Homework 5

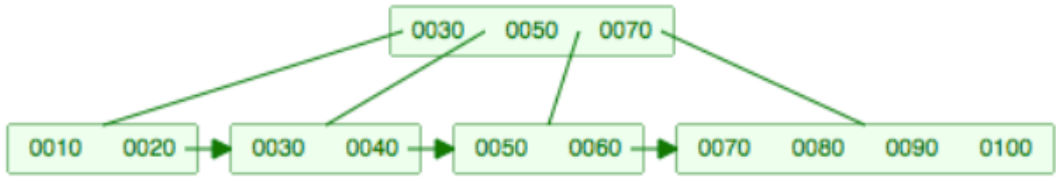
Prepare your answers using either a word processor or by neatly sketching diagrams and then scanning them into a single digital document (e.g., scan them into a word processing document and then create a pdf file of the word processing document).

1. Look at the hotel database shown in the exercises at the end of Chapter 4 and think about the types of queries that might be made on such a database. Feel free to examine the queries over that database shown in the exercises at the end of Chapter 6. However, I don't think many of the queries in those exercises would be frequently executed, and hence it would be hard to justify an index based on many of those queries. Conversely, I think there are many potential high-frequency queries that have been left out of those exercises (e.g., show me the rooms checking out today at the hotel I'm managing). Based on the queries you choose, and what was discussed in class about physical database design, describe what indices you would choose to create for each of the four relations and justify each index by indicating the queries that might make use of these indices. As an example of what I'm looking for, here is the Student relation from the courses database we used in class:

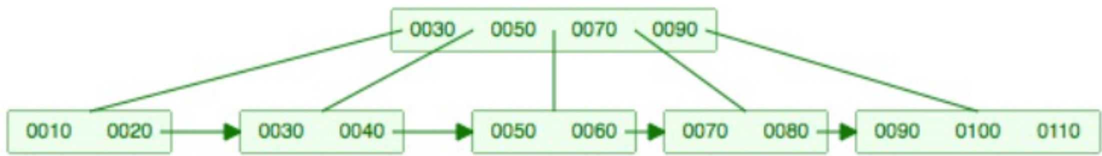
   Student(Id, Firstname, Lastname, GPA)
   Id: Secondary index because it will be frequently used in joins
   Firstname: No index since it may be printed as parts of other queries but we are unlikely to search directly for a first name
   Lastname: Secondary index because we will frequently need to retrieve a student's information based on their last name
   GPA: Clustering index because it is the attribute most likely to be used in a range query (e.g., give me all students with a GPA => 3.8). Note that it's a clustering index rather than a primary index because id is the primary key for this relation.

2. Perform the following operations on each of the example 5-way (M=5) B+-trees, with L = 4 (recall that M is the maximum number of children per interior node and L is the maximum number of records per leaf). I understand that you can go to the web-site used in class and simply copy the answer (and the tree the web-site produces might be wrong any way because when leaf nodes have an extra tuple it copies the extra tuple to the right node and I told you to copy the extra tuple to the left node). However, that won't help you on the exam, where the problem will count a lot more toward your final grade.
   a. Insert 32

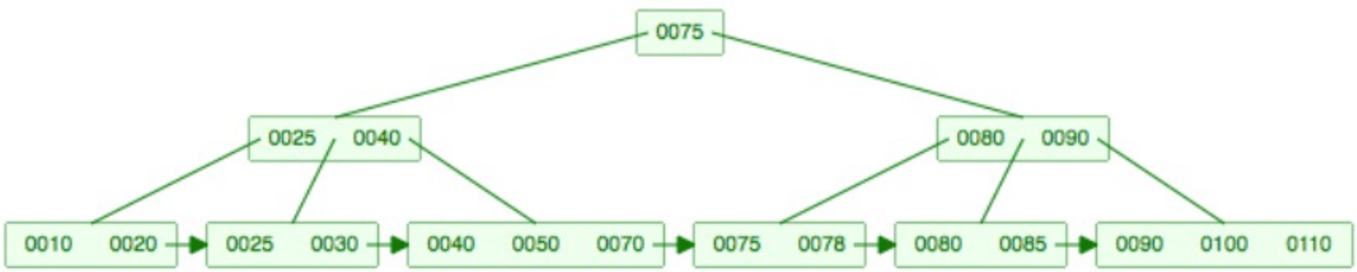
   b. Insert 110



c. Delete 40



d. Delete 75



3. Draw the extendible hash structure that results from inserting the following set of keys into an initially empty table with M = 4. Assume the keys are inserted in the order given: 10111101, 00000010, 10011011, 10111110, 01111111, 01010001, 10010110, 00001011, 11001111, 11011011, 11110000, 10011110. At the beginning of each bucket you should denote the number of leading bits being used in that bucket (put that number in parentheses as is done in the figures I drew in class).