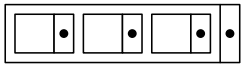
CS 5200: Module 10 HW: B+-tree Index

Name: Mengyuan Liu **NUID:** 001051575

Consider constructing a B^+-tree of order 4 (i.e. p=4 and p_{leaf} = 3)

(Q1) Show the resulting tree after inserting keys 10, 20, 30, 40, 50, 60, 70, 80, 90, 100 in this order.

- Inter node structure:  , - Leaf node structure: 

(Q2) Show the resulting tree after inserting keys 10, 20, 40, 50, 70, 80, 90, 100, 30, 60 in this order.

(Q3) Explain what is the difference between Q1 and Q2, and what do you learn from the difference?

(Q4) Index file

Given the following data file for EMPLOYEE,

EMPLOYEE [NAME (30 bytes), SSN (9 bytes), DEPT_CODE (9 bytes),
ADDRESS (40 bytes), PHONE(10 bytes), DOB (8 bytes), SEX (1 byte),
JOB_CODE (4 bytes), SALARY (4 bytes)]

Suppose that block size B=512 bytes, a block pointer P = 6 bytes, a record point P_R= 6, and the number of records r = 30,000. A Primary Key of EMPLOYEE is SSN (9 bytes). The records are fixed length and unspanned.

A. Calculate the total number of data file blocks b, the index entry size R_i in bytes, and the index blocking factor bfr_i, i.e., fan-out fo. bfr = floor(512/(30+9+9+40+10+8+1+4+4)) = floor(512/115) = 4

b = ceiling(30000/bfr) = ceiling(30000/4) = 7500

R_i = 9 + 6 = 15

bfr_i = floor(512/15) =34

B. Suppose that the file is ordered by the key field, i.e., SSN, and construct a **primary index** on SSN.
B-1. Calculate the number of first-level index entries L_1.

L_1 = b = 7500

B-2. Calculate the number of first-level index blocks b_{1i} (i.e., using single-level index)

b_{1i} = ceiling(7500/floor(512/15)) =221

B-3. Calculate the number of block accesses needed to run the following query using the primary index. You can use a binary search to access the first-level index.

```
SELECT * FROM EMPLOYEE WHERE SSN = "123456789";
```

binary search needs = 8

total needs = 8 + 1 = 9

C. Suppose that the file is not ordered by the key field SSN, and construct a **secondary index** on SSN. Note that the secondary index is a single level.
C-1. Calculate the number of first-level index entries

L_1 = r = 30000

C-2. Calculate the number of first-level index blocks (i.e., using the secondary index)

b_{1i} = ceiling(30000/floor(512/15)) = 883

C-3. Calculate the number of block accesses needed to run the following query using the secondary index. You can use a binary search to access the first-level index.

```
SELECT * FROM EMPLOYEE WHERE SSN = "123456789";
```

binary search needs = 10

total needs = 10 + 1 = 11