CS 351 DATA ORGANIZATION AND MANAGEMENT

Homework 2
Date Given : October 30, 2011
Date Due : November 16, 2011

Important Notes: 1. Please submit the Homework to Room EA 511 on the due date by 5:00 pm –on the same day you can also give it to me in the classroom- (no late submission will be accepted). 2. Answer the question in the order they are given using a standard size paper. 3. Handwritten submissions are accepted; however a word document is preferred and appreciated. 4. Staple all papers and write your name on them.

Q1. Perform heap sort on the following successively entered records with keys;

84, 20, 50, 57, 18, 90

Assume 6-record memory capacity, i.e., the maximum size of the complete binary tree (priority queue) is six nodes. In your work show the input value, tree, array's contents and output value when appropriate.

Q2. Perform replacement selection sort on the following successively entered records with keys;

84, 30, 50, 57, 28, 90, 92, 15, 24, 88, 12, 98

Assume that only 6-record memory capacity is available. In your work show the input value, tree, array's contents and output value when appropriate.

Q3. a. Is it possible to use heap sort without any merge and obtain a sorted file? Explain your answer. b. Answer the same question for replacement selection sort.

Q4. Assume that we have a file of size 2000 MB. Its records are in random order. Also assume that we have 10 MB available for storing the heap structures. (Assume IBM 3380 environment. R= 400 bytes. Our concern is the generation of sorted segments, merging is out of concern.) Show your work.

a. If we sort this file with heap sort how many (s+r) are we going to have?
b. What is the total sorting time excluding (s+r)?
c. Now consider the replacement selection sort and of course assume that we have two disk drives. How much time is needed for sorting (approximation is OK: you may ignore s+r)?

Q5. Consider a credit card company with 4,000,000 customers. Develop a record structure for keeping the information for customers (fixed size records, records can have a size between 200 to 400 bytes, B= 2400 bytes). The company wants to sort the customer data on a computer with a single IBM3380 hard disk structure. Assume that a certain amount of memory is available for sorting and then merging: it can have a size between 10 to 20 MB ([10, 20] MB), you choose the available memory size.

What is the total time (heap sort + merging) for each of the following cases?

a. The time needed to perform the heap sort part
b. 2–way merge time,
c. 4-way merge time,
d. P-way merge time where P is equal to the total number of available sorted segments (additional info: another name for a sorted segment is "run"),
e. Is there a significant difference between 2-way and 4-way merge? Please explain why.