COMP 3311: Database Management Systems

Lecture 15 Exercises
Query Processing: Join Operation

Exercise 1: Use the following information about the relations to estimate the page I/O cost to compute the query result using the stated join strategies.

Page size: 1000 bytes
buffer size *M*: 100 pages
Each attribute/pointer: 20 bytes
Sailor: 10,000 tuples; 12 tuples/page
Reserves: 40,000 tuples: 16 tuples/page
4 reservations/sailor on average

Sailor(sailorId, sName, rating, age)
Reserves(sailorId, boatId, rDate)

Query: Find the names of sailors who have reservations.

```
select *
from Sailor natural join Reserves;
```

a) block nested-loop join

i. using Sailor as the outer relation

ii. using Reserves as the outer relation

b) indexed nested-loop join with hash index on Reserves.sailorId (assume no overflow)

c) merge join

d) hash join (assume no overflow)

Do not upload this exercise sheet to Canvas.

Name: (1) _____/_____ **Student#:** (1) _____ **Date:** _____
Last/Family (PRINT) Given/First (PRINT)

Name: (2) _____/_____ **Student#:** (2) _____
Last/Family (PRINT) Given/First (PRINT)

NOTE: Only half credit will be given if done without a partner.

COMP 3311: Database Management Systems

Lecture 15 Exercises
Query Processing: Join Operation

Exercise 2: The relations $R_1(A, B, C)$ and $R_2(C, D, E)$ have the following properties:

- R_1 has 20,000 tuples
- 25 tuples of R_1 fit on one page
- R_1 requires 800 pages
- R_2 has 45,000 tuples
- 30 tuples of R_2 fit on one page
- R_2 requires 1500 pages

Assuming that there are 800 buffer pages available for processing a join, estimate the page I/O cost for each of the following join strategies for R_1 **JOIN** R_2.

a) nested-loop join
   i. using R_1 as the outer relation

      Join page I/O cost:

   ii. using R_2 as the outer relation

      Join page I/O cost:

b) block nested-loop join
   i. using R_1 as the outer relation

      Join page I/O cost:

   ii. using R_2 as the outer relation

      Join page I/O cost:

c) merge join (assume that both relations are <u>not</u> sorted initially)
   i. sorting cost

      Page I/O cost to sort R_1:

      Page I/O cost to sort R_2:

      Total page I/O cost to sort:

   ii. merge cost (join phase)

      Total page I/O cost to merge:

      Join page I/O cost:

d) hash join (assume no overflow occurs)

   Join page I/O cost:

You must upload this completed exercise sheet to Canvas by **11 p.m. today.**