COMP 3311
DATABASE MANAGEMENT SYSTEMS

LECTURE 10 EXERCISES
STORAGE AND FILE STRUCTURE

$bf = \lfloor$ # bytes per page / # bytes per record $\rfloor$
pages = $\lceil$ # records / bf_r $\rceil$

EXERCISE 1

A Student file has 20,000 records of fixed-length. Assume the page size is 512 bytes and each record has the following fields:

name (30 bytes), studentId (8 bytes), address (40 bytes), phone (8 bytes), birthdate (8 bytes), gender (1 byte), majorDeptCode (4 bytes), minorDeptCode (4 bytes), classCode (4 bytes), and degreeProgram (3 bytes). An additional byte is used as a deletion marker.

a) What is the record size in bytes?

***record size*:** 30 + 8 + 40 + 8 + 8 + 1 + 4 + 4 + 4 + 3 + 1 = 111 bytes

b) What is the blocking factor $bf_{Student}$?

***$bf_{Student}$*:** $\lfloor$512 bytes per page / 111 bytes per Student record$\rfloor$
= 4 records/page

c) How many pages are needed to store the file?

Pages needed: $\lceil$20,000 records / 4 records per page$\rceil$ = 5000 pages

Student records: 20,000
Page size: 512 bytes
$bf_{Student} = 4$

EXERCISE 2

How many page I/Os are needed to search for a record given its studentId value if the file of Exercise 1 is organized as

a) a heap file?

Search cost: $\lceil 5{,}000 / 2 \rceil$ = 2500 page I/Os (linear search)

b) a sequential file sorted on studentId?

Search cost: $\lceil \log_2 5{,}000 \rceil$ = 13 page I/Os (binary search)

$bf = \lfloor$ # bytes per page / # bytes per record $\rfloor$
pages = $\lceil$ # records / bf_r $\rceil$

EXERCISE 3

An Employee file has 30,000 records of fixed-length. Assume the page size is 1,000 bytes and each record has the following fields:

name (25 bytes), hkid (8 bytes), address (35 bytes), deptCode (8 bytes), phone (8 bytes), birthdate (8 bytes), gender (1 byte), jobCode (3 bytes), salary (4 bytes). An additional byte is used as a deletion marker.

a) What is the record size in bytes?

***record size*:** 25 + 8 + 35 + 8 + 8 + 8 + 1 + 3 + 4 + 1 = 101 bytes

b) What is the blocking factor $bf_{Employee}$?

$bf_{Employee}$**:** $\lfloor$1000 bytes per page / 101 bytes per record$\rfloor$ = $\lfloor$9.9$\rfloor$ = 9 records/page

c) How many pages are needed to store the file?

Pages needed: $\lceil$30000 records / 9 records per page$\rceil$ = $\lceil$3333.3$\rceil$ = 3334 pages

Employee records: 30,000
Page size: 1000 bytes
$bf_{Student}$ = 9
Pages: 3,334

EXERCISE 4

For the file of Exercise 3, how many page I/Os are needed to search for

a) a record given its hkid value if the file is organized as a sequential file sorted on hkid?

Search cost: $\lceil \log_2 3334 \rceil = \lceil 11.7 \rceil = 12$ page I/Os (binary search)

b) all the records with a given jobCode value if the file is organized as a sequential file sorted on hkid?

Search cost: 3334 page I/Os (linear search)

c) a record given its hkid value if the file is organized as a hash file hashed on hkid and there are no overflow pages?

Search cost: 1 page I/O