CSE 322 Final

Olcay Taner YILDIZ

I. QUESTION (16 POINTS)

Isik University database system contains information about instructors (identified by id) and courses (identified by courseid like CSE322.01). Instructors teach courses; each of the following situations concerns the Teaches relationship. For each situation, draw an ER diagram that describes it (assuming no further constraints hold).

- Instructors can teach the same course in several semesters, and each offering must be recorded in Campus-Online.
- Instructors can teach the same course in several semesters, and only the most recent such offering must be recorded in Course-Online.
- Every instructor must teach some course.
- Every instructor teaches exactly one course (no more, no less), and every course must be thought by some instructor.

II. QUESTION (16 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.

- $R1 \cup R2$
- $R1 \cap R2$
- $R1 - R2$
- $\pi_a(R1)$

III. QUESTION (16 POINTS)

Consider the following relations about the suppliers-parts-projects database. The significance of an shipment row is that the specified supplier supplies the specified part to the specified project in the specified quantity.

Suppliers(sid:integer, sname:string, status:integer, city:string)

Parts(pid:integer, pname:string, color:string, weight:integer, city:string)

Projects(prjid:integer, prjname:string, city:string)

Shipment(sid:integer, pid:integer, prjid:integer, quantity:integer)

Write the following queries in SQL.

- Find the names of suppliers who have supplied a project with part 1.
- Find the sids of all suppliers who have supplied red parts to projects but not green parts.
- Find the names of suppliers who have supplied all parts.
- Find the weight of the lightest part for each color.

IV. QUESTION (8 POINTS)

Give Java code for the following operations: Establish a connection to url 'jdbc:mysql:www.cse322.com:3083' with userID 'root' password 'root'. Do not forget exception handling codes (Use try-catch structure).

V. QUESTION (12 POINTS)

At most how many entries can we insert into a B+ tree of height 4 (There are 3 non-leaf levels and 1 leaf level in the tree) of order 2 without increasing the height of the tree?

VI. QUESTION (18 POINTS)

Consider a relation R(a, b, c, d, e) containing 5 million records, where each page of the relation holds 10 records. R is organized as a sorted file with secondary indexes. Assume that attribute a is a candidate key for R, with values lying in the range 0 to 4,999,999 and R is stored in R.a order. For each of the following relational algebra queries, state which of the following approaches is most likely to be the cheapest.

- Access the sorted file for R directly.
- Use a clustered B+ tree index on attribute R.a.
- Using a linear hash index on attribute R.a.
- Use a clustered B+ tree index on attributes (R.a, R.b).
- Use a linear hashed index on attributes (R.a, R.b).
- Use an unclustered B+ tree index on attribute R.b.

The queries are:

1) $\sigma_{a<50,000 \wedge b<50,000}$
2) $\sigma_{a=50,000 \wedge b<50,000}$
3) $\sigma_{a>50,000 \wedge b=50,000}$
4) $\sigma_{a=50,000 \wedge b=50,010}$
5) $\sigma_{a \neq 50,000 \wedge b=50,000}$
6) $\sigma_{a<50,000 \vee b=50,000}$

VII. QUESTION (14 POINTS)

Supppose there are 3 buffer pages available of main memory and we need to sort the following pages in sort-merge join: (6, 12), (8, 4), (6, 2), (3, 7), (6, 9), (8, 8), (10, 7), (2, 6). Show the iterations of external merge sort on those pages.