

Group 13 Revised Project Proposal

Members: Daniel Lindsey, Chase Keller, Long Nguyen

Resubmission Date: September 26th

Project Type: Fix and Win

Topic and Sources: SQL query unnesting for improving execution times and reducing total query duration.

- [A Formalization of Top-Down Unnesting](#)
- [Improving Unnesting of Complex Queries](#)

Problem(s) Addressed:

- Little real-world evidence is provided detailing the actual efficiency boost unnesting provides when executing SQL queries.
- Current unnesting algorithms only address specific query patterns/SQL constructs. Possibility of some constructs not being supported.
- Further efficiency improvements.

Project Goal and Motivation:

The goal of our project is to provide additional quantitative details covering just how much more efficient unnested SQL queries are compared to their nested counterparts. We will also attempt to identify additional edge cases that are not currently covered by the algorithms described within 'Improving Unnesting of Complex Queries'. For these cases we will then provide algorithms to efficiently unnest them in line with the current methodology. Additionally, we seek to further improve efficiency by testing modifications to the provided algorithms.

Our motivation for this project is based on increasing our understanding of how query operations can be interchanged and how this interchangeability can result in improved execution times. Unnesting initially seems like a simple concept but is deceptively complex which provides us with an interesting challenge to tackle. This project will also give us exposure to advanced queries with multiple sub-queries which is something we don't have much experience with so far in our academic careers.

Rough Idea of the proposed method:

- We will study unnesting, reverse query engineering, and SQL to Relational Algebra transformations.
- We will provide graphical evidence detailing the differences between query execution duration between nested and unnested SQL queries, identify and construct algorithms for unnesting uncovered edge cases, and analyze the provided algorithms for possible points of improvement.

Group 13 Revised Project Proposal

Testing Plan:

We will begin by further researching query unnesting as well as SQL-to-Relational-Algebra transformations to gain a better understanding of how SQL queries are unnested. We will then gather examples of real-world datasets as well as our own self-generated testing sets to graphically analyze the differences in query duration between nested and unnested SQL queries. From this we can definitively provide clear evidence of the efficiency and benefits of unnesting.

We will then analyze the provided algorithms and identify possible improvements. This portion of the project will be focused on identifying edge cases that are not supported. We will then create our own algorithm that covers the edge case and show how query duration has been reduced after implementing the modifications. There will be multiple rounds of testing to determine whether we may need to alter our modifications or commit more time to identifying edge cases if finding them proves difficult.

Project Schedule/Milestones:

Date	Milestone Event
September 22 nd	Research Ways to Improve Algorithms
September 29 th	Implement Improved Algorithm
October 6 th	Document Changes and Outline Future Improvements
October 13 th	Progress Report Due
October 20 th	Start Testing
October 27 th	Make Necessary Revisions/ Redo Testing
November 3 rd	Compile Test Results
November 10 th	Start Presentation
November 17 th	Presentation Due / Start Final Report
December 7 th	Complete Final Report

Member Contribution:

Each member provided equal contributions in the creation of this project proposal.

