October 20, 2017
Arnab Chakrabarti, Rihan Hai

Implementation of Databases (WS 17/18)

Exercise 2

Due until November 02, 2017, 10am.
Please submit your solution *in a single PDF file* before the deadline to the L^{2}P system!
Please submit solutions in groups of three students.

Exercise 2.1 (Query Languages) (10 pts)

1. What does "relational completeness" mean? Show that SQL is relationally complete by enumerating SQL constructs corresponding to selection, projection, cartesian product, union, and difference. (2 pt.)
2. Give two examples of SQL constructs/semantics not expressible in relational algebra (RA). (2 pt.)
3. Figure 1 shows the flow of a query through a DBMS, in which different forms are used to represent a query at different stages. Fill in the three blanks with the corresponding query languages (i.e., SQL, RC, RA). (3 pt.)

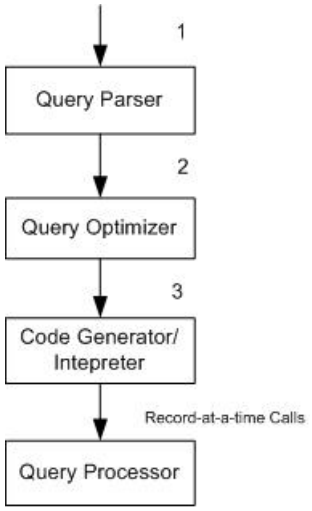

Figure 1: Query Flow through DBMS

4. List the basic operators of RA. Which of these operators are unary operators? Which of them are binary operators? Which of these operators stand for set operations? (3 pt.)

Exercise 2.2 (RA, TRC, DRC) (15 pts)

The following relations are given:

- `lives(pname,city,street)` which contains for every person the location(s) he lives,
- `works(pname,cname,salary)` which contains for every person the name of the company that the person works for, as well as his salary,
- `located(cname,city)` which contains the locations for every company (i.e. a company can be located in more than one city),
- `boss(pname,mname)`, which contains the persons that are supervised by a manager.

Formulate the following queries as expressions in **relational algebra**, **tuple relational calculus**, and **domain relational calculus**:

1. Find the names of all persons who are working in the same company as their boss and get in this company a higher salary than their boss.
2. Find the names of all persons, who work for at least two different companies (≥ 2).
3. Find the names of the persons with the highest salary (Note: There might be several persons with the same salary).
4. Find the names of all companies that are located in cities in which 'IBM' is **not** located.

Exercise 2.3 (Sorting) **(5 pts)**

Suppose you have a file of 15,000 pages and eight buffer pages and you are sorting it using general (external) merge-sort. Please answer the following questions:

1. How many runs will you produce? Remark: When a file is sorted, in intermediate steps subfiles are created. Each sorted subfile is called a *run*. See also slide 35 in Chapter 2.
2. How many passes will it take to sort the file completely?
3. How many buffer pages do you need at least to sort the file in two passes?
4. How many runs and passes would a Two-Way-Sort algorithm take?