COMP 3311: Database Management Systems

Lecture 15 Exercises
Query Processing: Introduction

Exercise 1: For the relation Apartment(code, year, price, area), we want to process the query

Find the apartments in a given set of areas.

There are 1,000,000 apartments uniformly distributed over 100 areas. 10 Apartment records fit in one page. The records are ordered on price. There is a hash index on code and a dense, non-clustering B$^+$-tree index on area. Each B$^+$-tree node can fit 20 pointers (19 values). 100 record pointers can fit on a page. Assume that each non-leaf B$^+$-tree node holds the _minimum_ *number of values*.

Is it advantageous, in terms of page I/Os, to use the B$^+$-tree on area to answer the query if there are 5 area values?

Cost to scan file:

Cost to use B$^+$-tree index:

Exercise 2: How many passes are needed to sort a file that has 108 pages, using only 5 pages of main memory?

Number of passes

Pass 0:

Pass 1:

Pass 2:

Pass 3:

What is the I/O cost?

COMP 3311: Database Management Systems

Lecture 15 Exercises
Query Processing: Introduction

Exercise 3: Assume that for the relation Sailor(<u>sailorId</u>, sName, rating, age) each rcord is 50 bytes, a page can hold 80 records, there are 500 pages of records and the file is unordered. We want to process the following two queries: i. $\sigma_{sailorId<50000}(\text{Sailor})$ ii. $\sigma_{sailorId=50,000}(\text{Sailor})$

Estimate the number of pages retrieved for each query given the information in a) and b) below.

a) Assume that we have a B$^+$-tree index on the search key sailorId with height 5, 50 index pages, the lowest index search-key value is 1 and highest index search-key value is 100,000.

   i. $\sigma_{sailorId<50000}(\text{Sailor})$

   ii. $\sigma_{sailorId=50,000}(\text{Sailor})$

b) Assume that we have a hash index on the search key sailorId with 50 index pages, no overflow, the lowest index search-key value is 1 and highest index search-key value is100,000.

   i. $\sigma_{sailorId<50000}(\text{Sailor})$

   ii. $\sigma_{sailorId=50,000}(\text{Sailor})$

Exercise 4: How many passes are needed to sort a file that has 600,000 pages, using 150 pages of main memory?

What is the I/O cost?

Exercise 5: Let B be the size of a file in pages. How many memory pages M do you need to sort the file in 2 passes?