Revisiting Conjunctive Query Entailment for $\mathcal{S}$ (Extended Abstract)

Yazmín Ibáñez-García[1], Jean Christoph Jung[2], Vincent Michielini[3] and Filip Murlak[4]

[1]Cardiff University, United Kingdom
[2]TU Dortmund University, Germany
[3]University of Bordeaux, France
[4]University of Warsaw, Poland

Abstract

We clarify the complexity of answering unions of conjunctive queries over knowledge bases formulated in the description logic $\mathcal{S}$, the extension of $\mathcal{ALC}$ with transitive roles. Contrary to what existing partial results suggested, we show that the problem is in fact 2ExpTime-complete, hardness already holding in the presence of two transitive roles, for Boolean conjunctive queries, and trivial ABoxes. We complement this result by showing that the problem remains in coNExpTime, and is in fact coNExpTime-complete, when the input query is rooted or is restricted to use at most one transitive role (but may use arbitrarily many non-transitive roles).

Keywords

Description Logics, Query Entailment, Transitivity, Computational Logics

1. Overview

In this paper, we aim to complete the complexity landscape for the problem of query answering over knowledge bases expressed in expressive description logics (DLs), that is, logics extending the basic DL $\mathcal{ALC}$. As is common in such an endeavor, we focus on the associated decision problem known as *query entailment*: given a DL knowledge base (KB) consisting of an ABox and a TBox, a query, and a tuple of individuals, the goal is to determine whether the query returns this tuple in every model of the given KB. As query languages we consider *unions of conjunctive queries (UCQs)* and fragments thereof. This problem has been heavily studied and is well understood. Existing results suggest a dichotomy:

- The problem is ExpTime-complete for $\mathcal{ALC}$ and its extensions with role hierarchies ($\mathcal{H}$) or qualified number restrictions ($\mathcal{Q}$) [1, 2], or regular path expressions ($_{reg}$) [3].
- The problem is 2ExpTime-complete for the extension $\mathcal{ALC}_{self}$ of $\mathcal{ALC}$ with the *self*-constructor [3], and for every extension that allows either inverse roles ($\mathcal{I}$) or both $\mathcal{H}$ and transitive roles ($\mathcal{S}$), as long as it remains inside the DL $\mathcal{SHIQ}$, which combines the extensions $\mathcal{S}, \mathcal{H}, \mathcal{I}, \mathcal{Q}$ [4, 5].

It is worth mentioning that even higher complexities are possible. For example, UCQ entailment is (at least) coN2ExpTime-hard yet still decidable for the DL $\mathcal{ALCQIO}$ [6, 7]. In many cases, the restriction to *rooted* queries, that is, connected queries with answer variables, leads to lower complexity. This is, for instance, the case for $\mathcal{ALC}_{self}$ [8] and all DLs between $\mathcal{ALCI}$ and $\mathcal{SHIQ}$, where rooted query entailment is coNExpTime-complete as long as transitive role names are disallowed in queries [1]. Without inverses, the complexity is even lower: entailment of (rooted or not) queries without transitive roles is ExpTime-complete for all DLs between $\mathcal{ALC}$ and $\mathcal{SHQ}$ [1]. Working with rooted queries does not necessarily help if transitive roles are allowed in queries: analyzing the 2ExpTime-hardness proof for query entailment in $\mathcal{SH}$ from [4] shows that rooted query entailment remains 2ExpTime-complete.

DL 2026: 39th International Workshop on Description Logics, July 17–19, 2026, Lisbon, Portugal
ibanezgarciay@cardiff.ac.uk (Y. Ibáñez-García); jean.jung@tu-dortmund.de (J. C. Jung); vincent.michielini@u-bordeaux.fr (V. Michielini); f.murlak@uw.edu.pl (F. Murlak)
0000-0002-1276-904X (Y. Ibáñez-García); 0000-0002-4159-2255 (J. C. Jung); 0000-0002-1413-9316 (V. Michielini); 0000-0003-0989-3717 (F. Murlak)

A notable gap remains for $\mathcal{S}$ without any extensions. For $\mathcal{S}$ the problem has been proven to be coNExpTime-complete when KBs and queries are allowed to use a single transitive role (and no other roles). The lower bound comes from [4] and the upper bound comes from [9]. A 2ExpTime upper bound for the general case follows from the mentioned result for $\mathcal{SH}$ [4]. In the same paper, Eiter et al. also claim an ExpTime upper bound for the case where the ABox is tree-shaped. These results have been regarded as a strong indication that the whole problem may be coNExpTime-complete, challenging the apparent dichotomy [9]. The complexity for rooted query entailment was open until now, but a coNExpTime-lower bound follows from [4].

The aim of this paper is to revisit the query entailment problem for $\mathcal{S}$ and close the mentioned gaps. Our first main result is that, surprisingly, when at least two transitive roles are allowed in the query, the problem is 2ExpTime-hard already for $\mathcal{S}$ (without role hierarchies), and hence 2ExpTime-complete. Importantly, our lower bound works with a trivial ABox $\{A(a)\}$ and thus contradicts the mentioned ExpTime upper bound for the case with tree shaped ABoxes [4]. Indeed, in their argument, Eiter et al. compile input conjunctive queries (CQs) to so-called *pseudo-tree queries (PTQs)*, designed to capture the behavior of CQs over tree-like interpretations, which is sufficient due to the tree-like model property of $\mathcal{S}$. The error seems to arise from a subtle mismatch between the formal definition of PTQs used in the compilation process, and their intuitive understanding as *trees of clusters*, on which the later algorithmic treatment relies.

Besides the presence of two transitive roles in the query, our 2ExpTime-hardness proof relies on the availability of non-rooted queries. We complement our lower bound by showing that both conditions are necessary: UCQ entailment remains in coNExpTime if at most one transitive role is allowed in the query or if the query is rooted.

We develop the two proofs in parallel, using common terminology and data structures whenever possible. At the core, we show a *small witness property* in the following sense: the query is not entailed iff there is a small structure witnessing that. Note that this structure cannot be simply a countermodel since UCQ entailment for $\mathcal{S}$ is not finitely controllable [10]; instead, a countermodel is obtained from the witness via a suitable unraveling. The NExpTime algorithm for *non*-entailment can then just guess a small structure and verify that it is indeed a witness. Our argument is divided into three steps.

Step 1. Reduce UCQ entailment to a special case with trivial ABoxes and UCQs of special shape.

Step 2. Reduce the special case of entailment to the existence of structures called *mosaics*, which are collections of interpretations called *tiles* that respect certain compatibility requirements.

Step 3. Show that both the size of all tiles in a mosaic and their number can be bounded exponentially.

While the overall strategy is familiar, the steps are subtle. Steps 1–2 for the single transitive role case rely on a careful refinement of Eiter et al.'s PTQs, which ensures correctness of the algorithmic treatment, but makes the compilation process significantly harder. The crux of Step 3 is to show that UCQ entailment *is* finitely controllable in the special case where the set of used role names consists of a single transitive role name and queries are acyclic. Going beyond existing results [9], we show that the size of the countermodel can be bounded independently from the query.

The following theorem summarizes our knowledge about UCQ entailment over KBs formulated in $\mathcal{S}$. Besides the discussed cases, it includes also the observation that the problem remains in ExpTime (and is thus not harder than in $\mathcal{ALC}$) when no transitive roles are allowed in the query [1]. Hence, our results also complete the picture in terms of the dependence of the complexity on the number of transitive roles used in the query.

Theorem 1. *UCQ entailment over $\mathcal{S}$ knowledge bases is* 2ExpTime*-complete in general, but:*

- coNExpTime*-complete, when restricted to rooted queries;*
- coNExpTime*-complete, when restricted to queries that use at most one transitive role;*
- ExpTime*-complete, when restricted to queries that do not use any transitive roles.*

This is an extended abstract of [11]; full proofs for all statements in that paper are available in [12].

Acknowledgments

Vincent Michielini was supported by the ANR grand INTENDED (ANR-19-CHIA-0014). Filip Murlak was supported by Poland's NCN grant 2018/30/E/ST6/00042.

Declaration on Generative AI

The authors did not employ any Generative AI tools.

References

[1] C. Lutz, The complexity of conjunctive query answering in expressive description logics, in: Proceedings of 4th International Joint Conference on Automated Reasoning, IJCAR 2008, volume 5195 of *Lecture Notes in Computer Science*, Springer, 2008, pp. 179–193. URL: https://doi.org/10.1007/978-3-540-71070-7_16. doi:10.1007/978-3-540-71070-7_16.

[2] M. Ortiz, M. Simkus, T. Eiter, Worst-case optimal conjunctive query answering for an expressive description logic without inverses, in: Proceedings of the Twenty-Third AAAI Conference on Artificial Intelligence, AAAI 2008, AAAI Press, 2008, pp. 504–510. URL: http://www.aaai.org/Library/AAAI/2008/aaai08-080.php.

[3] B. Bednarczyk, S. Rudolph, How to tell easy from hard: Complexities of conjunctive query entailment in extensions of ALC, J. Artif. Intell. Res. 78 (2023). URL: https://doi.org/10.1613/jair.1.14482. doi:10.1613/JAIR.1.14482.

[4] T. Eiter, C. Lutz, M. Ortiz, M. Simkus, Query answering in description logics with transitive roles, in: Proceedings of the 21st International Joint Conference on Artificial Intelligence, IJCAI 2009, 2009, pp. 759–764. URL: http://ijcai.org/Proceedings/09/Papers/131.pdf.

[5] B. Glimm, I. Horrocks, U. Sattler, Unions of conjunctive queries in SHOQ, in: Proceedings of the 11th International Conference on Principles of Knowledge Representation and Reasoning (KR), AAAI Press, 2008, pp. 252–262.

[6] B. Glimm, Y. Kazakov, C. Lutz, Status QIO: an update, in: R. Rosati, S. Rudolph, M. Zakharyaschev (Eds.), Proceedings of the 24th International Workshop on Description Logics (DL 2011), Barcelona, Spain, July 13-16, 2011, CEUR Workshop Proceedings, CEUR-WS.org, 2011. URL: https://ceur-ws.org/Vol-745/paper_44.pdf.

[7] S. Rudolph, B. Glimm, Nominals, inverses, counting, and conjunctive queries or: Why infinity is your friend!, J. Artif. Intell. Res. 39 (2010) 429–481. URL: https://doi.org/10.1613/jair.3029. doi:10.1613/JAIR.3029.

[8] B. Bednarczyk, Data complexity in expressive description logics with path expressions, in: Proceedings of the Thirty-Third International Joint Conference on Artificial Intelligence, IJCAI 2024, ijcai.org, 2024, pp. 3241–3249. URL: https://www.ijcai.org/proceedings/2024/359.

[9] M. Bienvenu, T. Eiter, C. Lutz, M. Ortiz, M. Simkus, Query answering in the description logic S, in: Proceedings of the 23rd International Workshop on Description Logics (DL 2010), volume 573 of *CEUR Workshop Proceedings*, CEUR-WS.org, 2010. URL: https://ceur-ws.org/Vol-573/paper_20.pdf.

[10] R. Rosati, On the finite controllability of conjunctive query answering in databases under open-world assumption, Journal of Computer and System Sciences 77 (2011) 572–594. doi:https://doi.org/10.1016/j.jcss.2010.04.011.

[11] Y. Ibáñez-García, J. C. Jung, V. Michielini, F. Murlak, Revisiting conjunctive query entailment for $\mathcal{S}$, in: Fortieth AAAI Conference on Artificial Intelligence, AAAI 2026, AAAI Press, 2026, pp. 19152–19159. doi:10.1609/AAAI.V40I23.38989.

[12] Y. A. Ibáñez-García, J. C. Jung, V. Michielini, F. Murlak, Revisiting conjunctive query entailment for $\mathcal{S}$, 2025. doi:10.48550/arXiv.2511.07933. arXiv:2511.07933.