Ontology-Based Query Answering for Probabilistic Temporal Data (Abstract)*

Patrick Koopmann[0000−0001−5999−2583]

Institute for Theoretical Computer Science, Technische Universität Dresden, Germany
firstname.lastname@tu-dresden.de

In *ontology-based query answering* (OBQA), queries are evaluated on a set of data with respect to an ontology, which specifies background knowledge about the domain. Specifically, a user may query information that is only implicit in the data, but logically entailed when combined with the ontology. While originally designed for querying data that are certain and static, applications such as situation recognition or querying of historical data motivate the need for OBQA where data can be both temporal and probabilistic. For instance, if temporal data are obtained using imprecise sensors, image recognition techniques or language recognition techniques, they can be more adequately represented using probabilities. To employ OBQA in such a setting, we propose *temporal probabilistic queries* (*TPQs*), a query language that can be used to describe temporal patterns involving probability bounds on subqueries.

We assume a representation of the data in form of a sequence of probabilistic data sets, which may have been obtained using further preprocessing and windowing operations. Assume for instance these data are obtained by a smartphone app for health monitoring, where motion and blood pressure sensors are used to detect with a certain probability whether a patient is exercising, and whether their blood pressure is high. We can then use the following TPQ to detect situations in which, during the last 10 time units, the patient was with a low probability exercising, until he had with a high probability a high blood pressure:

$$\bigcirc^{-10}\big(P_{<0.2}\mathsf{Excercising}(x)\,\mathcal{U}\,P_{>0.7}\mathsf{HighBloodPressure}(x)\big).$$

Our language is an extension of the temporal query language from [4,1], which extends conjunctive queries (CQs) with operators from linear temporal logic (LTL), and to which we add probability operators as in the example. As in [4,1], we consider classical DLs for specifying the ontology. This allows us to use *rigid roles* in the ontology, which are roles whose interpretation remains static over time. Rigid roles make reasoning easily undecidable in DLs that support LTL constructs as part of the DL [9]. For this reason, we focus on extensions of the query language rather than the DL in this work. (Extensions of both query language and DL are considered in [7].) The probabilistic component follows the semantics of probabilistic ABoxes from [6], which itself is based on semantics for probabilistic databases [5], and has since been used in other works on probabilistic OBQA [3,2].

* Supported by DFG in the CRC 912 (HAEC).

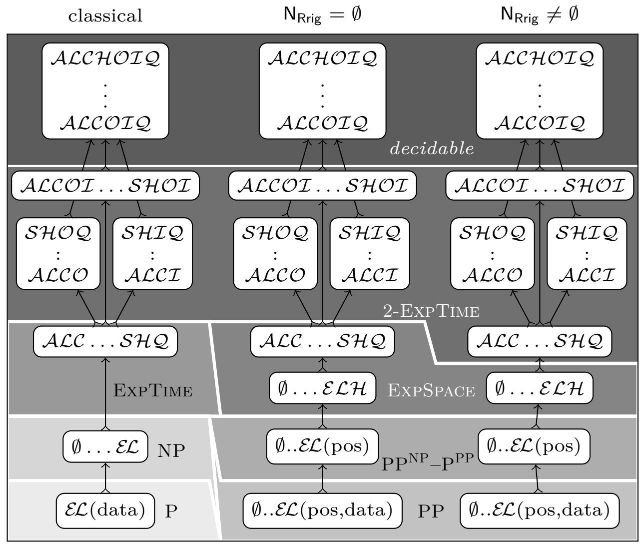

Fig. 1. Complexity of TPQ Entailment vs. classical CQ entailment.

We establish a more or less complete picture of the complexity of query entailment using our query language for ontologies expressed in various DLs, as shown in Figure 1. The figure compares the complexities of classical CQ entailment (on the left) with that of probabilistic temporal query entailment without rigid roles ($\mathsf{N}_{\mathsf{Rrig}} = \emptyset$), as well as with rigid roles ($\mathsf{N}_{\mathsf{Rrig}} \neq \emptyset$). Here, $\emptyset$ stands for the case without ontology, (data) refers to data complexity, and (pos) to a restricted query language without negation. Except for the P^{PP} upper bound, all results in this figure are tight.

For DLs that involve nominals or inverse roles, TPQ entailment is not harder than classical query entailment. However, as it turns out, TPQ entailment is also ExpSpace-hard already in the absence of an ontology, and if only a single probabilistic ABox is considered. This is a big increase compared to both temporal and probabilistic query entailment, which both can be performed within PSpace without ontologies [4,6]. The source of this high complexity comes from the explicit and implicit negation operators in the DL and the query language. By choosing a DL without negation, and restricting the query language to *positive TPQs*, which may not use negation or specify probability upper bounds, we obtain a drop in complexity to P^{PP}, a complexity class contained in PSpace. In fact, if the nesting depth of probability operators is bounded, positive TPQ entailment is not harder than classical CQ entailment in probabilistic database systems (PP in data and $\mathrm{PP}^{\mathrm{NP}}$ in combined complexity).

The full paper will be published in the proceedings of the 33rd AAAI Conference on Artificial Intelligence (AAAI'19) [8].

References

1. Baader, F., Borgwardt, S., Lippmann, M.: Temporal query entailment in the description logic $\mathcal{SHQ}$. J. Web Sem. **33**, 71–93 (2015)
2. Baader, F., Koopmann, P., Turhan, A.: Using ontologies to query probabilistic numerical data. In: Proceedings of FroCoS 2017. pp. 77–94. Springer International (2017)
3. Borgwardt, S., Ceylan, İ.İ., Lukasiewicz, T.: Ontology-mediated queries for probabilistic databases. In: Proceedings of AAAI 2017. pp. 1063–1069. AAAI Press (2017)
4. Borgwardt, S., Thost, V.: Temporal query answering in DL-Lite with negation. In: Proceedings of GCAI 2015. pp. 51–65 (2015), `http://www.easychair.org/publications/paper/245305`
5. Dalvi, N.N., Suciu, D.: Efficient query evaluation on probabilistic databases. VLDB J. **16**(4), 523–544 (2007)
6. Jung, J.C., Lutz, C.: Ontology-based access to probabilistic data with OWL QL. In: Cudré-Mauroux, P., Heflin, J., Sirin, E., Tudorache, T., Euzenat, J., Hauswirth, M., Parreira, J.X., Hendler, J., Schreiber, G., Bernstein, A., Blomqvist, E. (eds.) Proceedings of ISWC 2012. Lecture Notes in Computer Science, vol. 7649, pp. 182–197. Springer (2012). https://doi.org/10.1007/978-3-642-35176-1_12, `https://doi.org/10.1007/978-3-642-35176-1_12`
7. Koopmann, P.: Maybe eventually? Towards combining temporal and probabilistic description logics and queries. In: Proceedings of DL 2019. CEUR-WS.org (2019)
8. Koopmann, P.: Ontology-based query answering for probabilistic temporal data. In: Hentenryck, P.V., Zhou, Z.H. (eds.) Proceedings of AAAI 2019. AAAI Press, Honolulu, USA (2019), to appear.
9. Lutz, C., Wolter, F., Zakharyaschev, M.: Temporal description logics: A survey. In: Demri, S., Jensen, C.S. (eds.) Proceedings of TIME 2008. pp. 3–14. IEEE Press (2008). https://doi.org/10.1109/TIME.2008.14