UCLA
Computer Science Department
Fall 2014

Instructor: J. Cho

Student Name and ID: ______________________________

CS143 Midterm: Closed Book, 110 minutes

(IMPORTANT PLEASE READ **):**

- There are 5 problems on 9 pages for a total of 100 points. *You should look through the entire exam before getting started, in order to plan your strategy.*
- *Simplicity and clarity of your solutions will count.* You may get as few as 0 point for a problem if your solution is far more complicated than necessary, or if we cannot understand your solution.
- Unless otherwise indicated, assume the SQL92 standard and the notation that we learned in the class.
- If you need to make any assumption to solve a question, please write down your assumptions.
- To get partial credits, you may want to write down how you arrived at your answer step by step.
- You may use one-page double-sided cheat-sheet during exam. You are also allowed to use a calculator.
- Please, write your answers neatly. Attach extra pages as needed. Write your name and ID on the extra pages.

Problem	**Score**	
1	20	
2	30	
3	20	
4	10	
5	20	
Total	100	

Problem 1: 20 points

In this problem, you need to come up with equivalent queries if possible. In providing your answer, pay attention to the possibility of duplicates. In all problems in this exam, keys are underlined.

1. Consider the following SQL query on the table `R(A, B, C)`. You may assume that there are no NULLs.

   ```
   SELECT R1.A
   FROM   R R1, R R2
   WHERE  R1.A=R2.A AND R1.B='UCLA' AND R2.C='blue';
   ```

   Is it possible to write an equivalent relational algebra expression? If yes, write such an expression succinctly. If no, briefly explain why. (10 points)

   ANSWER:
   ```
   No.  When we project on A, there can be multiple tuples that satisfy the
   condition with the same A values.  A relational algebra expression will
   always remove such duplicates while the above SQL query will not.
   ```

2. Consider the following relational algebra expression on the relation $R(\underline{A})$:

   $$R - \Pi_{R.A}(\sigma_{R.A<S.A}(R \times \rho_S(R)))$$

   Again, you may assume that there are no NULLs. Is it possible to write an equivalent SQL query without using subqueries? If yes, write such an SQL query succinctly. If no, briefly explain why. (10 points)

   ANSWER:
   ```
   SELECT MAX(A) FROM R;
   ```

Problem 2: 30 points

1. Consider the following two relational algebra expressions Q_1 and Q_2 over the relation $R(\underline{A}, \underline{B})$.

$$\begin{aligned} Q_1 &: \quad \Pi_A(\sigma_{R.A=5}(R \bowtie \rho_{S(A,C)}(R))) \\ Q_2 &: \quad \sigma_{A=5}(\Pi_A(R)) \end{aligned}$$

   Are Q_1 and Q_2 equivalent? If yes, write YES below. If no, write NO and provide a small instance on which the above expressions return different answers. (10 points)

   ANSWER:
   `YES`

2. Consider the following two relational algebra expressions Q_1 and Q_2 on the relation $R(\underline{A}, \underline{B})$:

$$\begin{aligned} Q_1 &: \quad \Pi_A(\sigma_{R.A=5 \wedge B \neq C}(R \bowtie \rho_{S(A,C)}(R))) \\ Q_2 &: \quad \sigma_{A=5}(\Pi_A(R)) \end{aligned}$$

   Are Q_1 and Q_2 equivalent? If yes, write YES below. If no, write NO and provide a small instance on which the above expressions return different answers. (10 points)

   ANSWER:
   `No.  R=(5,3) returns empty set for` Q_1 `but returns {5} for` Q_2`.`

3. Consider a SQL table `R(K,V)`. There are no NULL values. Consider the following three queries:

```
Q1: SELECT V FROM R
    WHERE V <= ALL (SELECT V FROM R)

Q2: SELECT DISTINCT V FROM R R1
    WHERE V < ALL (SELECT V FROM R R2 WHERE R2.K <> R1.K)

Q3: SELECT MIN(V) FROM R
```

Circle one of the following statements.

(a) Q1, Q2 and Q3 are all equivalent.

(b) Q1 and Q2 are equivalent; Q3 may produce a different result on some databases.

(c) Q1 and Q3 are equivalent; Q2 may produce a different result on some databases.

(d) Q2 and Q3 are equivalent; Q1 may produce a different result on some databases.

(e) Q1, Q2 and Q3 may all produce different answers on some databases.

If you choose any answer other than 1, show the *simplest* single instance of `R` that you can come up with that demonstrates the nonequivalence(s). In this case, show the result of all three queries on your instance of `R`. (10 points)

ANSWER:

`(e). When R = {(1,1), (2,1)}: Q1 returns {1, 1}. Q2 returns {}. Q3 returns {1}.`

Problem 3: 20 points

Consider the following database of teachers and students:

`Student(sname, saddr, pname),` `Teacher(tname, taddr)`

Here, `sname` is the name of a student, `pname` is the name of the student's parent and `tname` is the name of a teacher. Similarly, `saddr` and `taddr` are the addresses of the student and and the teacher, respectively. We assume that every person's name is unique. A student may have a teacher as his/her parent, in which case the `Student.pname` appears in `Teacher.tname`. Assume a reasonable type for each attribute in the tables.

Now we want to enforce the following constraint on this database: *If a student has a teacher as his/her parent, the student and the teacher should have the same address.*

1. What database modifications can potentially violate the constraint? In the following table write YES for the operations that can potentially violate the constraint. For example, if the deletion of a tuple from the `Student` table may violate the constraint, write YES in the cell corresponding to the `Student` row and the `DELETE` column. (10 points)

	INSERT	DELETE	UPDATE
`Student`			
`Teacher`			

ANSWER:
`Student: YES, NO, YES, Teacher: YES, NO, YES`

2. Is it possible to write one (or more) tuple-based CHECK constraints and reject any modifications that will violate the constraint? If yes, write the CHECK constraint(s) and indicate which table(s) the constraint(s) is/are associated with. If not, briefly explain why. Assume the general SQL92 CHECK constraint, not the limited CHECK constraint supported by MySQL. (10 points)

   ANSWER:

   Yes. To Student table, add CHECK(pname NOT IN (SELECT tname From Teacher) OR saddr = (SELECT taddr FROM Teacher WHERE tname = pname)). To Teacher table, add CHECK(tname NOT IN (SELECT pname From Student) OR taddr = ALL (SELECT saddr FROM Student WHERE pname = tname))

Problem 4: 10 points

A is the owner of the table `R`. Users of our database system have executed the following sequence of commands:

1. By User A: `GRANT SELECT ON R TO` B`,` E `WITH GRANT OPTION;`
2. By User B: `GRANT SELECT ON R TO` C `WITH GRANT OPTION;`
3. By User C: `GRANT SELECT ON R TO` D `WITH GRANT OPTION;`
4. By User E: `GRANT SELECT ON R TO` C`;`
5. By User A: `REVOKE GRANT OPTION FOR SELECT ON R FROM` B `CASCADE;`

In the following table, indicate whether each user has the `SELECT` privilege on `R` with/without `GRANT OPTION`. If the user has the corresponding privilege write YES in the corresponding entry. (10 points)

User	`SELECT ON R?`	`GRANT OPTION?`
A		
B		
C		
D		
E		

ANSWER:

`A: YES, YES. B: YES, NO. C: YES, NO. D: NO, NO. E: YES, YES.`

Problem 5: 20 points

Consider a B+Tree where each node can have at most 3 keys and 4 pointers (i.e., $n = 4$). Suppose the tree initially has a root node with two children leaf nodes, which contain keys $\{1, 3, 5\}$ and $\{7, 9, 11\}$ respectively. The root node has a single key 7. *When a node is split after insertion, assume that we place one more key on the left node than the right node.*

1. After inserting 0 and 6 to the tree, we inserted two more distinct keys from the set $\{2, 4, 8, 10, 12\}$ in some order, which caused the tree to grow by one level. Which two keys were added to the tree? Write the two keys separated by commas together with the final structure of the B+tree after the two keys are inserted. If there are multiple solutions, write any one solution. (10 points)

   ANSWER:
   `Sequence of 2 keys:  4 and one of {8, 10, 12} (as an example, numbers 4,8 is a valid solution)`

2. Suppose we are back to the original tree with two leaf nodes having keys $\{1, 3, 5\}$ and $\{7, 9, 11\}$. Now X distinct keys from the set $\{0, 2, 4, 6, 8, 10, 12\}$ are added to the tree in some order, and the tree does not grow a level. What is the maximum possible value of X? (10 points)

   ANSWER:
   `Maximum value of X: 6.  For example, insert {0, 6, 2, 10, 8, 12} in this sequence.`

This page is intentionally left empty. You can use this page as scratch paper.