UNIVERSITY OF EDINBURGH

COLLEGE OF SCIENCE AND ENGINEERING

SCHOOL OF INFORMATICS

INFR11011 ADVANCED DATABASES

Thursday 4 th May 2017

14:30 to 16:30

INSTRUCTIONS TO CANDIDATES

Answer any TWO of the three questions. If more than two questions are answered, only QUESTION 1 and QUESTION 2 will be marked.

All questions carry equal weight.

CALCULATORS MAY BE USED IN THIS EXAMINATION

Year 4 Courses

Convener: I. Murray
External Examiners: A. Cohn, A. Donaldson, S. Kalvala

THIS EXAMINATION WILL BE MARKED ANONYMOUSLY

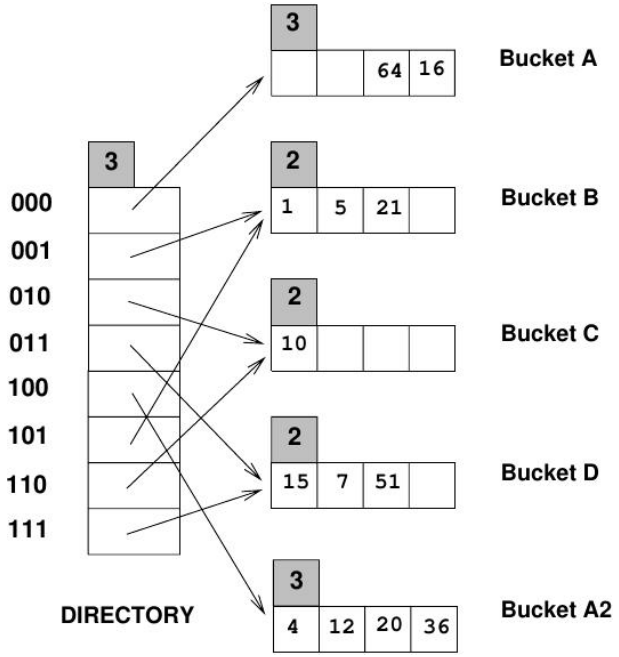
1. Answer the following questions about **Extensible Hashing**:

   (a) Explain what *global depth* and *local depth* are and why they are needed. Can the information provided by the local depth be obtained by other means if it is not given explicitly? [*7 marks*]

   (b) After an insertion that causes the directory size to double, how many buckets have exactly one directory entry pointing to them? Justify your answer. [*4 marks*]

   (c) Does Extensible Hashing guarantee at most one disk access to retrieve a record with a given key value? Justify your answer. [*4 marks*]

   (d) Consider the following Extensible Hashing index:



   (i) Draw the index after inserting an entry with hash value 68. [*5 marks*]

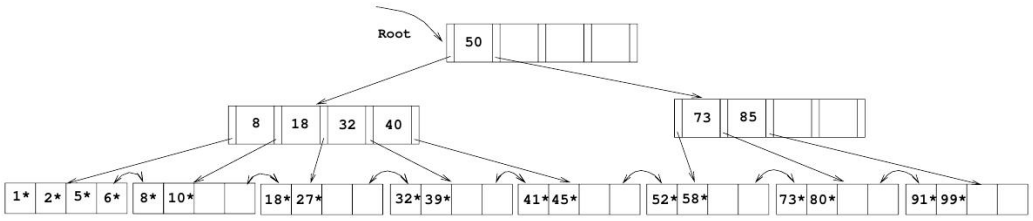
   (ii) Draw the index after inserting entries with hash values 17 and 69 (into the original index). [*5 marks*]

2. Answer the following questions about **Tree-Based Indexing**.

(a) What is the minimum space utilization for a B+ tree index and for an ISAM index, respectively? Justify your answer. [*5 marks*]

(b) If your database system supported both a static and a dynamic tree index (say, ISAM and B+ trees), would you ever consider using the static index in preference to the dynamic index? Justify your answer. [*5 marks*]

(c) Consider the following B+ tree index of order $d = 2$:



(i) Show the B+ tree that would result from inserting a data entry with key 9 into this tree. [*3 marks*]

(ii) Show the B+ tree that would result from inserting a data entry with key 3 **into the original tree**. How many page reads and page writes does the insertion require? Justify your answer. [*4 marks*]

(iii) Show the B+ tree that would result from deleting the data entry with key 8 **from the original tree**, assuming that the left sibling is checked for possible redistribution. [*4 marks*]

(iv) Show the B+ tree that would result from deleting the data entry with key 8 **from the original tree**, assuming that the right sibling is checked for possible redistribution. [*4 marks*]

3. Answer the following questions about **Memory Management** and **External Sorting**.

   (a) Explain what the buffer manager must do to process a read request for a page. [*5 marks*]

   (b) What does it mean that a page is *pinned* in the buffer pool? Who is responsible for pinning pages? Who is responsible for unpinning pages? Justify your answers. [*3 marks*]

   (c) What happens if a page is requested when all pages in the buffer pool are dirty? [*3 marks*]

   (d) Explain what *sequential flooding* of the buffer pool is, and why LRU is the worst possible replacement policy in this case. Give a concrete example. [*4 marks*]

   (e) We want to sort a file with 20,000 pages using the most general external sort algorithm we have seen in class and 5 buffer pages.

      (i) How many runs will you produce in the first pass? [*2 marks*]

      (ii) How many passes will it take to sort the file completely? [*2 marks*]

      (iii) What is the total I/O cost of sorting the file? [*2 marks*]

      (iv) How many buffer pages do you need to sort the file completely in just two passes? [*4 marks*]

   Justify your answers.