \cup \{A \sqsubseteq B\}, \{A(a)\})$ are $\{B\}$-query inseparable. We now establish matching lower bounds in the technically challenging cases.

Theorem 25. *For combined complexity, Σ-query entailment and inseparability are* (*i*) *2*ExpTime*-hard for Horn-$\mathcal{ALCI}$ KBs and* (*ii*) ExpTime*-hard for DL-Lite$^{\mathcal{H}}_{core}$ KBs.*

The obtained tight complexity bounds apply to the membership problem for universal UCQ-solutions and to Σ-query inseparability of TBoxes. As a consequence of Theorems 3, 23 and 25 we obtain:

Theorem 26. *For combined complexity, the membership problem for universal UCQ-solutions is 2*ExpTime*-complete for Horn-$\mathcal{ALCHI}$ and Horn-$\mathcal{ALCI}$;* ExpTime*-complete for Horn-$\mathcal{ALCH}$, Horn-$\mathcal{ALC}$, DL-Lite$^{\mathcal{H}}_{horn}$ and DL-Lite$^{\mathcal{H}}_{core}$; and* P*-complete for $\mathcal{EL}$ and $\mathcal{ELH}$. For data complexity, all these problems are* P*-complete.*

Σ-query inseparability of *DL-Lite*$^{\mathcal{H}}_{core}$ TBoxes was known to sit between PSpace and ExpTime [12]. Using the fact that witness ABoxes for *DL-Lite*$^{\mathcal{H}}_{core}$ TBox separability can always be chosen among the singleton ABoxes [12, Theorem 8], we can modify the proof of Theorem 25 to improve the PSpace lower bound:

Theorem 27. *Σ-query inseparability of DL-Lite$^{\mathcal{H}}_{core}$ TBoxes is* ExpTime*-complete.*

For more expressive DLs, TBox Σ-query inseparability is often harder than KB inseparability: for *DL-Lite*$_{horn}$, the space of relevant witness ABoxes for TBox separability is of exponential size and, in fact, TBox inseparability is NP-hard, while KB inseparability is in P. Similarly, Σ-query inseparability of $\mathcal{EL}$ KBs is tractable, while Σ-query inseparability of TBoxes is ExpTime-complete [17]. The complexity of TBox inseparability for Horn-DLs extending *Horn-$\mathcal{ALC}$* is not known.

As for future work, from a theoretical point of view, it would be of interest to investigate the complexity of Σ-query inseparability for KBs in more expressive Horn DLs (e.g., Horn-$\mathcal{SHIQ}$) and non-Horn DLs extending $\mathcal{ALC}$. We conjecture that the game technique developed in this paper can be extended to those DLs as well. Our games can also be used to define *efficient approximations* of Σ-query entailment and inseparability for KBs. The existence of a forward strategy, for example, provides a sufficient condition for Σ-query entailment for all of our DLs. Thus, one can extract a Σ-query module of a given KB $\mathcal{K}$ by exhaustively removing from $\mathcal{K}$ those inclusions and assertions α such that player 1 has a winning strategy in the game $G^{f}_{\Sigma}(\mathcal{G}_2, \mathcal{G}_1)$, where $\mathcal{G}_1$ and $\mathcal{G}_2$ are generating structures for $\mathcal{K} \setminus \{\alpha\}$ and $\mathcal{K}$, respectively. The resulting modules are minimal for our DLs without inverse roles, and we conjecture that in practice they are often minimal for DLs with inverse roles as well; see [12] for experiments testing similar ideas for module extraction from TBoxes.

References

1. Arenas, M., Botoeva, E., Calvanese, D., Ryzhikov, V.: Exchanging OWL 2 QL knowledge bases. In: Proc. of the 23rd Int. Joint Conf. on Artificial Intelligence (IJCAI 2013) (2013)
2. Arenas, M., Botoeva, E., Calvanese, D., Ryzhikov, V., Sherkhonov, E.: Exchanging description logic knowledge bases. In: Proc. of the 13th Int. Conf. on Principles of Knowledge Representation and Reasoning (KR 2012). AAAI Press (2012)
3. Baader, F., Brandt, S., Lutz, C.: Pushing the EL envelope. In: Proc. of the 19th Int. Joint Conf. on Artificial Intelligence (IJCAI-05). pp. 364–369 (2005)
4. Baader, F., Calvanese, D., McGuinness, D., Nardi, D., Patel-Schneider, P.F. (eds.): The Description Logic Handbook: Theory, Implementation and Applications. Cambridge University Press (2003), (2nd edition, 2007)
5. Baier, C., Katoen, J.P.: Principles of Model Checking. MIT Press (2007)
6. Botoeva, E., Kontchakov, R., Ryzhikov, V., Wolter, F., Zakharyaschev, M.: Query inseparability for description logic knowledge bases. In: Proc. of the 14th Int. Conf. on Principles of Knowledge Representation and Reasoning (KR 2014). AAAI Press (2014), full version is available at `www.dcs.bbk.ac.uk/~roman/KR2014.pdf`.
7. Calvanese, D., De Giacomo, G., 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), 385–429 (2007)
8. Eiter, T., Gottlob, G., Ortiz, M., Simkus, M.: Query answering in the description logic Horn-$\mathcal{SHIQ}$. In: Proc. of the 11th European Conf. on Logics in Artificial Intelligence (JELIA 2008). Lecture Notes in Computer Science, vol. 5293, pp. 166–179. Springer (2008)
9. Hustadt, U., Motik, B., Sattler, U.: Data complexity of reasoning in very expressive description logics. In: Proc. of the 19th Int. Joint Conf. on Artificial Intelligence (IJCAI-05). pp. 466–471 (2005)
10. Jiménez-Ruiz, E., Cuenca Grau, B., Horrocks, I., Berlanga Llavori, R.: Supporting concurrent ontology development: Framework, algorithms and tool. Data & Knowledge Engineering 70(1), 146–164 (2011)
11. Kazakov, Y.: Consequence-driven reasoning for Horn-SHIQ ontologies. In: Proc. of the 21st Int. Joint Conf. on Artificial Intelligence (IJCAI 2009). pp. 2040–2045 (2009)
12. Konev, B., Kontchakov, R., Ludwig, M., Schneider, T., Wolter, F., Zakharyaschev, M.: Conjunctive query inseparability of OWL 2 QL TBoxes. In: Proc. of the 25th AAAI Conf. on Artificial Intelligence (AAAI 2011). AAAI Press (2011)
13. Kontchakov, R., Lutz, C., Toman, D., Wolter, F., Zakharyaschev, M.: The combined approach to query answering in DL-Lite. In: Proc. of the 12th Int. Conf. on Principles of Knowledge Representation and Reasoning (KR 2010). AAAI Press (2010)
14. Krötzsch, M., Rudolph, S., Hitzler, P.: Complexities of Horn description logics. ACM Transactions on Computational Logic 14(1), 2 (2013)
15. Lin, F., Reiter, R.: Forget it! In: In Proc. of the AAAI Fall Symposium on Relevance. pp. 154–159 (1994)
16. Lutz, C., Toman, D., Wolter, F.: Conjunctive query answering in the description logic EL using a relational database system. In: Proc. of the 21st Int. Joint Conf. on Artificial Intelligence (IJCAI-09). pp. 2070–2075 (2009)
17. Lutz, C., Wolter, F.: Deciding inseparability and conservative extensions in the description logic EL. Journal of Symbolic Computation 45(2), 194–228 (2010)
18. Mazala, R.: Infinite games. In: Automata, Logics, and Infinite Games. pp. 23–42 (2001)
19. Stuckenschmidt, H., Parent, C., Spaccapietra, S. (eds.): Modular Ontologies: Concepts, Theories and Techniques for Knowledge Modularization, Lecture Notes in Computer Science, vol. 5445. Springer (2009)

20. Wang, Z., Wang, K., Topor, R.W.: Revising general knowledge bases in description logics. In: Proc. of the 12th Int. Conf. on Principles of Knowledge Representation and Reasoning (KR 2010). AAAI Press (2010)