COMP 3311: Database Management Systems

Lecture 13 Exercises
Indexing: Hash Index and Bitmap Index

Exercise 1: A company database has the following file and sizes of each field

Employee(employeeId: 6 bytes, employeeName: 10 bytes, departmentId: 4 bytes)

where departmentId is the id of the department where the employee works. There are 100,000 employee records and 1,000 departments (each department has 100 employees). A page is 1,000 bytes and a pointer is 4 bytes. Assume that the file is sorted on departmentId and there is no index. We want to build a *hash index on employeeId* on the above file where each entry has the form <employeeId, *pointer*>.

a) How many index entries are needed? (Briefly explain your answer.)

Index entries needed:

Explanation:

b) How many pages are required for these index entries (assuming full pages)?

Pages needed:

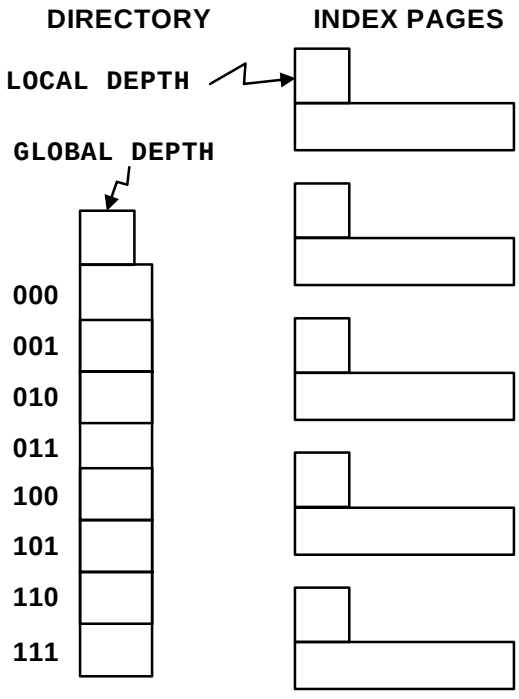
c) Using the hash index, what is the page I/O cost of retrieving the record of an employee with a given employeeId, assuming that there are *no overflow pages*? (Briefly explain your answer.)

Page I/O cost:

Explanation:

Exercise 2: Using the template below, construct a file that contains records with the given search-key values using extendable hashing. Use the hash function h(x) = x mod 8 and insert records one at a time in order into an empty file. Assume index pages can hold 4 records.

key value	1	4	5	7	10	12	15	16	20	24
h(x)	1	4	5	7	2	4	7	0	4	0
binary value	001	100	101	111	010	100	111	000	100	000



Do not upload this exercise sheet to Canvas.

Name: (1) _____/_____ **Student#:** (1) _____ **Date:** _____
Family/Given (PRINT) Given/First (PRINT)

Name: (2) _____/_____ **Student#:** (2) _____
Family/Given (PRINT) Given/First (PRINT)

NOTE: You are highly encouraged to do this exercise with a partner.

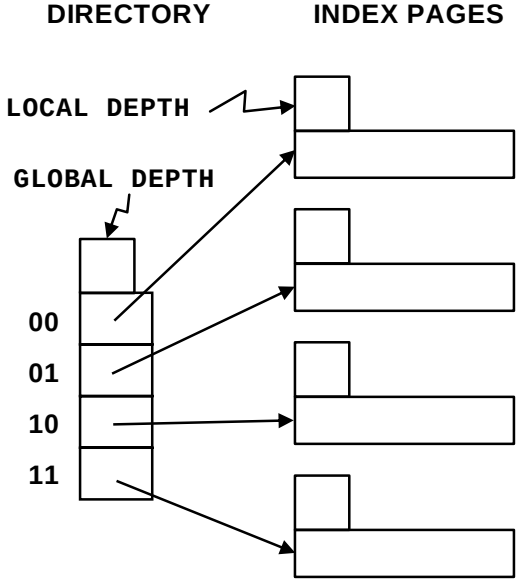
INDEX PAGES Database Management Systems

Lecture 13 Exercises
Indexing: Hash Index and Bitmap Index

Exercise 3: A global e-commerce website maintains a Customer file with attributes customerId, name, address, email and country. The file is organized as a hash file on customerId. There is also a secondary hash index on country. For each country there is only one index entry in the hash index. Assume that on average there are 9,000 customers for each country, there are 90 countries and that a page can hold 100 record pointers. How many page I/Os are required to retrieve the records of all the customers for a given country using the hash index on country?

Exercise 4: Using the template below, construct a file that contains records with the given search-key values using extendable hashing. Use the hash function h(x) = x mod 8 and insert records one at a time in order into an empty file. Assume index pages can hold 3 records.

key value	2	3	5	7	11	17	19	23	29	31
h(x)	2	3	5	7	3	1	3	7	5	7
binary value	010	011	101	111	011	001	011	111	101	111



You must upload this completed exercise sheet to Canvas by **11 p.m. today.**

Name: (1) _____/_____ **Student#:** (1) _____ **Date:** _____
Family/Given (PRINT) Given/First (PRINT)

Name: (2) _____/_____ **Student#:** (2) _____
Family/Given (PRINT) Given/First (PRINT)

NOTE: You are highly encouraged to do this exercise with a partner.

COMP 3311: Database Management Systems

Lecture 13 Exercises
Indexing: Hash Index and Bitmap Index

Exercise 5: Given the Customer relation and only the two bitmap indexes on gender and rating shown below, explain how you would use the bitmap indexes to answer the following queries? If the bitmap indexes are not useful, explain why.

gender index

male	female
1	0
1	0
0	1
1	0

Customer

id	name	gender	rating
112	Joe	m	1
115	Ram	m	5
119	Sue	f	5
112	Woo	m	2

rating index

1	2	3	4	5
0	1	0	0	0
0	0	0	0	1
0	0	0	0	1
1	0	0	0	0

DO NOT CALCULATE THE RESULT OF A QUERY*. Explain how to obtain the result using the bitmaps.*

a) Explain how you would use the two bitmap indexes to determine how many customers with a rating less than 3 are male? If the bitmap indexes are not useful, explain why.

b) Explain how you would use the two bitmap indexes to determine what percentage of customers are male? If the bitmap indexes are not useful, explain why.

c) Explain how you would use the two bitmap indexes to determine how many customer there are? If the bitmap indexes are not useful, explain why.

d) Explain how you would use the two bitmap indexes to determine how many customer are named Woo? If the bitmap indexes are not useful, explain why.

You must upload this completed exercise sheet to Canvas by **11 p.m. today.**