Homework #1
(95 points)

Due Date: February 16th, 2022, at the beginning of class

Directions

1. This is an INDIVIDUAL assignment; do your own work! Submitting answers created by other people is NOT doing your own work.

2. Start early! Getting help is much easier n days before the due date/time than it will be n hours before.

3. Write complete answers to each of the following questions, in accordance with the given directions. Create your solutions as a PDF document such that each question is on a separate page; all parts of a multi-part question may be on the same page. Show your work, when appropriate, for possible partial credit.

4. Questions of the form “α.β” are found in the Connolly/Begg text, available via D2L, as question β at the end of chapter α. For questions with sub-parts, note that a notation of “(w,z)” is asking you to complete parts w and z only, not parts x and y.

5. If you have questions about any aspect of this assignment, help is available from the class staff via piazza.com and our office hours.

6. When your answers are ready to be turned in, do so on gradescope.com. Be sure to assign pages to problems after you upload your PDF. Need help? Visit https://help.gradescope.com/ and search for “Submitting an Assignment.”

7. Remember that you can use at most one late day on a homework assignment, because we will be distributing solutions after that time.

1. (5 points) 1.4
2. (5 points) 2.2
3. (5 points) 3.2
4. (5 points) 3.18(a), but use only the MySQL DBMS (www.mysql.com) as your subject. We’re looking for a good-faith effort at finding relevant information.
5. (5 points) 4.2
6. (5 points) 4.5
7. (5 points) 12.2 (skip quaternary, and provide examples not used in class or found in the text)
8. (5 points) 12.7 (provide an example not used in class or found in the text)
9. (20 points) 12.12 (all parts)
10. (5 points) 13.10

(Continued on back . . .)
11. (10 points) Assume a RAID 0 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.

(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?

12. (10 points) In Program #2, you extendibly-hashed using decimal digits (Base 10) extracted from their key. For this exercise, assume dynamic hashing with keys that are in Base 3 instead of Base 10, and assume that buckets are disk blocks that can hold at most three keys each. Insert into the dynamic hashing index the keys listed below, in the order provided, and draw the final structure.

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

13. (10 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. You may show additional intermediate trees if you wish to do so.

(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.