Cardinality Queries over DL-Lite Ontologies (Extended abstract)

Meghyn Bienvenu
CNRS, University of Bordeaux, Bordeaux INP, LaBRI
Talence, France
meghyn.bienvenu@u-bordeaux.fr

Quentin Manière
CNRS, University of Bordeaux, Bordeaux INP, LaBRI
Talence, France
quentin.maniere@u-bordeaux.fr

Michaël Thomazo
Inria, DI ENS, ENS, CNRS, University PSL
Paris, France
michael.thomazo@inria.fr

ABSTRACT

We summarize our recent work [5] on classifying the complexity of answering cardinality queries over DL-Lite ontologies.

KEYWORDS

Ontology-mediated query answering, Counting queries

A major topic in ontology-mediated query answering (OMQA) research has been to understand the complexity of OMQA and identify tractable settings [6, 11, 12]. Nowadays, for the most commonly considered query language, namely, conjunctive queries (CQs), we have an almost complete picture of the complexity landscape for ontologies formulated in a wide range of different description logics (DLs) [2] and rule-based languages [3, 7]. In particular, it has been shown that CQ answering is tractable in data complexity for ontologies expressed in the most commonly considered dialects of the DL-Lite family [1, 9], which are often employed in OMQA. A crucial property of such DL-Lite dialects and other Horn DLs is that they admit a *canonical model*, which is a single (possibly infinite) model that, by virtue of being homomorphically embeddable into every model, is guaranteed to give the correct answers to all CQs.

While CQs are a natural and well-studied class of queries, there are many other relevant forms of database queries that could be potentially be employed in OMQA. In the present work, our focus will be on counting queries, which together with other forms of aggregate queries, are widely used for data analysis, yet still not well understood in the context of OMQA. A natural way to equip CQs with counting is to count the number of distinct query matches for each answer. As the count value may differ between models, [10] advocated a form of certain answer semantics that considers lower and upper bounds on the count value across different models. Their work provided the first investigation of the complexity of answering counting CQs in the presence of ontologies, revealing such queries to be much more challenging to handle than plain CQs: coNP-complete in data complexity for the well-known DL-Lite$_{core}$ and DL-Lite$^{\mathcal{H}}_{core}$ dialects. A recent work by [4] refined and generalized the complexity results from [10] to a wider class of counting queries and identified a restricted scenario with very low (TC^0-complete) data complexity: rooted CQs coupled with DL-Lite$_{core}$ ontologies. A similar tractability result for connected rooted CQs was proven independently by [8], who also initiated a study of the impact of other restrictions on query shape and developed the first query rewriting procedure for counting CQs. Notably, both the aforementioned TC^0 result and the rewriting procedure crucially relied upon showing that the canonical model gives the right answers under the considered restrictions.

While recent studies have improved our understanding of the complexity of counting CQs, there nevertheless remain many unanswered questions. In this work, we focus on Boolean atomic counting queries of the form $\exists z.\mathrm{A}(z)$ and $\exists z_1, z_2.\mathrm{R}(z_1, z_2)$, which we term *cardinality queries* as they correspond to the natural task of determining (bounds on) the cardinality of a given concept or role name. The data complexity of answering such basic counting queries remains completely open for DL-Lite$_{core}$ ontologies, whilst for DL-Lite$^{\mathcal{H}}_{core}$, the problem is known to be P-hard and in coNP [8]. The main results of our investigation are displayed in Table 1. We show that when ontologies are expressed in DL-Lite$_{core}$, cardinality query answering is tractable in data complexity and enjoys the lowest possible complexity (TC^0-complete). For cardinality queries based upon a concept atom, TC^0 membership holds even for the fragment of DL-Lite$^{\mathcal{H}}_{core}$ obtained by disallowing negative role inclusions. By contrast, for role cardinality queries, we show that coNP-hard situations arise in DL-Lite$^{\mathcal{H}}_{pos}$, which allows only positive concept and role inclusions. In fact, we obtain a complete data complexity classification for DL-Lite$^{\mathcal{H}}_{pos}$, showing that every ontology-mediated query is either TC^0-complete, coNP-complete, or is in P and logspace-equivalent to the complement of Perfect Matching (whose precise complexity is a longstanding open problem). The preceding classification does not extend to DL-Lite$^{\mathcal{H}}_{core}$: we identify new sources of coNP-hardness and further exhibit L-complete cases. We find it intriguing that such complex behaviour arises in what appears at first glance to be a simple OMQA setting. Moreover, in all of the tractable cases we identify, the canonical model may not yield the minimum cardinality, and query answering involves solving non-trivial optimization problems. This led us to devise an entirely new approach based upon exploring a space of strategies to find the optimal way of merging witnesses for existential axioms.

	Concept	Role
DL-Lite$_{core}$	TC^0-c	TC^0-c
DL-Lite$^{\mathcal{H}}_{pos}$	TC^0-c†	TC^0-c | co-PM-c | coNP-c
DL-Lite$^{\mathcal{H}}_{core}$	TC^0-c | L-c | coNP-c | ?	TC^0-c | L-c | co-PM-c | coNP-c | ?

Table 1: Data complexity of cardinality queries based upon concept / role atoms for various DL-Lite dialects. †: holds for all DL-Lite$^{\mathcal{H}}_{core}$ ontologies without negative role inclusions.

ACKNOWLEDGMENTS

Partially supported by ANR project CQFD (ANR-18-CE23-0003)

REFERENCES

[1] Alessandro Artale, Diego Calvanese, Roman Kontchakov, and Michael Zakharyaschev. 2009. The DL-Lite Family and Relations. *Journal of Artificial Intelligence Research (JAIR)* 36 (2009), 1–69.

[2] Franz Baader, Ian Horrocks, Carsten Lutz, and Ulrike Sattler. 2017. *An Introduction to Description Logic.* Cambridge University Press.

[3] Jean-François Baget, Michel Leclère, Marie-Laure Mugnier, and Eric Salvat. 2011. On rules with existential variables: Walking the decidability line. *Journal of Artificial Intelligence (JAR)* 175, 9-10 (2011), 1620–1654.

[4] Meghyn Bienvenu, Quentin Manière, and Michaël Thomazo. 2020. Answering Counting Queries over DL-Lite Ontologies. In *Proc. of the 29th International Joint Conference on Artificial Intelligence (IJCAI).* 1608–1614.

[5] Meghyn Bienvenu, Quentin Manière, and Michaël Thomazo. 2021. Cardinality Queries over DL-Lite Ontologies. In *Proc. of the 30th International Joint Conference on Artificial Intelligence (IJCAI).* 1801–1807.

[6] Meghyn Bienvenu and Magdalena Ortiz. 2015. Ontology-Mediated Query Answering with Data-Tractable Description Logics. In *Tutorial Lectures of the 11th Reasoning Web International Summer School.* 218–307.

[7] Andrea Calì, Georg Gottlob, and Thomas Lukasiewicz. 2012. A general Datalog-based framework for tractable query answering over ontologies. *Journal of Web Semantics (JWS)* 14 (2012), 57–83.

[8] Diego Calvanese, Julien Corman, Davide Lanti, and Simon Razniewski. 2020. Counting Query Answers over a DL-Lite Knowledge Base. In *Proc. of the 29th International Joint Conference on Artificial Intelligence (IJCAI).* 1658–1666.

[9] Diego Calvanese, Giuseppe De Giacomo, Domenico Lembo, Maurizio Lenzerini, and Riccardo Rosati. 2007. Tractable Reasoning and Efficient Query Answering in Description Logics: The DL-Lite Family. *Journal of Automated Reasoning (JAR)* 39, 3 (2007), 385–429.

[10] Egor V. Kostylev and Juan L. Reutter. 2015. Complexity of answering counting aggregate queries over DL-Lite. *Journal of Web Semantics (JWS)* 33 (2015), 94–111.

[11] Antonella Poggi, Domenico Lembo, Diego Calvanese, Giuseppe De Giacomo, Maurizio Lenzerini, and Riccardo Rosati. 2008. Linking Data to Ontologies. *Journal of Data Semantics* 10 (2008), 133–173.

[12] Guohui Xiao, Diego Calvanese, Roman Kontchakov, Domenico Lembo, Antonella Poggi, Riccardo Rosati, and Michael Zakharyaschev. 2018. Ontology-Based Data Access: A Survey. In *Proc. of the 27th International Joint Conference on Artificial Intelligence (IJCAI).* 5511–5519.