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). When the two normal form problems say "describe and illustrate the process of normalizing", I want to see three sets of relations--the 1st NF, 2nd NF, and 3rd NF relations--**and the functional dependencies you used to go from 1st NF to 2nd NF and then from 2nd NF to 3rd NF**.
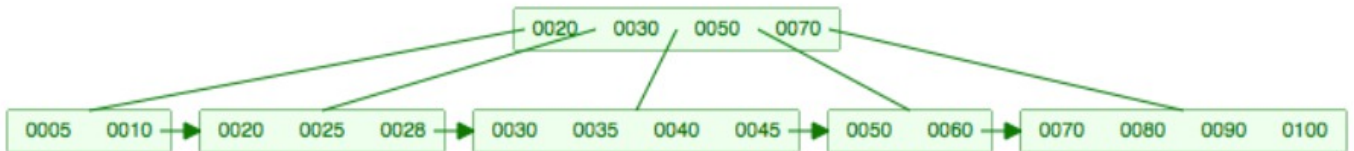
1. 14.14. You may need to read parts of the Wellmeadows Hospital case study in Appendix B. You may make the following assumptions:
   a. A drug's dosage and method of administration is determined by its drug no
   b. Each ward has unique bed numbers, but they may not be unique throughout the hospital (i.e., the bed number may be replicated in another ward).
   c. The relation only stores patients currently registered at the hospital and hence each patient will be in at most one bed and one ward
   Use reasonable assumptions about the remaining fields on the form. Make sure that you look at the fields at the top of the form, such as full name and patient number, and not just at the fields in the table.

2. 14.15. Assume that this appointments relation not only contains information about the appointments, but is also the only repository of information about the dentists in the practice, and the patients in the practice. Assume that surgeries has another relation which is referenced by the surgery number and contains information about the surgeries. You may make the following assumptions:
   a. When the book says "assigned to a surgery" the book is talking about the location (e.g., such as a building or clinic) in which the operation is to be performed, not the operation itself.
   b. a patient can be registered for only one surgery (i.e., clinic). You will note that patients may have multiple appointments but are registered for the same surgery at each appointment. The book makes this assumption as well in its solution.
   c. a patient may have multiple appointments on the same day.

3. 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.
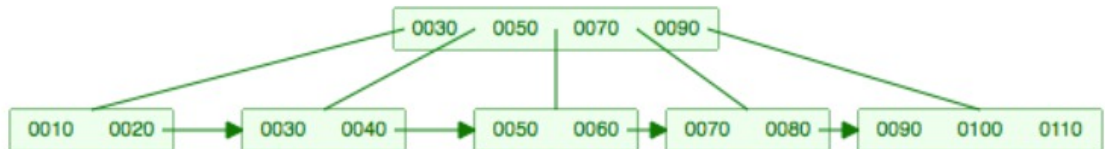
4. Perform the following operations on each of the example 5-way (M=5) B+-trees, with L = 4 (recall that M is the maximum number of children per interior node and L is the maximum number of records per leaf). I understand that you can go to the web-site used in class and simply copy the answer (and the 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

| 0020 | 0030 | 0050 | 0070 |

| 0005 | 0010 | → | 0020 | 0025 | 0028 | → | 0030 | 0035 | 0040 | 0045 | → | 0050 | 0060 | → | 0070 | 0080 | 0090 | 0100 |

    b. Insert 110

| 0030 | 0050 | 0070 |

| 0010 | 0020 | → | 0030 | 0040 | → | 0050 | 0060 | → | 0070 | 0080 | 0090 | 0100 |

    c. Delete 40

| 0030 | 0050 | 0070 | 0090 |

| 0010 | 0020 | → | 0030 | 0040 | → | 0050 | 0060 | → | 0070 | 0080 | → | 0090 | 0100 | 0110 |

    d. Delete 75

| 0075 |

| 0025 | 0040 | ... | 0080 | 0090 |

| 0010 | 0020 | → | 0025 | 0030 | → | 0040 | 0050 | 0070 | → | 0075 | 0078 | → | 0080 | 0085 | → | 0090 | 0100 | 0110 |