CMPSCI 445 — Homework 1

5 Problems, 100 Points

Due September 22, at the beginning of class.
Typed solutions preferred. If hand-written, solutions must be **legible**.

1. (18 pts) **Operator result sizes**

   Given two relations R1 and R2, where R1 contains N1 tuples, R2 contains N2 tuples, and N2 > N1 > 0, give the minimum and maximum possible sizes (in tuples) for the resulting relation produced by each of the following relational algebra expressions. Also, in each case, state any assumptions about the input schemas for R1 and R2 needed to make the expression meaningful:

   (a) R1 ∩ R2
   (b) R1 ∪ R2
   (c) R1 − R2
   (d) R1 × R2
   (e) $\sigma_{a=5}(R1)$
   (f) $\pi_a(R1)$

2. (10 pts) **Translating relational algebra expressions**

   Consider the following schema:

   - Suppliers(sid: integer, sname: string, address: string)
   - Parts(pid: integer, pname: string, color: string)
   - Catalog(sid: integer, pid: integer, cost: real)

   The domain of each field is listed after the field name. Naturally, the Suppliers and Parts relations represent supplier entities and part entities and their attributes. The Catalog relation lists the prices charged for parts by suppliers. Describe in words the output of these relational algebra expressions:

   (a) $\pi_{sname}(\pi_{sid}((\sigma_{color=green}Parts) \bowtie (\sigma_{cost<100}Catalog)) \bowtie Suppliers)$
   (b) $(\pi_{sname}((\sigma_{color=red}Parts) \bowtie (\sigma_{cost<100}Catalog) \bowtie Suppliers)) \cap (\pi_{sname}((\sigma_{color=green}Parts) \bowtie (\sigma_{cost<100}Catalog) \bowtie Suppliers))$

3. (12 pts) **Monotonicity**

   A query or operator on relations is said to be *monotone* if whenever we add a tuple to one of the input relations, the result contains all the tuples that it contained before adding the tuple, plus perhaps more tuples. That is, there is no way to remove tuples from the output by adding tuples to the input.

   For each relational algebra operator below, state whether it is monotone. If an operator is not monotone, provide an example that demonstrates this fact.

   (a) $\cup$

   (b) $\cap$

   (c) $\times$

   (d) σ

   (e) π

   (f) $-$

4. (6 pts) **Relational algebra to SQL**

   Given relational tables with schemas $R(A, B), S(A, C), T(C, D)$, express the following relational algebra query in SQL.

$$\Pi_{R.A,R.B,S.C}(\sigma_{R.A<10,T.D=11}(R \bowtie S \bowtie T)))$$

   (Note: There is a tricky part to this question—remember that the relational algebra and SQL handle duplicates differently.)

5. (54 pts) **Practical SQL exercises**

You will use Postgres to execute queries on a sample dataset consisting of the 3 tables described again below:

- Suppliers(sid: integer, sname: varchar(50), address: varchar(60))
- Parts(pid: integer, pname: varchar(50), color: varchar(20))
- Catalog(sid: integer, pid: integer, cost: real)

First load the database into the Postgres server, which is running in the edlab. Each table is in a separate text file (one tuple per line, with attributes separated by commas). Instructions for loading the data, links to Postgres documentation, and other resources are available on the course web page under "Assignments".

Write SQL expressions for each of the following queries and execute them. Please submit for each query below (i) a valid SQL query, and (ii) the output of the query on the sample data.

(1) Find the pnames of parts for which there is some supplier.

(2) Find the snames of suppliers who supply every part.

(3) Find the snames of suppliers who supply every red part.

(4) Find the sids of suppliers who charge more for some part than the average cost of that part (averaged over all the suppliers who supply that part).

(5) For each part, find the sname of the supplier who charges the most for that part.

(6) Find the sids of suppliers who supply only red parts.

(7) Find the sids of suppliers who supply a red part and a green part.

(8) Find the sids of suppliers who supply a red part or a green part.

(9) For every supplier, print the name of the supplier and the total number of parts that she supplies.

(10) EXTRA CREDIT: For every supplier that only supplies green parts, print the name of the supplier and the total number of parts that she supplies.