Improving BAO’s worst-case performance

R244 Large-scale data processing and optimisation
Presentation by Martin Graf on 30/11/2022
Quick Recap: Bao

• Bao = Machine Learning + Existing Query Optimiser
• Impressive speed-ups
• Bao comes “batteries included”
There Is Always a Catch

• Performance of fast queries negatively affected
Contribution: Improving BAO’s worst-case performance

• “Disable” Bao where queries are predicted to be fast
• Will not reduce latency through Bao’s architecture
• ➔ Will not benefit supercomputers
• **Will** benefit small systems
Deliverables

• Implement “testing” stage to prevent Bao from running when query is predicted to be fast
• Run performance measurements with IMDb workload for default / adopted Bao implementation on a machine with 2 / 4 CPUs where queries are executed sequentially / 4 at a time when using recommended 48 hint sets
References

Improving BAO’s worst-case performance

R244 Large-scale data processing and optimisation
Presentation by Martin Graf on 30/11/2022