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?

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.

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 and eliminates all unwanted attributes during the initial sort pass and also removes duplicate tuples on-the-fly during the merge passes. Furthermore, assume the following:

- R is 10 pages and each R tuple is 300 bytes.
- S is 100 pages and 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 main memory buffer size *M* is 3 pages.
- Only the (optimized) block nested-loop join method is implemented.

a) What is the cost of processing the query $\pi_{A,B,C,D}(R \bowtie_{A=C} S)$?

i. join cost (using block nested-loop)

ii. projection cost (using external sorting and eliminating unwanted attributes during initial sort pass)

b) What is the cost of processing the query $\pi_{A,B,C,D}(R \bowtie_{A=C} S)$ if merge join is used instead of block nested-loop join with a main memory buffer of 3 pages and unwanted attributes are eliminated during sorting?

i. cost to sort R

ii. cost to sort S

iii. cost to join

Total cost:

You must upload this completed exercise sheet to Canvas by **11:00 p.m. today.**