York University **Midterm Exam / Fall 2007**
Department of Computer Science and Engineering **CSE3421**

- This is a closed book, **75 minutes** test.
- **No questions are allowed during the test. If in doubt, write down your doubts and assumptions and proceed with your answer.**

LAST NAME	**SOLUTIONS**
FIRST NAME	**SOLUTIONS**
YORK ID#	**SOLUTIONS**
CS LOGIN	**SOLUTIONS**

Exercise 1 [15 points]
Consider the following Supplier-Parts-Catalog schema:

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

State what the following 3 queries compute:

Query 1:

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

Query 2:

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

Query 3:

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

Answer :
Query 1: This Relational Algebra statement does not return anything because of the sequence of projection operators. Once the sid is projected, it is the only field in the set. Therefore, projecting on *sname* will not return anything.
Query 2: Find the Supplier names of the suppliers who supply a red part that costs less than 100 dollars and a green part that costs less than 100 dollars.
Query 3: Find the Supplier names of the suppliers who supply a red part that costs less than 100 dollars and a green part that costs less than 100 dollars.

Exercise 2 [20 points]
Consider the following Supplier-Parts-Catalog schema:

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

Write the following queries in relational algebra and in SQL

Query 1:
Find the sids of suppliers who supply every red part or supply every green part.

Query 2:
Find the pids of parts supplied by every supplier at less than $200. (If any supplier either does not supply the part or charges more than $200 for it, the part is not selected.)

Answer:

Query 1:
RA:

$$\rho(R1, ((\pi_{sid,pid}Catalog)/(\pi_{pid}\sigma_{color='red'}Parts)))$$
$$\rho(R2, ((\pi_{sid,pid}Catalog)/(\pi_{pid}\sigma_{color='green'}Parts)))$$
$$R1 \cup R2$$

SQL:

```
SELECT C.sid
FROM   Catalog C
WHERE  (NOT EXISTS (SELECT P.pid
                    FROM   Parts P
                    WHERE  P.color = 'red' AND
                    (NOT EXISTS (SELECT C1.sid
                                 FROM   Catalog C1
                                 WHERE  C1.sid = C.sid AND
                                        C1.pid = P.pid))))
       OR ( NOT EXISTS (SELECT P1.pid
                        FROM   Parts P1
                        WHERE  P1.color = 'green' AND
                        (NOT EXISTS (SELECT C2.sid
                                     FROM   Catalog C2
                                     WHERE  C2.sid = C.sid AND
                                            C2.pid = P1.pid))))
```

Query 2:
RA:

$$\pi_{sid}\left(Suppliers\right) \rightarrow R1$$
$$\pi_{sid,pid}\left(\sigma_{\cos t \leq 200}\left(Cata\log\right)\right) \rightarrow R2$$
$$R2 / R1 \rightarrow answer$$

SQL:
Similar to above

Exercise 3 [10 points]

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. (assume that the schemas for R1 and R2 are appropriate for performing the related operation).

(1) R1/R2
(2) R2/R1

Answer:

Expression	Assumption	Min	Max
$R1/R2$	The set of attributes of $R2$ is a subset of the set of attributes of $R1$	0	0
$R2/R1$	The set of attributes of $R1$ is a subset of the set of attributes of $R2$	0	$\lfloor$N2 / N1$\rfloor$

Exercise 4 [15 points]
Consider the following relations:

- PARENT: a tuple (p, c) in PARENT means that p is parent of c (i.e., c is child of p).
- BROTHER: a tuple (b, x) in BROTHER means that b is brother of x (x may be male or female).
- SISTER: a tuple (s, f) in SISTER means that s is sister of f (f may be male or female).

Use SQL to produce a relation FIRST_COUSIN with all the tuples (w, z) such that w and z are first cousins, based on the information given from the above 3 tables.

Answer:

```
SELECT    P1.c2, P2.c2
FROM      PARENT P1, PARENT P2, BROTHER B, SISTER S
WHERE     (P1.c1 = B.c1 and B.c2 = P2.c1) OR
          (P1.c1 = S.c1 and S.c2 = P2.c1)
```

Note: c1, c2 are column 1 and column 2 of the corresponding relation.
Note: the idea is:
A tuple (c, d) is in FIRST_COUSIN, if: a tuple (p, c) is in PARENT and (((p, x) is in BROTHER and (x, d) is in PARENT) or ((p, x) is in SISTER and (x, d) is in PARENT))).