COMP 3311: Database Management Systems

Tutorial 9
Query Optimization

Exercise 1: Sailor(sailorId, sName, rating, age) 6,000 tuples   10 tuples/page
      Reserves(sailorId, boatId, rDate)   1,500 tuples   15 tuples/page
      Boat(boatId, bName, color)   3,000 tuples   20 tuples/page

    There is no index on any relation. 15% of boats are red.

a) Estimate the size in tuples of $\pi_{boatId}(\sigma_{color='red'}Boat)$.

b) Estimate the size in tuples of (Sailor Join Reserves Join Boat). Briefly explain the reason(s) for your answer.

c) i. Calculate the result size in tuples of the join order (Sailor Join Boat) Join Reserves

  ii. Calculate the result size in tuples of the join order (Sailor Join Reserves) Join Boat

  iii. Which join order is better and why?

COMP 3311: Database Management Systems

Tutorial 9
Query Optimization

Exercise 2: For the query $\pi_{A,B,C,D}(R \bowtie_{A=C} S)$ assume that the projection routine uses external sorting, eliminates all unwanted attributes during the initial sort pass and removes duplicate tuples on-the-fly during the merge passes. Furthermore, assume the following:

- R is 10 pages; each R tuple is 300 bytes.
- S is 100 pages; each S tuple is 500 bytes.
- The combined size of attributes A, B, C and D is 450 bytes.
- A and B are in R and have combined size 200 bytes; C and D are in S.
- A is a key for R.
- Each S tuple joins with exactly one R tuple.
- The page size is 1024 bytes.
- The buffer size M is 3 pages.
- Only the (optimized) block nested-loop join method is implemented.

What is the query processing page I/O cost?

a) Join Cost (using (optimized) block nested-loop and eliminating unwanted attributes during the join)

Join page I/O cost:

Write join result page I/O cost:

b) Projection Cost (using external sorting)

Sort page I/O cost pass 0:

Merge page I/O cost:

Query processing page I/O cost:

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

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

NOTE: You are highly encouraged to do this exercise with a partner.

Exercise 3: What is the query processing page I/O cost if merge join is used for Exercise 2 instead of block-nested loop using a buffer of 3 pages?

a) <u>Cost to sort R</u>

Sort pass page I/O cost:

Merge pass page I/O cost:

Page I/O cost to sort R:

b) <u>Cost to sort S</u>

Sort pass page I/O cost:

Merge pass page I/O cost:

Page I/O cost to sort s:

c) <u>Cost to join</u>

Join page I/O cost:

<u>Query processing page I/O cost</u>:

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