

CSE 4701 Fall 2023

Homework 2: Due Oct. 18 (Wed) 2023, midnight, at HuskyCT

Problem 1. (35 pts)

Consider a disk with block size $B = 1024$ bytes. A block pointer is $P = 6$ bytes long, and a record pointer is $P_R = 7$ bytes long. A file has $r = 1,000,000$ EMPLOYEE records of fixed length. Each record has the following fields: NAME (25 bytes), SSN (10 bytes), DEPARTMENTCODE (9 bytes), ADDRESS (35 bytes), BIRTHDATE (8 bytes), JOBCODE (2 bytes), and SALARY (3 bytes).

- a. Calculate the record size R in bytes.
- b. Calculate the blocking factor bfr and the number of file blocks b , assuming an unspanned organization.
- c. Suppose that the file is ordered by the key field SSN and we want to construct a primary index on SSN. Calculate
 - (i) the index blocking factor $bfri$
 - (ii) the number of first-level index entries and the number of first-level index blocks
 - (iii) the number of levels needed if we make it into a multilevel index
 - (iv) the total number of blocks required by the multilevel index, and
 - (v) the number of block accesses needed to search for and retrieve a record from the file—given its SSN value—using the primary index.

Problem 2. (15 pts)

Given the same specifications of Problem 1, consider this time you are building a primary index on SSN using B-tree. Calculate (i) the order p for the B-tree, (ii) the number of levels needed if blocks are approximately 69% full (round up for convenience), and (iii) the worst-case number of blocks needed to search for and retrieve a record from the file—given its SSN value—using the B-tree you are estimating.

Problem 3. (30 pts)

A PARTS file with Part# as key field includes records with the following Part# values: 3, 29, 39, 61, 21, 19, 6, 55, 22, 28, 14, 42, 2, 26, 8, 15, 9, 11. Suppose that the search field values are inserted in the given order in a B+-tree of order $p = 4$ and $p_{leaf} = 3$.

- (i) Show how the tree will expand (show all steps as in Fig 17.12 (7th ed)) and what the final tree will look like.
- (ii) What is the fill ratio of the B+-tree you created? (Note: we learned 69% is the average fill ratio in class.)

Note: 5% penalty for one day late submission. **Cut-off date 10/20/2023, 11:59 PM at HUSKYCT as the HW solution will be published 10/21/2023, morning at HUSKYCT.**