CSCI 6333/6315 Database Systems

Spring 2020

ASSIGNMENT 4: Data Storage and Querying, Transaction Management

All answers shall be typed using a word processor and some drawing utilities. A cover page shall be prepared with course title, homework number, submitted date and time, and contact info including your email address.

Due: Midnight, Tuesday, April 28, 2020

The total score of this assignment is 110. Each part of the problem has 10 points.

Problem 1. This problem has two parts:

a. Construct a B^+-tree for the following key values:

   2, 3, 5, 7, 11, 17, 19, 23, 29, 31.

   Assume that the number of pointers that will fit in one internal node is 4 and each leaf node can store 3 key values.

b. After the B^+-tree is constructed for Part a, show the final tree after the following operations:

   b.1. Insert 9
   b.2. Insert 10
   b.3. Insert 8
   b.4. Delete 23
   b.5. Delete 19

Problem 2. This problem has two parts:

a. Suppose that we are using extendable hashing on a file that contains records with the following search key values:

   2, 3, 5, 7, 11, 17, 19, 23, 29, 31.

   Show the extendable hash structure for this file if the hash function is $h(x) = x\%8$, and each bucket can hold three records.

b. After Completing Part a, show the extendable hash structures after the following operations:

   b.1. Delete 11
   b.2. Delete 31
   b.3. Insert 1
   b.4. Insert 15

Problem 3. Consider the following SQL query for our University Database:

```
select T.dept_name
from department as T, department as S
where T.budget > S.budget and S.building = "MAG"
```

Write an efficient relational-algebra expression that is equivalent to this query. Justify your choice.

Problem 4. Let the relations $r_1(A,B,C)$ and $r_2(C,D,E)$ have the following properties: r_1 has 20,000 tuples, r_2 has 45,000 tuples, 25 tuples of r_1 fit on one block, and 30 tuples of r_2 fit on one block. Estimate the number of block accesses required, using each of the following join strategies for $r_1 \triangleright\triangleleft r_2$:

a. Nested-loop join
b. Block nested-loop join
c. Merge join
d. Hash join

Problem 5. Show that the following equivalences hold.

c. $E_1 \triangleright\triangleleft_\theta (E_2 - E_3) = (E_1 \triangleright\triangleleft_\theta E_2 - E_1 \triangleright\triangleleft_\theta E_3)$

d. $\sigma_{\theta_1 \wedge \theta_2}(E_1 \triangleright\triangleleft_{\theta_3} E_2) = \sigma_{\theta_1}(E_1 \triangleright\triangleleft_{\theta_3} (\sigma_{\theta_2}(E_2)))$, where θ_2 involves only attributes from E_2.