Database Systems, CSCI 4380-01
Homework # 7
Due Thursday April 21, 2011 at 2 pm

Answer the following questions. Turn in a single text or PDF file in the assignment drop box.

Question 1 [20 points]. Suppose you are given a B-tree where each leaf node can address at most 5 and at least 2 tuples. Each internal node can have at most 5 and at least 2 pointers (i.e. between 1-4 key values).

(a) Given this B-tree structure and the given B-tree. Insert the following points: 107, 105, 103, 137, 140. Draw the resulting tree.

(b) Given this B-tree structure and the given B-tree. Delete the following points: 5, 8, 10, 24, 26, 31. Draw the resulting tree.

Note. The sibling pointers at the leaf level are not shown for simplicity.

Question 2 [15 points]. Suppose you are given a B-tree on $R(A, B)$ of height 3 where each node (leaf or internal) contains about 500 entries approximately (key value, pointer pairs). R has a total of 10 million tuples.

Suppose you are given the following queries:

(a) SELECT * FROM R WHERE A = 10 AND B = 20 AND C = 30

(b) SELECT * FROM R WHERE A >= 1 AND A < 10 AND B = 20

(c) SELECT * FROM R WHERE A = 10 AND B >= 1 AND B < 10

Suppose, the number of tuples satisfying the above conditions are as given below:

Conditions	Number of tuples
A = 10	10,000
B = 20	50,000
C = 30	500
A = 10 and B = 20	2,000
A = 10 AND B = 20 AND C = 30	20
A >= 1 AND A < 10	100,000
A >= 1 AND A < 10 AND B = 20	5,000
B >= 1 AND B < 10	500,000
A = 10 AND B >= 1 AND B < 10	8,000

For each query, assume you are using the B-tree. Write down how many nodes of the B-tree are scanned and how many tuples are read from the relation to answer this query.

Question 3 [10 points]. Suppose you are given a relation R that spans 5,000 pages ($PAGES(R) = 5,000$). What is the cost of sorting this relation if the available memory for the sort is:

(a) $M = 50$ pages

(b) $M = 100$ pages

Show your work.

Question 4 [10 points]. What is the cost of block-nested loop join of R and S where

(a) $PAGES(R) = 5,000$, $PAGES(S) = 500$, $M = 101$.

(b) $PAGES(R) = 5,000$, $PAGES(S) = 500$, $M = 1,001$.

Use the cheapest ordering among R or S (i.e. choosing which relation is outer and which one is the inner relation).

Question 5 [10 points]. Suppose you perform sort based merge join between R and S where $PAGES(R) = 5,000$, $PAGES(S) = 500$, $M = 500$. What is the total cost of the join algorithm (including sort)? Show your work.