

Answers to some of these questions are available at the solutions manual for the textbook by Ramakrishnan. Google: Wisconsin dbbook solutions.

Q1. Drugwarehouse.com has offered you a free life-time supply of prescription drugs (no questions asked) if you design its database schema. Given the rising cost of health care, you agree. Here is the information that you gathered:

- Patients are identified by their SSN, and we also store their names and age.
  - Doctors are identified by their SSN, and we also store their names and specialty.
  - Each patient has one primary care physician, and we want to know since when the patient has been with her primary care physician.
  - Each doctor has at least one patient.
- a) Draw an ER diagram that captures the above information
- b) Write the DDLs for the ER diagram you have drawn. Include primary and foreign key constraints. You may assume any reasonable data types.

Q2. Consider the relation Stocks(B,O,I,S,Q,D), whose attributes may be thought of informally as broker, office (of the broker), investor, stock, quantity (of the stock owned by the investor), and dividend (of the stock). Let the set of FD's for Stocks be  $\{S \rightarrow D, I \rightarrow B, IS \rightarrow Q, B \rightarrow O\}$

- a) What are all the keys for Stocks ?
- b) What normal forms (BCNF or 3NF) does Stocks conform to ? (If not, indicate violating FDs).
- c) Decompose Stocks into the strongest normal form so that the decomposition is both lossless join and dependency preserving.

Q3. [Ex. 9.6] Consider again the disk specifications from Ex. 9.5 (see Ex.17), and suppose that a block size of 1024 bytes is chosen. Suppose that a file containing 100,000 records of 100 bytes each is to be stored on such a disk and that no record is allowed to span two blocks.

- a) How many records fit onto a block?
- b) How many blocks are required to store the entire file? If the file is arranged sequentially on the disk, how many surfaces are needed?
- c) How many records of 100 bytes each can be stored using this disk?
- d) If pages are stored sequentially on disk, with page 1 on block 1 of track 1, what page is stored on block 1 of track 1 on the next disk surface? How would your answer change if the disk were capable of reading and writing from all heads in parallel?
- e) What time is required to read a file containing 100,000 records of 100 bytes each sequentially? Again, how would your answer change if the disk were capable of reading/writing from all heads in parallel (and the data was arranged optimally)?
- f) What is the time required to read a file containing 100,000 records of 100 bytes each in a random order? To read a record, the block containing the record has to be fetched from disk. Assume that each block request incurs the average seek time and rotational delay.

Q4. [Ex. 12.5] Consider again the schema with the Sailors relation:

Sailors(sid: integer, sname: string, rating: integer, age: real)

Assume that each tuple of Sailors is 50 bytes long, that a page can hold 80 Sailors tuples, and that we have 500 pages of such tuples. For each of the following selection conditions, estimate the number of pages retrieved, given the catalog information in the question.

- a) Assume that we have a B+-tree index T on the search key 'Sailors.sid', and assume that  $\text{IndexHeight}(T) = 4$ ,  $\text{IndexLeafPages}(T) = 50$ ,  $\text{LowestKey}(T) = 1$ , and  $\text{HighestKey}(T) = 100,000$ .
- a)  $\sigma_{\text{Sailors.sid} < 50,000}(\text{Sailors})$
- b)  $\sigma_{\text{Sailors.sid} = 50,000}(\text{Sailors})$
- b) Assume that we have a hash index T on the search key 'Sailors.sid', and assume that  $\text{IndexHeight}(T) = 2$ ,  $\text{IndexLeafPages}(T) = 50$ ,  $\text{LowestKey}(T) = 1$ , and  $\text{HighestKey}(T) = 100,000$ .

- a)  $\sigma_{Sailors.sid < 50,000}(Sailors)$
- b)  $\sigma_{Sailors.sid = 50,000}(Sailors)$

---

Q5. [Ex. 16.3.] Consider a database with objects X and Y and assume that there are two transactions T1 and T2. Transaction T1 reads objects X and Y and then writes object X. Transaction T2 reads objects X and Y and then writes objects X and Y.

- a) Give an example schedule with actions of transactions T1 and T2 on objects X and Y that results in a write-read conflict.
  - b) Give an example schedule with actions of transactions T1 and T2 on objects X and Y that results in a read-write conflict.
  - c) Give an example schedule with actions of transactions T1 and T2 on objects X and Y that results in a write-write conflict.
  - d) For each of the three schedules, show that Strict 2PL disallows the schedule.
-