UCLA
Computer Science Department
Spring 2011

Instructor: J. Cho

Student Name and ID: _____

CS143 Final: Closed Book, 90 minutes

(IMPORTANT PLEASE READ **):**

- There are 4 problems on 8 pages to be completed in 90 minutes. *You should look through the entire exam before getting started, in order to plan your strategy.*
- *Simplicity and clarity of your solutions will count.* You may get as few as 0 point for a problem if your solution is far more complicated than necessary, or if we cannot understand your solution.
- If you need to make any assumption to solve a question, please write down your assumptions.
- To get partial credits, you may want to write down how you arrived at your answer step by step.
- You may use two double-sided cheat-sheets during exam. You are also allowed to use a calculator.
- Please, write your answers neatly. Attach extra pages as needed. Write your name and ID on the extra pages.

Problem	**Score**	
1	20	
2	14	
3	15	
4	01	
Total	50	

Problem 1 (Index): 20 points

1. Consider a B+tree of order (or branching factor) n.

   (a) What is the minimum number of records that the tree can index, when it has 2 levels? (3 points)

   ANSWER: $2 \times \lceil (n-1)/2 \rceil$

   (b) What is the minimum number of records that the tree can index, when it has k levels ($k \geq 2$)? (3 points)

   ANSWER: $2 \times \lceil n/2 \rceil^{k-2} \times \lceil (n-1)/2 \rceil$

   (c) What is the maximum number of records that the tree can index, when it has 2 levels? (3 points)

   ANSWER: $n \times (n-1)$

   (d) What is the maximum number of records that the tree can index, when it has k levels ($k \geq 2$)? (3 points)

   ANSWER: $n^{k-1} \times (n-1)$

   (e) Suppose that we are using a B+tree with $n = 99$ to index a table with 300,000 records. In the worst case, how many nodes in the index will one have to examine while looking up a record? (3 points)

   ANSWER: $2 \times \lceil n/2 \rceil^{k-2} \times \lceil (n-1)/2 \rceil \leq 300,000$ `where` $n = 99$. `So` $2 \times 50^{k-2} \times 49 \leq 300,000$. `The minimum such` k `is 4.`

2. Consider the following movie-ratings database with two tables:

```
Movie(title, director)       // every movie title is unique
Rating(person, title, score) // every person name is unique
```

Suppose there are three types of queries commonly asked on this schema:

- Given a movie title, find the director of the movie.
- Match each person with the directors of movies the person has rated.
- Given a person, find the titles of all movies the person has rated.

(a) What is the *minimum* number of indexes needed to speed up *all three* types of queries? (2 points)

ANSWER: `2`

(b) On which attributes should these indexes be created? (3 points)

ANSWER: `title, person`

Problem 2 (Database Design): 14 points

1. The following functional dependencies hold in a relation $R(A, B, C, D, E, F)$:

   $AB \rightarrow C$
   $DE \rightarrow F$
   $F \rightarrow B$
   $C \rightarrow E$

   List all the keys for R. Make sure to clearly denote when a set of attributes together constitute a key (as opposed to each attribute being a key on its own). Remember that a key should be minimal. (5 points)

ANSWER: `ABD, ACD, ADE, ADF.`
`Since neither A or D is funtionally dependent on any other attribute, a key should always contain these two attributes. But AD cannot be a key since` $\{AD\}+ = \{AD\}$. `AD and one more attribute -- ABD, ACD, ADE, ADF --, they are all keys since their closures are all attributes.`

2. Consider the following two relational schemas:

   Schema 1: $R(A, B, C)$
   Schema 2: $R1(A, B)$, $R2(A, C)$

   (a) Suppose that the only dependencies (functional or multivalued) that hold on the relations in these schemas are $A \rightarrow BC$ and all dependencies that follow from this one. Identify the two correct statements below by circling them (3 points)

   i. Schema 1 is in neither BCNF nor 4NF
   ii. Schema 1 is in BCNF but not 4NF
   iii. Schema 1 is in 4NF but not BCNF
   iv. Schema 1 is in both BCNF and 4NF
   v. Schema 2 is in neither BCNF nor 4NF
   vi. Schema 2 is in BCNF but not 4NF
   vii. Schema 2 is in 4NF but not BCNF
   viii. Schema 2 is in both BCNF and 4NF

   `ANSWER: iv, viii`

   (b) Now suppose that the only dependencies (functional or multivalued) that hold on the relations in these schemas are $BC \rightarrow A$, $A \rightarrow C$, and all dependencies that follow from these two. Identify the two correct statements below by circling them (3 points)

   i. Schema 1 is in neither BCNF nor 4NF
   ii. Schema 1 is in BCNF but not 4NF
   iii. Schema 1 is in 4NF but not BCNF
   iv. Schema 1 is in both BCNF and 4NF
   v. Schema 2 is in neither BCNF nor 4NF
   vi. Schema 2 is in BCNF but not 4NF
   vii. Schema 2 is in 4NF but not BCNF
   viii. Schema 2 is in both BCNF and 4NF

   `ANSWER: i, viii`

   (c) Now suppose that the only dependencies (functional or multivalued) that hold on the relations in these schemas are $A \twoheadrightarrow B$ and all dependencies that follow from this one. Identify the two correct statements below by circling them (3 points)

   i. Schema 1 is in neither BCNF nor 4NF
   ii. Schema 1 is in BCNF but not 4NF
   iii. Schema 1 is in 4NF but not BCNF
   iv. Schema 1 is in both BCNF and 4NF
   v. Schema 2 is in neither BCNF nor 4NF
   vi. Schema 2 is in BCNF but not 4NF
   vii. Schema 2 is in 4NF but not BCNF
   viii. Schema 2 is in both BCNF and 4NF

   `ANSWER: ii, viii`

Problem 3 (Transactions): 15 points

Consider table `Lakers(player,salary)` where player is a key, and the following two transactions:

```
T1: Begin Transaction
S1: update Lakers set salary = 2*salary where player = 'Kobe Bryant'
S2: update Lakers set salary = 3*salary where player = 'Kobe Bryant'
T1: Commit

T2: Begin Transaction
S3: update Lakers set salary = salary-20 where player = 'Kobe Bryant'
S4: update Lakers set salary = salary-10 where player = 'Kobe Bryant'
T2: Commit
```

You may assume that **the individual statements S1, S2, S3, and S4 always execute atomically**. Let Kobe's salary be 50 before either transaction executes.

1. Suppose both transactions T1 and T2 execute to completion with isolation level Serializable. What are Kobe's possible final salaries? (3 points)

   ANSWER: `120, 270`
   `Only serializable schedules are allowed when both transactions are serializable. Therefore, either T1 T2 or T2 T1.`

2. Suppose both transactions T1 and T2 execute to completion with isolation level Read-Committed. What are Kobe's possible final salaries? (3 points)

   ANSWER: `120, 270`
   `S1 S2 S3 S4 and S3 S4 S1 S2 are both possible because they are serial schedules. But we cannot intermix the statements from T1 and T2 because it leads to dirty read for at least one statement. For example, S1 S3 S2 S4 leads to dirty reads for S3 and S2.`

3. Suppose transaction T1 executes with isolation level Read-Committed, transaction T2 executes with isolation level Read-Uncommitted, and both transactions execute to completion. What are Kobe's possible final salaries? (3 points)

   ANSWER: `120,210,270.`
   `Other than two serial schedules, the only other schedule that is allowed is S1 S3 S4 S2.`

4. Suppose both transactions T1 and T2 execute to completion with isolation level Read-Uncommitted. What are Kobe's possible final salaries? (3 points)

   ANSWER: `120,150,170,210,230,270.`
   `There are 6 possible schedules: S1 S2 S3 S4. S1 S3 S2 S4. S1 S3 S4 S2. S3 S1 S2 S4. S3 S1 S4 S2. S3 S4 S1 S2.`

5. Suppose both transactions T1 and T2 execute with isolation level Serializable. Transaction T1 executes to completion, but transaction T2 rolls back after statement S3 and does not re-execute. What are Kobe's possible final salaries? (3 points)

   ANSWER: `300.`
   `A serializable transaction will never see an uncommitted value from a rolledback transaction.`

Problem 4 (CS143): 1 point

Please write down

1. One thing that you liked most in this class.

2. One thing you hated most (or hope to be changed) in this class.