Miniproject 1 Reflection

Initial Query Development (Task 1)

Our team started with low to medium SQL experience, which provided an interesting baseline for comparing manual versus LLM-assisted query development.

Our manual development process followed these steps:
1. Started with basic query components
2. Gradually increased complexity through nesting
3. Iteratively tested and improved the query
4. Generated a query from an LLM for comparison
5. Noticed that both queries were wrong (interpretation of “friends” was wrong)
6. Started over
6. Found a correct query
7. Improved its performance by consulting Google & LLM

Key findings:
- Both the manual as well as the generated query were wrong in the first try
- Generating with ChatGPT 4o gave wrong results, o1-preview (with reasoning) was almost correct, there is just one foreign like missing for one message. When asking to correct the query it failed.
- The manual approach took ~4 hours, o1-preview took 1,5 minute
- The first working human generated query took 50s to execute. After researching how to optimize it, we could improve it to an approximate execution time of 500ms. The LLM-generated query has a similar execution time.

Our interaction with ChatGPT:
https://chatgpt.com/share/673def55-67cc-8004-96c3-0460f0d12fce

Query Extension (Task 2)

When extending the queries with an LLM:
- Both extensions worked well, the results from the original queries were not changed (the error of the original LLM query stayed)

When manually extending a generated query:

- It was hard to understand the generated query while working on it. Our understanding of the manually written query is way deeper, so it was easier to manipulate it.

We personally prefer the chat interaction with the LLM for query development, as it allows for fast and comprehensive explanations. Nevertheless, the LLM generated queries were not correct and we were not able to prompt the LLM to improve its query.

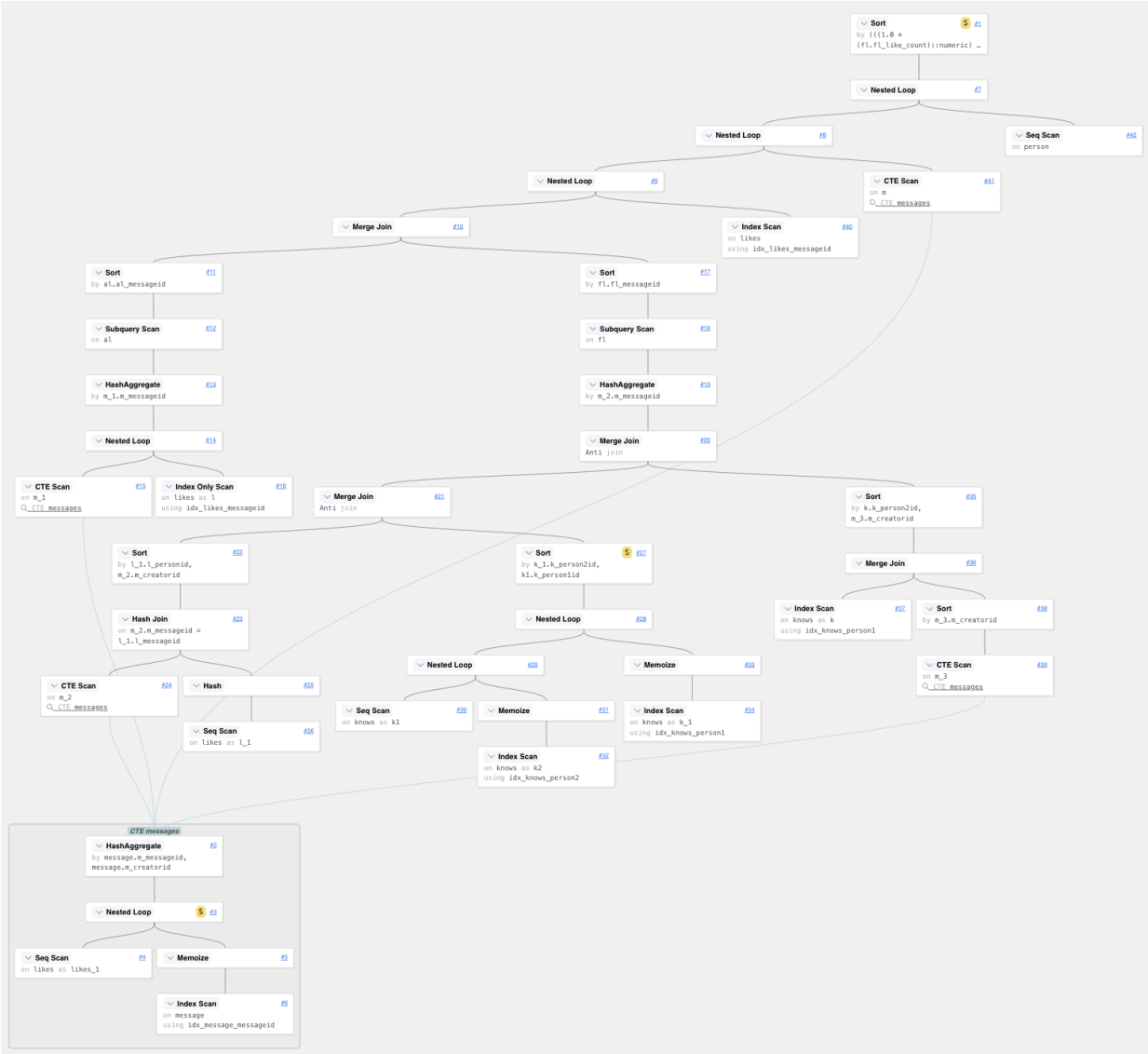
Query Execution (Task 3)

We looked at the manually crafted query and its manually crafted extension
We provided the sql scheme and the query to our LLM and asked it to find a feasible query plan.

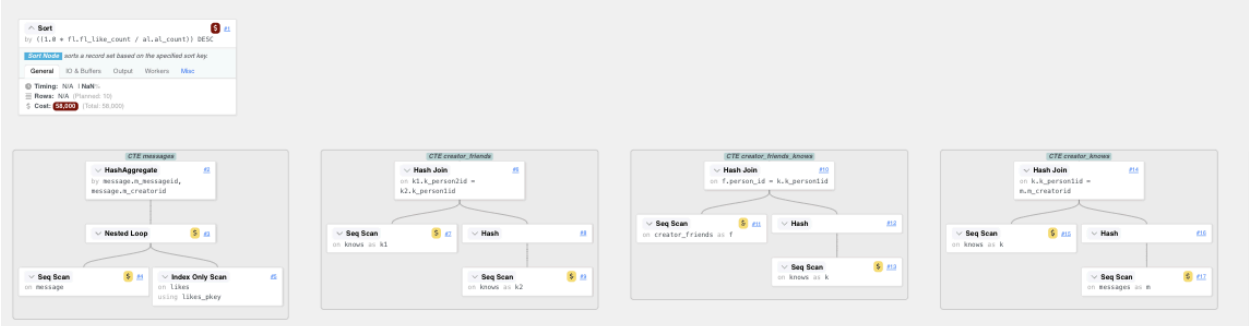
- First observation: for different prompts, the LLM gives completely different execution plans

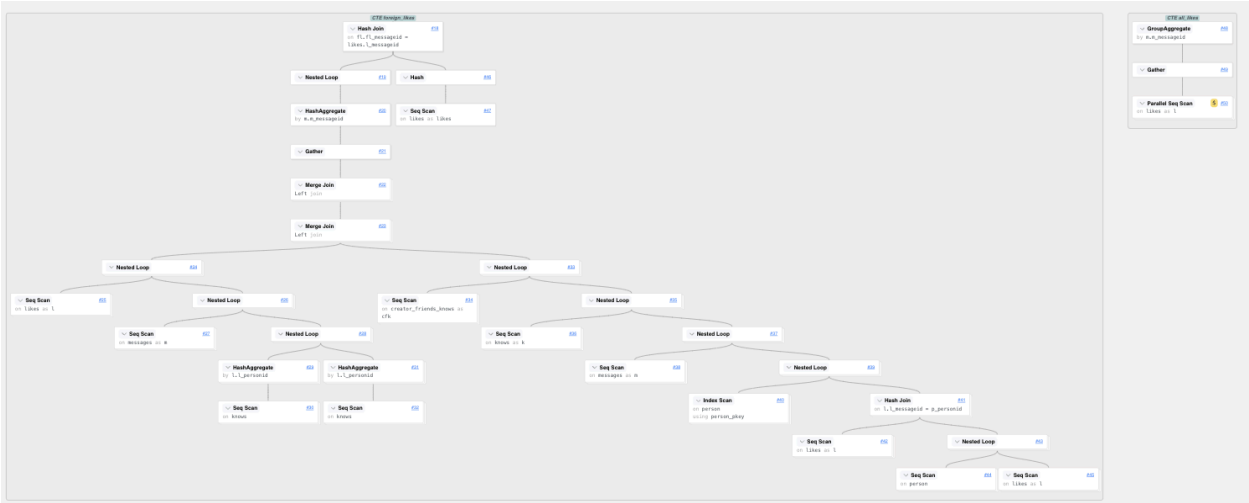
We tried different prompting techniques like giving examples of the result, formulating a Q&A and explicitly asking for the correct format but were not able to get a correct result.

Correct Execution Plan



Example for a generated execution plan





- Some plans couldn't be displayed because of syntax errors
- Other plans contain multiple disconnected trees
- Looks more like hallucination than actual reasoning

In addition, we asked the LLM to explain the real query plan in natural language. It successfully came up with a 120 word description that could recreate parts of the original task description from the query plan.

Our interaction with ChatGPT:
https://chatgpt.com/share/673defe7-2280-8004-bba1-06e44109cc11