Mixed-World Reasoning with Existential Rules under Active Domain Semantics (Abstract)

Meghyn Bienvenu[1] and Pierre Bourhis[2]

[1] CNRS & University of Bordeaux, Talence, France
[2] CNRS, University of Lille, Inria Lille, Lille, France

Abstract. This extended abstract summarizes our recent study [4] of existential rules with closed predicates and active-domain semantics.

The vast majority of work on ontology-mediated query answering (OMQA) adopts the *open world assumption*, whereby facts that are not present in the data instance are treated as unknown. Formally, each knowledge base (KB) $(\mathcal{R}, I)$, consisting of an ontology $\mathcal{R}$ and data instance I, give rise to a set of models, defined as the first-order structures that make true all facts in I and satisfies all rules (or axioms) in $\mathcal{R}$. When querying a KB, we are interested in the *certain answers* that hold for every model of the KB. By contrast, relational databases make the *closed-world assumption*, where each instance I is interpreted as the finite structure whose domain is the *active domain* of I (i.e., the constants explicitly mentioned in I) and which makes true precisely the facts in I.

In practice, it seems natural to assume that applications may involve some predicates whose contents are fully known and others for which we have only partial information. This motivates the consideration of *mixed-world semantics*, where the set of predicates is partitioned into closed and open predicates, and models are required to coincide with the instance on the closed predicates. Mixed-world OMQA was first explored for description logic (DL) ontologies [10–12] and has only recently been studied for existential rules [3], another prominent class of ontology languages [1, 2, 9, 5]. Further variants of traditional OMQA semantics can be obtained by placing additional restrictions on the models. In particular, as the classical semantics considers arbitrary models with possibly infinite domains, It is interesting to consider the restriction to finite models [8, 7], or models with fixed or bounded-size domains as in [6, 13].

The research summarized here combines elements of these different lines of research by exploring OMQA with existential rules under a hybrid mixed-world active-domain semantics, in which we can use both closed and open predicates, and the semantics is based upon active-domain models whose domains are equal to the active domain of the instance. Such a semantics is appropriate for scenarios in which all relevant constants are made available in the instance. A possible use case, which served as a motivation for our work, is in analyzing the trajectories of people circulating in a geographically restricted area (e.g., industrial facility, closed medical facility) that is equipped with captors and secure entry and exit (so that the set of people in the area is known at each timepoint).

We have analyzed the data complexity of three central reasoning tasks in this setting which are to determine whether a mixed-world knowledge base (MWKB) is satisfiable, and whether a Boolean conjunctive query (BCQ) is certain or possible w.r.t. a MWKB (the latter can be equivalently viewed as determining whether a BCQ is not certainly false). We show that all three tasks are intractable (NP- or coNP-complete in data complexity) for MWKBs based upon arbitrary existential rules. This is unsurprising in light of previous negative results obtained for querying lightweight DL ontologies with closed predicates [10]. However, for the linear fragment (where rule bodies may contain at most one atom with an open predicate, and arbitrarily many closed atoms), we establish a maximal model property, which we exploit to to derive tractability results. Specifically, we show that satisfiability, possibility of BCQs, and certainty of facts, are all PTIME-complete in data complexity. Moreover, all of the PTIME upper bounds remain valid in the presence of useful extensions, like (in)equality atoms, negated closed atoms, and disjunction in ruleheads.

Motivated by these encouraging results, we investigate the linear fragment in more detail. It is a well-known result in descriptive complexity that semi-positive Datalog (i.e., Datalog rules that may use negated extensional atoms in their bodies) captures all PTIME queries over ordered instances with min and max, i.e., instances I with a binary predicate $Succ$ providing a successor relation among all constants in $\mathsf{adom}(I)$, and unary predicates Min and Max containing the smallest and largest constants. We study the expressive power of the linear fragment and prove our most technically challenging result, namely, that atomic queries coupled with mixed-world linear existential rules, extended with either closed negated atoms or disjunctive ruleheads and interpreted under either certain or possible active-domain semantics, can express all PTIME queries over ordered instances (we do not require min and max to be given as input, as we can compute them from the $Succ$ relation).

While the (extended) linear fragment has desirable computational properties, it cannot express some useful constructs, such as functionality. To broaden its applicability, we introduce a general method of approximating unrestricted existential rules by means of linear rules, which roughly works as follows. Given an arbitrary ruleset, we can first compute the certain and possible facts using only its linear rules, which yields a subset of the 'true' certain facts and a superset of the possible ones. We store these facts in new closed predicates linked by rules to the original open predicates, create new linear rules from the non-linear rules by replacing all open atoms but one with one with a closed atom (using the new closed predicates), then recompute the certain and possible facts. Iterating this process until fixpoint allows us to obtain more refined approximations of the certain and possible facts of the original ruleset.

We envision several continuations to this work, including: studying the combined complexity and the impact of using a fixed but succinctly represented domain, further exploring the expressive power of different fragments, and developing and implementing efficient PTIME algorithms for the linear fragment (avoiding the construction of full maximal models).

References

1. Baget, J., Leclère, M., Mugnier, M., Salvat, E.: Extending decidable cases for rules with existential variables. In: Proceedings of IJCAI (2009)
2. Baget, J.F., Leclère, M., Mugnier, M.L., Salvat, E.: On rules with existential variables: Walking the decidability line. Artificial Intelligence **175**(9-10), 1620–1654 (2011)
3. Benedikt, M., Bourhis, P., ten Cate, B., Puppis, G.: Querying visible and invisible information. In: Proceedings of LICS (2016)
4. Bienvenu, M., Bourhis, P.: Mixed-world reasoning with existential rules under active domain semantics. In: Proceedings of IJCAI (2019)
5. Calì, A., Gottlob, G., Lukasiewicz, T.: A general datalog-based framework for tractable query answering over ontologies. Journal of Web Semantics **14**, 57–83 (2012)
6. Gaggl, S.A., Rudolph, S., Schweizer, L.: Fixed-domain reasoning for description logics. In: Proceedings of ECAI (2016)
7. Gogacz, T., Ibáñez-García, Y.A., Murlak, F.: Finite query answering in expressive description logics with transitive roles. In: Proceedings of KR (2018)
8. Ibáñez-García, Y.A., Lutz, C., Schneider, T.: Finite model reasoning in Horn description logics. In: Proceedings of KR (2014)
9. Krötzsch, M., Rudolph, S.: Extending decidable existential rules by joining acyclicity and guardedness. In: Proceedings of IJCAI (2011)
10. Lutz, C., Seylan, I., Wolter, F.: Ontology-based data access with closed predicates is inherently intractable (sometimes). In: Proceedings of IJCAI (2013)
11. Lutz, C., Seylan, I., Wolter, F.: Ontology-mediated queries with closed predicates. In: Proceedings of IJCAI (2015)
12. Ngo, N., Ortiz, M., Simkus, M.: Closed predicates in description logics: Results on combined complexity. In: Proceedings of KR (2016)
13. Rudolph, S., Schweizer, L.: Not too big, not too small... complexities of fixed-domain reasoning in first-order and description logics. In: Proceedings of EPIA (2017)