

Worst-Case Optimal Joins in Relational Database

Reading

Freitag et al. Adopting Worst-Case Optimal Joins in Relational Database Systems. PVLDB 2020.
<https://db.in.tum.de/~freitag/papers/p1891-freitag.pdf>

Work Packages

WP1. Implement and optimize the algorithms that are specified in Algorithms 1-3 (Hash Trie).

You can define your own data format to store the data (it does not need to be in a relational database). Think about different design choices here and the impact of data layout on algorithm performance and implementations.

You can use a subset of queries in this paper as your target workload — you don't have to run the full TPCH, but some representative queries (potentially simplified).

WP2. Implement and optimize Algorithm 4.

You can assume that you have complete information about a query optimizer and cardinality estimator (you do not have to build a query optimizer).

What are the impacts of this optimization on the execution of your implementations? Can your performance metrics reflect this?