

## **CPSC 4660 – Database Management Systems Fall 2018**

### **Assignment 1 (80 marks) – Due October 5<sup>th</sup>, 4pm (note change!)**

#### **Assignment Submission: Moodle**

- 1) [4 marks]
  - a. Briefly describe one advantage and one disadvantage of using fixed-length records in a heap file
  - b. Briefly describe one advantage and one disadvantage of using variable-length records in a heap file
  
- 2) [20 marks] (modified from p. 595 of the course text) A file STUDENT has  $r=20,000$  records of fixed length. Each record has the following fields: Name (25 bytes), SIN (9 bytes), Address (44 bytes), Phone (10 bytes), BirthDate (8 bytes), MajorDeptCode (4 bytes), ClassCode (4 bytes) and DegreeProgram (4 bytes). An additional byte is used as a deletion marker. The file is stored on a hard disk with the following parameters: 15 single-sided disks, block size  $B=512$  bytes, 20 blocks per track, and 40 tracks per surface, seek time of 4ms, transfer time of 0.1ms a block.
  - a. Calculate the record size  $R$  in bytes.
  - b. Assuming an unspanned record organization, calculate the blocking factor  $bfr$ , the number of blocks  $b$ , and the number of track  $t$  occupied by the file.
  - c. Calculate the average time it takes to find a record by doing a linear search on the file if the file blocks are stored contiguously (both contiguous blocks and tracks). You can assume after the first seek to a new disk that moving the read/write head over an existing disk incurs negligible cost.
  - d. Assume that the file is ordered by SIN. by doing a binary search, calculate the worst-case time it takes to search for a record given its SSN value.

(Note: For c and d, there will probably be some variations to what I provide in the answer key. These will be accepted if some justification is provided as to how you derived the calculation that you did.)

- 3) [20 marks] Construct a B+-tree for the following set of key values. Assume that a node size of 4 pointers (i.e. 3 keys) is used. (If you want to be considered for part marks in the event you make a mistake, you should show some of the intermediate steps).

5,6,8,10,14,22,26,32,34

4) [10 marks] Starting with the B+-tree you created in Question 3, show the resulting B+ -tree after each of the following operations are performed.

- a. Insert 12
- b. Insert 13
- c. Insert 11
- d. Delete 26
- e. Delete 22

5) [20 marks] Construct an extendible Hashing Structure for the following set of key values. Assume that a bucket size of 3 entries, and the hashing function  $h(x) = x \bmod 8$ , are used. (If you want to be considered for part marks in the event of a mistake, you should show some of the intermediate steps).

5,6,8,10,14,22,26,32,34

6) [6 marks] Starting with the extendible hashing structure from Question 5, show the resulting structure after each of the following operations are performed.

- a. Delete 14
- b. Delete 34
- c. Insert 4
- d. Insert 18