PTime Combined Complexity and FPT in Ontology-Mediated Querying

Pablo Barceló[1], Cristina Feier[2], Carsten Lutz[2], and Andreas Pieris[3]

[1] DCC, U of Chile & IMFD Chile
[2] University of Bremen, Germany
[3] University of Edinburgh, United Kingdom

Ontology-mediated queries (OMQs) based on description logic (DL) ontologies and their complexity have been a subject of intense study [5, 7, 8]. In the full paper [1] reported on in this abstract, we explore the frontiers of two important notions of tractability for OMQs, PTime combined complexity and fixed-parameter tractability (FPT) where the parameter is the size of the OMQ. Given that ontologies can get large in practice, these notions of tractability are arguably more realistic than PTime data complexity as frequently considered in the literature [9, 11, 16, 17].

As usual, we use $(\mathcal{L}, \mathcal{Q})$ to denote the OMQ language where ontologies are formulated in the DL $\mathcal{L}$ and queries are from the query language $\mathcal{Q}$. From now on, we generally mean combined complexity when speaking of complexity. There are only few OMQ languages that have PTime complexity or are FPT (with the parameter being the size of the OMQ) without imposing serious restrictions on the shape of the query or the ontology. An important example for the former is $(\mathcal{ELH}^{dr}_{\bot}, \text{AQ})$ where $\cdot^{dr}$ stands for domain and range restrictions and AQ refers to the class of atomic queries of the form $A(x)$, A a concept name; this result is implicit in [18]. An important example for an OMQ language that is FPT is $(\mathcal{ELHI}_{\bot}, \text{AQ})$; we are not aware of this being stated explicitly anywhere, but it is not too hard to prove using standard means. Note that the (unrestricted) use of the popular conjunctive queries (CQs) and unions thereof (UCQs) as the query language $\mathcal{Q}$ rules out both of the considered complexities independently of the choice of $\mathcal{L}$ since (U)CQ-evaluation (without an ontology) is NP-complete and W[1]-hard, thus most likely not fixed-parameter tractable [14].

A seminal result by Grohe precisely characterizes the (recursively enumerable) classes of CQs over schemas of bounded arity that can be evaluated in PTime: this is the case if and only if for some k, every CQ in the class is equivalent to a CQ of tree width k, unless the assumption from parameterized complexity theory that FPT $\neq$ W[1] fails [15]. Grohe's result also establishes that PTime complexity and FPT coincide for evaluating CQs (for schemas of bounded arity). A generalization to UCQs has been observed by Chen [10]. It has further been observed in [6] that whenever $\mathcal{Q}$ is a class of CQs that can be evaluated in PTime, then the same is true for OMQs from $(\mathcal{ELH}, \mathcal{Q})$. In particular, $\mathcal{Q}$ might be the class of CQs of tree width bounded by some k.

The main aim of the work that we report about is to precisely analyze the frontiers of PTime complexity and FPT for OMQs in which the ontology language is from the $\mathcal{EL}$ and $\mathcal{ELI}$ families of DLs and where $\mathcal{Q}$ are (U)CQs. An

OMQ has *bounded tree width* if the actual query in it has. Our main contributions are the following, assuming that FPT $\neq$ W[1]:

1. the subclasses of $(\mathcal{ELH}_\bot^{dr}, \text{UCQ})$ that admit PTime evaluation are exactly those in which each OMQ is equivalent to an OMQ of bounded tree width;
2. the subclasses of $(\mathcal{ELHI}_\bot, \text{UCQ})$ for which evaluation is in FPT are exactly those in which each OMQ is equivalent to an OMQ of bounded tree width.

In Point 1 (but not in Point 2), we assume that the ABox signature is full. Regarding Point 2, we also show that the runtime of the FPT algorithm can be made single exponential in the parameter. Given that $\mathcal{ELH}_\bot^{dr}$ is a fragment of $\mathcal{ELHI}_\bot$, Points 1 and 2 imply that PTime complexity and FPT coincide in $(\mathcal{ELH}_\bot^{dr}, \text{UCQ})$ when the ABox signature is full. For the 'upper bound' of Point 2, we use existential pebble games adapted in a careful way to OMQs. For the rather non-trivial 'lower bound', we build on Grohe's result. Dealing with non-full ABox signatures is a serious challenge as standard techniques from relational databases such as using the core of a CQ must be replaced by more subtle ones. In Points 1 and 2, equivalence to an OMQ Q of bounded tree width includes the case that Q uses a different ontology than the original OMQ. We also show, however, that in most cases there is no benefit in changing the ontology.

We point out that our tractability results are stronger than those in [6]: adding an ontology can lower the complexity of a (U)CQ and it is in fact not hard to see that there are classes of OMQs from $(\mathcal{EL}, \text{CQ})$ that can be evaluated in PTime, but the class of CQs used in them cannot. More loosely related studies of the combined complexity of OMQs in which the ontology is formulated in DL-Lite$^\mathcal{R}$ and DL-Lite$^\mathcal{R}_\mathsf{horn}$ are in [3, 4]. There is also a loose connection to the rewriting of OMQs based on CQs and expressive DLs such as $\mathcal{ALC}$ into OMQs based on instance queries (and expressive DLs) [12]. For a study of FPT in the context of subsumption, see [19].

We further study the complexity of the meta problem of deciding whether a given OMQ is equivalent to an OMQ of bounded tree width. Our results range from Π_2^p between (DL-Lite$^\mathcal{R}$, CQ) and (DL-Lite$^\mathcal{R}_\mathsf{horn}$, UCQ) via ExpTime between $(\mathcal{EL}, \text{CQ})$ and $(\mathcal{ELH}_\bot^{dr}, \text{UCQ})$ to 2ExpTime between $(\mathcal{ELI}, \text{CQ})$ and $(\mathcal{ELHI}_\bot, \text{UCQ})$. As an important special case, we consider the full ABox signature. There, the complexity drops considerably, to NP, NP, and ExpTime, respectively. The case of the full ABox signature is also interesting because it admits constructions that are close to the case of relational databases, such as (a suitably adapted version of) retracts. Under the full ABox signature, the problems studied here are related to the the evaluation of (U)CQs of bounded tree width over relational databases with integrity constraints [2].

We also take a first glimpse at OMQ languages based on DL-Lite$^\mathcal{F}$. This turns out to be closely related to the evaluation of UCQs over relational databases in the presence of key dependencies, as studied by Figueira [13]. We show that evaluating OMQs that are equivalent to an OMQ of tree width bounded by some k is in FPT and even in PTime when $k = 1$, and that the meta problem of deciding whether an OMQ belongs to this class is decidable in 3ExpTime and NP-complete when $k = 1$. In this part, we assume the full ABox signature and that queries are Boolean. When $k > 1$, we further assume that the ontology cannot be changed.

References

1. Pablo Barceló, Cristina Feier, Carsten Lutz, and Andreas Pieris. When is ontology-mediated querying efficient? In *LICS*. IEEE Computer Society, 2019.
2. Pablo Barceló, Georg Gottlob, and Andreas Pieris. Semantic acyclicity under constraints. In *PODS*, pages 343–354, 2016.
3. Meghyn Bienvenu, Stanislav Kikot, Roman Kontchakov, Vladimir V. Podolskii, Vladislav Ryzhikov, and Michael Zakharyaschev. The complexity of ontology-based data access with OWL 2 QL and bounded treewidth queries. In *PODS*, pages 201–216. ACM, 2017.
4. Meghyn Bienvenu, Stanislav Kikot, Roman Kontchakov, Vladimir V. Podolskii, and Michael Zakharyaschev. Ontology-mediated queries: Combined complexity and succinctness of rewritings via circuit complexity. *J. ACM*, 65(5):28:1–28:51, 2018.
5. Meghyn Bienvenu and Magdalena Ortiz. Ontology-mediated query answering with data-tractable description logics. In *Reasoning Web*, volume 9203 of *LNCS*, pages 218–307. Springer, 2015.
6. Meghyn Bienvenu, Magdalena Ortiz, Mantas Simkus, and Guohui Xiao. Tractable queries for lightweight description logics. In *IJCAI*, pages 768–774. IJCAI/AAAI, 2013.
7. Meghyn Bienvenu, Balder ten Cate, Carsten Lutz, and Frank Wolter. Ontology-based data access: A study through disjunctive datalog, CSP, and MMSNP. *ACM Trans. Database Syst.*, 39(4):33:1–33:44, 2014.
8. Diego Calvanese, Giuseppe De Giacomo, Domenico Lembo, Maurizio Lenzerini, Antonella Poggi, Mariano Rodriguez-Muro, and Riccardo Rosati. Ontologies and databases: The DL-Lite approach. In *Reasoning Web*, pages 255–356, 2009.
9. Diego Calvanese, Giuseppe De Giacomo, Domenico Lembo, Maurizio Lenzerini, and Riccardo Rosati. Data complexity of query answering in description logics. *Artif. Intell.*, 195:335–360, 2013.
10. Hubie Chen. On the complexity of existential positive queries. *ACM Trans. Comput. Log.*, 15(1):9:1–9:20, 2014.
11. Thomas Eiter, Georg Gottlob, Magdalena Ortiz, and Mantas Simkus. Query answering in the description logic Horn-$\mathcal{SHIQ}$. In *JELIA*, volume 5293 of *LNCS*, pages 166–179. Springer, 2008.
12. Cristina Feier, Carsten Lutz, and Frank Wolter. From conjunctive queries to instance queries in ontology-mediated querying. In *IJCAI*, pages 1810–1816, 2018.
13. Diego Figueira. Semantically acyclic conjunctive queries under functional dependencies. In *LICS*, pages 847–856, 2016.
14. Jörg Flum and Martin Grohe. *Parameterized Complexity Theory*. Texts in Theoretical Computer Science. An EATCS Series. Springer, 2006.
15. Martin Grohe. The complexity of homomorphism and constraint satisfaction problems seen from the other side. *J. ACM*, 54(1):1:1–1:24, 2007.
16. Ullrich Hustadt, Boris Motik, and Ulrike Sattler. Reasoning in description logics by a reduction to disjunctive datalog. *J. Autom. Reasoning*, 39(3):351–384, 2007.
17. Adila Krisnadhi and Carsten Lutz. Data complexity in the $\mathcal{EL}$ family of description logics. In *LPAR*, volume 4790 of *LNCS*, pages 333–347. Springer, 2007.
18. Carsten Lutz, David Toman, and Frank Wolter. Conjunctive query answering in the description logic $\mathcal{EL}$ using a relational database system. In *IJCAI*, pages 2070–2075, 2009.
19. Frantisek Simancik, Boris Motik, and Ian Horrocks. Consequence-based and fixed-parameter tractable reasoning in description logics. *Artif. Intell.*, 209:29–77, 2014.