Midterm 2 - 5:30

CS386D Database Systems
Date: 4/23/2020
Version: 1.0

Name:

There are 4 sections totaling 200 points. Point weighting is in parenthesis. Place your answers on the question sheet OR if this is not currently feasible for you, a) begin each of the 4 sections on a new page, being careful to clearly label the page with the section number, b) on the first page of the exam, be sure to write your name. You have one hour and fifteen minutes. Good luck. D.M.

I. **Definitions** – Define the following terms (60)

A) "System Catalog"

B) "View"

C) "Materialized View"

D) "Join Selectivity" (preferred answer is expressed succinctly using mathematical notation)

E) Define "logical query plan" and "physical query plan", being careful to explain the differences.

II. **Short Answer**(60)

A) (20) i)What advantages, if any, are gained by restricting query optimization to right or left deep query plans?

ii) What are the disadvantages, if any?

B) (20) Given a relation R(a,b,c), and its size in blocks B(R) = 10,000 and column cardinalities V(R,a) = 10,000, V(R,b) = 2000, V(R,c) = 11 and the number of memory buffers, B(M) = 1000. Per the text book's linear cost model (omit the final write),

i. How long will a two phase sort take when sorting on column a?

ii. How long will a two phase sort take when sorting on column b?

iii. Can we use two phase sort when sorting on column c? (yes/no and explain) (Hint: Consider each step and/or phase of the algorithm.)

C) (20) In some RDBMSs foreign-key constraints are restricted to reference only the primary key of the parent table.

i. Explain the advantage(s) of such a restriction, if any.

ii. Explain the disadvantage(s) of such a restriction, if any.

III. Query Optimization (50)

Consider the following SQL create table commands, including the consistency constraints, the query and values from the data catalog.

T(R) = 1000	V(R,b) = 1000	V(R,c) = 50	
T(S) = 500	V(S,a) = 50	V(S,b) = 50	
T(W) = 1500	V(W,c) = 100	V(W,d) = 50	V(W,e) = 15

```
CREATE TABLE R (b float NOT NULL, c int);
CREATE TABLE S (a float, b float NOT NULL REFERENCES R)
CREATE TABLE W (c int, d float, e varchar(64));

SELECT *
FROM R, S, W
WHERE R.b = S.b and S.a = W.d and W.c = R.c and W.e = “student”;
```

a) Draw the query graph. (10)

b) Illustrate a right deep logical query plan with the leaves in alphabetical order.(10)

c) Annotate your illustration with the detail (predicate for) each join operator, the size of the intermediate results and the size of the final result.(30)

IV. **Semantics of foreign key constraints.** (Multiple choice, choose the best answer) (30)

1) Consider this snippet of code declaring a table in SQL that contains the information of all UT student employees, including your TAs:

```
Create Table UT_StudentEmployees(
        Name char(200),
        SSN char(11) UNIQUE,
        EID char(7) Primary Key,
        Phone Integer);
```

Consider two additional tables, UT_Employees and UT_Students. Assume both tables also contain (at minimum) name and EID fields and that EID is primary key in both. Which of the options below BEST describes the most appropriate foreign key constraints for this database?

A. UT_Employees.EID references UT_StudentEmployees(EID);
   UT_Students.EID references UT_StudentEmployees(EID)

B. UT_Employees.name references UT_StudentEmployees(name);
   UT_Students.name references UT_ Employees(name)

C. UT_StudentEmployees.EID references UT_Students (EID);
   UT_StudentEmployees.EID references UT_ Employees(EID)

D. UT_StudentEmployees.name references UT_Students(name);
   UT_StudentEmployees.name references UT_ Employees(name)

E. Foreign key constraints are not suitable for this problem

Consider the following SQL fragment
"(T1.attribute1, T1.attribute2) references T2(attribute1,attribute2) ON UPDATE SET NULL"

2) Suppose a tuple in T2 is updated from (a,1) to (a,5). The modification succeeds. What happens next?

A. Nothing.
B. All occurrences of (a,1) in T1 would be set to (NULL,NULL)
C. All occurrences of (a,1) in T1 would be set to (a, NULL)
D. All occurrences of (a,5) in T1 would be set to (a, NULL)
E. All occurrences of (a,5) in T1 would be set to (NULL,NULL)

3) Now I change (b,3) in T1 to (m,3). The modification succeeds. What happens next?

A. Nothing
B. All occurrences of (b,3) in T2 would be set to (NULL,NULL)
C. All occurrences of (m,3) in T2 would be set to (NULL, NULL)
D. All occurrences of (b,3) in T2 would be set to (NULL,3)
E. All occurrences of (m,3) in T2 would be set to (NULL,3)