CS 411: Database Systems Fall 2018 Homework 3 (Due by 23:59 CDT on October 26)

October 15 2018

1 Relational database design

1. Consider the following relations $R1$ and $R2$. What are the normal forms (BCNF, 3NF) that each of the relations are in? Justify why $R1/R2$ is or is not in each of these normal forms. (4 Points)

   (a) $R1(A, B, C, D, E)$ with a set of functional dependencies $FD = \{D \rightarrow B, CE \rightarrow A\}$

   (b) $R2(A, B, C, D, E)$ with a set of functional dependencies $FD = \{A \rightarrow E, BC \rightarrow A, DE \rightarrow B\}$

2. Consider the relation $R(A, B, C, D, E, F, G, H)$ with the functional dependencies $FD = \{A \rightarrow BC, ABCD \rightarrow E, EF \rightarrow GH, ABDF \rightarrow EC\}$

   (a) Compute the minimal cover of FD. Show the steps of your computation. (4 Points)

   (b) Decompose the relation, as necessary, into a collection of relations that are in 3NF. (4 Points)

3. Consider the relation $R(A, B, C, D, E, F)$ and the functional dependencies $F = \{A \rightarrow B, C \rightarrow AF, B \rightarrow D\}$

   Are the following decompositions lossless? Explain your answer. (6 Points)

   (a) $(AEF)(ABCDE)$

(b) $(CB)(ACDEF)$

(c) $(ABC)(BCDF)$

4. Consider the relation $R(A, B, C, D, E, F)$ with the functional dependencies $F = \{A \to B, C \to AD, BE \to DF\}$

   (a) Decompose the relation, as necessary, into a collection of relations that are in BCNF. Explain which dependency violation you are correcting by your decompositions. (6 Points)

   (b) Do we have a unique BCNF decomposition for R? (2 Points)

5. Consider the relation $R(A, B, C, D, E, F)$ with the functional dependencies $FD = \{A \to BD, AC \to DE, D \to F\}$.

   Is there a BCNF decomposition for R that preserves the functional dependencies? Justify your answer. (4 Points)

2 Updates and Transactions

1. Consider the following schedules S_1 and S_2 by transactions T_1, T_2 and T_3 on database objects A and B.

   $\mathbf{S}_1 : R_1(A), R_3(B), W_1(B), R_2(B), W_3(A), W_1(A), R_3(A)$

   $\mathbf{S}_2 : W_1(A), R_1(B), W_2(B), W_1(B), R_2(A), W_3(B), R_3(A)$

   (a) Determine if schedules S_1 and S_2 are conflict-serializable by drawing a precedence graph of transactions. (2+2 = 4 points)

   (b) For the same schedules S_1 and S_2, determine if they are serializable. If serializable, then give a corresponding serial schedule. If not serializable, provide a suitable justification. (2+2 = 4 points)

2. Consider another set of schedules S_1 and S_2 by transactions T_1, T_2 and T_3 on database objects A, B and C

   $\mathbf{S}_1 : R_2(A), R_1(B), R_3(C), R_2(B), R_1(C), W_2(A), W_1(B), W_3(C)$

S$_2$: $R_1(C), W_2(A), W_3(B), W_3(C), R_2(B), W_2(B), R_2(A), R_3(C)$

(a) Determine whether schedules S_1 and S_2 can be produced by a Two Phase Lock (2PL) scheduler. If yes, provide a table with placements of shared locks (Slock), exclusive locks (Xlock) and lock releases (Rel) that follows 2PL and obeys lock-compatibility restriction. Structure the table so that each column corresponds to a single transaction and each row has a single operation. Or else, give a partial table marked with the first point of the schedule where a transaction fails to get a required lock. (4 points)

(b) For the same schedules S_1 and S_2, determine whether schedules S_1 and S_2 can be produced by a *strict* Two Phase Lock (2PL) scheduler. Similar to part (a), provide the corresponding lock placement table. (4 points)

3. Given two transaction T_1 and T_2 operating on objects A, B and C.

T_1 : $R_1(A), R_1(B), W_1(B), R_1(C)$

T_2 : $R_2(A), W_2(C), R_2(A), W_2(A)$

Consider two operating scenarios of isolation levels for these transactions:

- Scenario 1: Read Committed (T_1) and Repeatable Read (T_2).
- Scenario 2: Repeatable Read (T_1) and Repeatable Read (T_2).

For each scenario, answer the following questions:

(a) Rewrite each transaction with appropriate locks for each of the below scenarios. For full credit, place the locks to maximize concurrency in each transaction. (2 + 2 = 4 points)

(b) Check whether a deadlock is possible. If deadlock is possible, then give an example of a schedule in which deadlock happens under the assumed isolation levels (Use a table to show the place of deadlock). If it is not possible, provide a clear explanation that covers all allowed schedules. (2 + 2 = 4 points)

4. Consider two transactions T_1 and T_2 operating on database objects A and B.

T_1 : $R_1(A), R_1(B), W_1(B)$

T_2 : $W_2(A), R_2(B), W_2(B)$

(a) What is the total number of possible schedules of transactions T_1 and T_2 ? (1 point)

(b) Among the possible schedules of transactions T_1 and T_2, how many are conflict-equivalent to the serial order (T_1, T_2) ? (2 point)

(c) Determine the number of possible conflict serializable schedules. (3 points)