

Learning in Query Optimization over Knowledge Graphs: From Adaptive Techniques to Neuro-Symbolic Optimizers ... And Back?

Maribel Acosta¹

¹Technical University of Munich, Germany

Abstract

Query optimization has traditionally relied on the optimize-then-execute paradigm, which, while effective in static settings, faces significant limitations in dynamic and complex environments¹ such as knowledge graphs on the web. In this keynote, I will explore how early database adaptive techniques² provided the first steps toward online learning query optimization over knowledge graphs³, allowing systems to adjust during execution. I will present results on when adaptivity enhances continuous performance⁴ in knowledge graphs and when it falls short.

Following this, I will discuss the rise of neuro-symbolic optimizers⁵, which are recent innovations that combine machine learning with symbolic processing (e.g., rules, statistics, data summaries, etc.). Therefore, neuro-symbolic optimizers promise to deliver more accurate results than traditional optimizers⁶, especially in the presence of increasingly complex workloads and scenarios. This raises a fundamental question: Can neuro-symbolic optimizers eliminate the need for adaptive techniques, or will systems still require the flexibility to adapt during execution? I will conclude with an exploration of this open question, the challenges of embedding machine learning into query optimization, and the interplay of machine learning and adaptivity, with a special focus on knowledge graphs and their unique requirements.

¹A. Deshpande, Z. Ives, V. Raman, et al., Adaptive query processing, Foundations and Trends in Databases 1 (2007) 1–140.

²R. Avnur, J. M. Hellerstein, Eddies: Continuously adaptive query processing, in: Proceedings of the 2000 ACM SIGMOD International conference on Management of data, 2000, pp. 261–272.

³M. Acosta, M.-E. Vidal, Networks of linked data eddies: An adaptive web query processing engine for RDF data, in: The Semantic Web-ISWC 2015: 14th International Semantic Web Conference, Bethlehem, PA, USA, October 11–15, 2015, Proceedings, Part I 14, Springer, 2015, pp. 111–127.

⁴M. Acosta, M.-E. Vidal, Y. Sure-Vetter, Diefficiency metrics: measuring the continuous efficiency of query processing approaches, in: International Semantic Web Conference, Springer, 2017, pp. 3–19.

⁵G. Li, X. Zhou, L. Cao, Machine learning for databases, in: Proceedings of the First International Conference on AI-ML Systems, 2021, pp. 1–2.

M. Acosta, C. Qin, T. Schwabe, Neuro-symbolic query optimization in knowledge graphs, in: Handbook on Neurosymbolic AI and Knowledge Graphs, volume 400 of Frontiers in Artificial Intelligence and Applications, IOS Press, 2025, pp. 624–643.

⁶T. Schwabe, M. Acosta, Cardinality estimation over knowledge graphs with embeddings and graph neural networks, Proceedings of the ACM on Management of Data 2 (2024) 1–26.

AMW 2024: 16th Alberto Mendelzon International Workshop on Foundations of Data Management, September 30th–October 4th, 2024, Mexico City, Mexico

✉ maribel.acosta@tum.de (M. Acosta)

🌐 <https://www.cs.cit.tum.de/cde/people/maribel-acosta/> (M. Acosta)

🆔 0000-0002-1209-2868 (M. Acosta)



© 2024 Copyright for this paper by its authors. Use permitted under Creative Commons License Attribution 4.0 International (CC BY 4.0).