INF-2700 Mandatory Assignment Nr. 3

Deadline: Monday, 6 November 2023 23.59

06.07.2023

Now, you are going to extend the features of the `db2700` DBMS with a *natural join* operation.

If you like, you may start with the provided base program without your implementation in Assignment 2. Before you start, make a `git pull` so that you get the lastest corrections in the base program.

Task 1. Implementing natural join

Implement the *natural join* operation with *nested-loop join* (Section 15.5.1 of our textbook) and *block nested-loop join* (Section 15.5.2 of our textbook). We assume that the two join tables only have one common attribute and it has the *int* type.

Task 2. Performance

Run your programs with relatively large tables (e.g. thousand times of the available buffer pages). Profile the runs and compare the performance of the two algorithms.

Task 3. Think out of the box

This task is only on paper. No programming is required.

Now suppose that both tables in the join are stored in B^+-tree organized files where the search keys happen to be the common attribute of the tables. Could you suggest a join algorithm that makes use of this file organization? Compare your suggested algorithm with the block nested-loop join algorithm.

Hand-in

Commit your final solution in the master branch and push it to the INF-2700 git server `inf2700.cs.uit.no`.

Do *not* include the data files for database tables and the executables. These should be able to be re-generated with your code. (That is, run `make cleanall` before you add files for git commits.)

In addition to the source code, you must hand in a report `report-assignment3.pdf` that includes

- a description of your design and implementation,
- instructions on how to run your program and experiments,
- performance of your join algorithms,
- your findings about the performance,
- your design and discussions on the B^+-tree organized files.

Place your report in the `assignments-2-3/docs/` directory.

Enjoy coding and good luck!