CSc 460 — Database Design
Fall 2009 (McCann)

http://www.cs.arizona.edu/classes/cs460/fall09/

Homework #1

(100 points)

Due Date: October 1st, 2009, at the beginning of class

Remember: Solutions to this homework must be electronically submitted as a single PDF file (using lectura's `turnin`). Diagrams may be created with a suitable diagramming tool; several open-source and/or free options exist (xfig, dia, OpenOffice, etc.). You may also draw your diagrams by hand and scan/photograph them for inclusion within your solutions document, so long as the result is readable and reasonably sized.

Write complete, legible answers to each of the following questions. A problem identified as "M.N(x,y)" references parts x and y of exercise N from chapter M of the Connolly/Begg text, 5th edition. Show your work, when appropriate, for possible partial credit. This is not a group project; do your own work. We will set aside time before the exam to answer any questions you might have about these questions. We will post our solutions 24 hours after the due date (remember, you can use one late day on homeworks, so we can't give solutions on the due date).

On the due date, by the start of class, hand-in a printout of your solutions **and** submit your electronically–formatted version of your solutions (the turnin folder is `cs460h1`). If you need to submit your solutions within the 24-hour late window, place your printout in one of our mailboxes in CS 713 as soon as you are able to do so. Solutions submitted more than 24 hours after the due date and time will not be accepted.

1. (5 points) 1.4
2. (5 points) 2.2
3. (5 points) 3.2
4. (5 points) 3.12(a) , but use the Postgres DBMS (`www.postgresql.org`) as your subject
5. (5 points) 4.2
6. (5 points) 4.5
7. (5 points) 12.2 (skip quaternary, and find examples not in the text)
8. (5 points) 12.7 (find an example not used in class or the text)
9. (20 points) 12.12 (all parts)
10. (5 points) 13.10
11. (5 points) You have a RAID system of four hard drives. Two of the drives have a failure rate of 1.5% ($p_f = 0.015$), and the other two have a failure rate of 3%. Assume that failures of the hard drives are independent events. (Reminder: The lecture slides covering this subject were presented on Sept. 3rd.)
    (a) What is the probability of failure of your RAID system?
    (b) Assuming that your quantities of 1.5% and 3% drives remain equal (that is, you have n 1.5% and n 3% drives), what is the minimum total number of hard drives your RAID system would have to have to exceed a 50% failure rate?

(Continued on back . . .)

12. (10 points) In Program #2, you extendibly-hashed using decimal digits of a key. For this exercise, assume that our keys are in Base 3 instead of Base 10. Also assume that buckets are disk blocks that can hold at most three keys each. Build a **dynamic** hashing (not extendible hashing!) index structure using the keys listed below, and draw the final structure.

    1112, 2210, 0001, 0111, 1221, 0120, 1210, 2100, 1102, 0122, 0012, 2112

13. (20 points) Assuming a B^+-tree of Order 2 (using the order concept of Comer's definition).

    (a) Insert the values J, A, F, C, R, S, and L, in the order presented. Show the tree after each insertion that causes the tree to grow by a level, and show the final tree.

    (b) Delete, from your final tree of part (a), the following keys, and show the final tree: L and F (in that order). Again, you may show intermediate trees if you so desire.