Understanding and Improving Reformulation-Based Query Answering Performance in RDF

Damian Bursztyn[1], François Goasdoué[2] and Ioana Manolescu[1]

[1] INRIA & U. Paris-Sud, France [2] U. Rennes 1 & INRIA, France

Abstract. Answering queries over Semantic Web data, i.e., RDF graphs, must account for both *explicit* and *implicit* data, entailed by the explicit data and the *semantic constraints* holding on them. Two main query answering techniques have been devised, namely *Saturation*-based (Sat) which precomputes and adds to the graph all implicit information, and *Reformulation*-based (Ref) which reformulates the query based on the graph constraints, so that evaluating the reformulated query directly against the explicit data (i.e., without considering the constraints) produces the query answer.

While Sat is well known, Ref has received less attention so far. In particular, reformulated queries often perform poorly if the query is complex. Our work [2] includes a large set of Ref techniques, including but not limited to one we proposed recently [1]. The audience will be able to analyze and understand the performance challenges they raise. In particular, we show how a *cost-based* Ref *approach* allows avoiding reformulation performance pitfalls.

References

1. D. Bursztyn, F. Goasdoué, and I. Manolescu. Optimizing reformulation-based query answering in RDF. In *EDBT*, 2015.
2. D. Bursztyn, F. Goasdoué, and I. Manolescu. Reformulation-based query answering in RDF: alternatives and performance". In *VLDB*, 2015.