CS3550 (DBMS-I)
Segment 3-4
Final Exam
Sunday, Nov 15, 2020
Time Limit: -

Name (Print): ______________________

This exam contains 8 pages (including this cover page) and 10 problems. Check to see if any pages are missing. Enter all requested information on the top of this page, and put your initials on the top of every page, in case the pages become separated.

You may *not* use your books, notes, or any calculator on this exam.

You are required to show your work on each problem on this exam. The following rules apply:

- You have to *return* the question paper along with the answers written in the empty spaces provided. Please use the extra sheets provided to do the rough work and *not* on the question paper.
- Make sure that the answers for the subjective type and the reasons for the objective type are *clearly* written. Note that we value *precision* and *conciseness* in the answers. If you need to make any additional assumptions, write them clearly.
- The marks for each question and each sub-question are given.

Do not write in the table to the right.

Problem	Points	Score
1	10	
2	10	
3	10	
4	10	
5	10	
6	10	
7	10	
8	10	
9	10	
10	10	
Total:	100	

1. (10 points) Consider the two relation schemas for **EMPLOYEE** and **DEPARTMENT** as follows:

Relation E (for EMPLOYEE)

```
enr  ename   dept
---  -----   ----
1    Bill     A
2    Sarah    C
3    John     A
```

Relation D (for DEPARTMENT)

```
dnr   dname
---   ----------
A     Marketing
B     Sales
C     Legal
```

Given a SQL statement as follows:

*select * from E, D where dept = dnr*

(a) (6 points) Write the *Relational algebra expression* for the given SQL (without using a natural or inner join) and

(b) (4 points) the resulting tuples.

2. (10 points) Given two relations R_1 and R_2, where R_1 contains N_1 tuples, R_2 contains N_2 tuples, and $N_2 > N_1 > 0$, give both the **minimum** and **maximum** possible sizes (in tuples) for the resulting relation produced by each of the following relational algebra expressions. In each case, state any **assumptions or necessary conditions** about the schemas for R_1 and R_2 needed in order to make the expression meaningful. Full marks will only be awarded for correct minimum, maximum and assumption/condition necessary. No partial credits will be given.

   (a) (2 points) $R_1 \cup R_2$

   (b) (2 points) $R_1 \cap R_2$

   (c) (2 points) $R_1 - R_2$

   (d) (2 points) $R_1 \times R_2$

   (e) (2 points) $\sigma_{a=5}(R_1)$

3. (10 points) Given a keyword like *TOP (k)* that returns the top-k tuples (e.g.: SELECT TOP(K) $\langle A_1, A_2, ..., A_n\rangle$), write a SQL query to find the 10-th highest employee salary from an **EMPLOYEE** table. Explain your answer.

   Note: You may assume that there are at least 10 records in the **EMPLOYEE** table.

4. (10 points) Given a table **TBL** with a field **Nmbr** that has rows with the following values:

   $1, 0, 0, 1, 1, 1, 1, 1, 0, 0, 1, 0, 1, 0, 1, 0, 1$

   Write a query to add 2 where Nmbr is 0 and add 3 where Nmbr is 1.

5. (10 points) Consider the following schema and provide the *relational algebra expression* :

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

The primary keys of the table are outlined in bold font.

Find the Supplier ids of the suppliers who supply both a red part and a green part where each costs less than 100 dollars.

6. (10 points) For the following SQL query provide the equivalent relational algebra expression. Note that relational algebra has no *MAX* operator!

```
SELECT E.eid
FROM Employees E
WHERE E.salary = ( SELECT MAX(E2.salary)
                   FROM   Employees E2 )
```

7. (10 points) As a continuation to problem 6, for the following SQL query provide the equivalent relational algebra expression. Note that relational algebra has no *MAX* operator!

```
SELECT E.eid
FROM Employees E
WHERE E.salary = (SELECT MAX(E2.salary)
                  FROM Employees E2
                  WHERE E2.salary != (SELECT MAX(E3.salary)
                                      FROM Employees E3 ))
```

8. (10 points) Consider the schema from problem 5 and provide the *SQL* for :

   For every supplier that supplies both a green part and a red part, print the name and price of the most expensive part that he/she supplies.

9. (10 points) Given the following tables:

```
SELECT * FROM users;

user_id  username
-------  --------
1        John Doe
2        Jane Don
3        Alice Jones
4        Lisa Romero

SELECT * FROM training_details;

user_training_id  user_id  training_id  training_date
----------------  -------  -----------  -------------
1                 1        1            "2015-08-02"
2                 2        1            "2015-08-03"
3                 3        2            "2015-08-02"
4                 4        2            "2015-08-04"
5                 2        2            "2015-08-03"
6                 1        1            "2015-08-02"
7                 3        2            "2015-08-04"
8                 4        3            "2015-08-03"
9                 1        4            "2015-08-03"
10                3        1            "2015-08-02"
11                4        2            "2015-08-04"
12                3        2            "2015-08-02"
13                1        1            "2015-08-02"
14                4        3            "2015-08-03"
```

Write a query to to get the list of users who took a training lesson (identified by training_id) more than once in the same day, grouped by user and training lesson, each ordered from the most recent lesson date to the oldest date.

10. (10 points) Given the following table named **A**

```
  x
------
   2
  -2
   4
  -4
  -3
   0
   2
```

Write a single query to return a single tuple with two attributes: (i) the sum of all positive values of x and (ii) the sum of all negative values of x.

Good luck!