

Homework 4

Normalization, data storage and transaction management

Total points: 40

Complete the following exercises (some are from the textbook). Submit a document with your solutions.

Chapter 7

8.21 (4 pts.) Use the schema given in this exercise. Ignore the double arrow and assume that it is a single arrow that represents a functional dependency. Take each of the two relations and specify which normal form they are in. Then normalize them to 3NF. Hint: AccessionNo and userid are the primary keys.

```
books(accessionno, isbn, title, author, publisher)
users(userid, name, deptid, deptname)
accessionno → isbn
isbn → title
isbn → publisher
isbn →→ author
userid → name
userid → deptid
deptid → deptname
```

8.35 (5 pts.) – This question is not included in the textbook.

Ø Consider the relation R(A,B,C,D,E,F,G,H) in which ABC is the primary key. If the following dependencies hold in this relation, is this relation in 3NF? If not, which normal form does it satisfy? Reduce it to 3NF.

```
AB -> D
C -> H
EF-> G
```

Chapter 11

11.15 (2 pts.) When is it preferable to use a dense index rather than a sparse index? Explain your answer.

11.16 (2 pts.) What is the difference between a clustering index and a secondary index?

11.19 (3 pts.) Explain the distinction between closed and open hashing. Discuss the relative merits of each technique in database applications.

11.21 (2 pts.) Why is a hash structure not the best choice for a search key on which range queries are likely?

11.29 (5 pts.) – This question is not included in the textbook. This material is covered in the lectures.

Ø Consider a disk with the following parameters:

```
B = 512 bytes
P = 6 bytes
Pr = 7 bytes
```

and an Employee data file with the following parameters:

r = 30000

R = 115 bytes

bfr = 4 records/block

Suppose the file is not ordered by the key field SSN (9 bytes) and we want to construct a *secondary index* with record pointers on SSN. Calculate the following:

- a. The index blocking factor, bfr_i .
- b. The number of first-level index entries and the number of number of first-level index blocks.
- c. The number of levels needed if we make it into a multi-level index
- d. The total number of blocks required by the multi-level index
- e. The number of block accesses needed to search for and retrieve a record from the file – given its SSN value – using the primary index.

Chapter 14

14.12 (5 pts.) List the ACID properties. Explain the usefulness of each.

14.13 (2 pts.) No explanations needed. During its execution, a transaction passes through several states, until it finally commits or aborts. List all possible sequences of states through which a transaction may pass. Explain why each state transition may occur.

14.14 (2 pts.) Explain the distinction between the terms serial schedule and serializable schedule.

Chapters 15 & 16

15.40 (8 pts.) – This question is not included in the textbook.

Define or explain briefly the following terms.

Ø 2PL

Ø Deadlock

Ø Deferred database modification

Ø Checkpoint