Planning Regular Data Path Queries

Thomas Mulder*

{t.mulder@tue.nl}
Technische Universiteit Eindhoven

Abstract

We investigate the problem of query planning for Regular Data Path Queries (RDPQ). RDPQs are an extension of Regular Path Queries (RPQ) which allow for the interaction with data associated with nodes in a graph through a finite number k of registers [2]. RPQs enjoy a wide range of applications in analyzing connected data from social-, information-, technological- and biological networks [1]. The interaction with data, as opposed to only considering topological constraints as RPQs do, is vital in order to utilise the value that exists in this data. While the theoretical capabilities and limitations of RDPQs and related formalism's have been explored [3][4], the more practical challenges of query planning and execution have, to the best of our knowledge, not yet been addressed. As a first step towards the future goal of presenting a complete query optimizer that covers RDPQs, we show that optimising a query plan with respect to its interactions with the topology of a graph and its interactions with the data of a graph are orthogonal problems. That is, well-performing query plans cannot, in general, be obtained by optimising for topology and data independently. Rather, these two types of interactions must be considered together at every stage of query planning.

References

[1] Angela Bonifati, George, Hannes Voigt, and Nikolay Yakovets. *Querying graphs.* Morgan & Claypool Publishers, 2018

[2] Leonid Libkin, Wim Martens, and Domagoj Vrgoč. 2016. Querying Graphs with Data. *J. ACM* 63, 2, Article 14 (March 2016), 53 pages.

[3] Pablo Barceló, Gaëlle Fontaine, and Anthony Widjaja Lin. Expressive path queries on graphs with data. *Logical Methods in Computer Science*, 11(4:1), 2015.

[4] Jelle Hellings, Bart Kuijpers, Jan Van den Bussche, and Xiaowang Zhang. Walk logic as a framework for path query languages on graph databases. In *Proceedings of the 16th International Conference on Database Theory*, ICDT '13, pages 117-128, New York, NY, USA, 2013. ACM.

*Under supervision of Nikolay Yakovets and George Fletcher