CSEP 544 Homework 4

Objectives
To understand query execution, recovery and concurrency control.

Reading Assignments
Chapters 16,17,18 from **R&G** and Chapter 18 from **DS:CB**.

Number of points:
100

Due Date
Wednesday, Nov 10, at 12:01 am. Submit a pdf file containing the answers via `https://catalyst.uw.edu/collectit/dropbox/joyleung/11920`.

Notes
- **R&G** refers to our class textbook by Ramakrishnan and Gehrke.
- **DS:CB** refers to "Database Systems: The Complete Book 1E." (Not required).

Questions
1. **[15 points]** After a systems crash, the undo-log using non-quiescent checkpointing contains the following data:

<START T1>
<T1, X1, 1>
<START CKPT ????>
<START T2>
<T2, X2, 2>
<T1, X1, 3>
<START T3>
<COMMIT T1>
<END CKPT>
<START CKPT ????>
<T2, X2, 4>
<T3, X3, 5>
<START T4>
<COMMIT T2>
<T4, X4, 6>
<COMMIT T3>
<END CKPT>
<START T5>
<T5, X5, 7>
<START CKPT ????>
<T4, X4, 8>
CRASH !!!

   (a) What are the correct values of the three <START CKPT ????> records ? You have to provide three correct values for the three ????s.

   (b) Assuming that the three <START CKPT ???> records are correctly stored in the log, according to your answer at a., show which elements are recovered by the undo recovery manager and compute their values after recovery.

   (c) Indicate what fragment of the log the recovery manager needs to read.

2. **[10 points]** (**Exercise 18.2.4 DS:CB**) For each of the following schedules:

a) $r_1(A)$; $r_2(A)$; $r_3(B)$; $w_1(A)$; $r_2(C)$; $r_2(B)$; $w_2(B)$; $w_1(C)$;

b) $r_1(A)$; $w_1(B)$; $r_2(B)$; $w_2(C)$; $r_3(C)$; $w_3(A)$;

d) $r_1(A)$; $r_2(A)$; $w_1(B)$; $w_2(B)$; $r_1(B)$; $r_2(B)$; $w_2(C)$; $w_1(D)$;

Answer the following questions:

i. What is the precedence graph for the schedule?

ii. Is the schedule conflict-serializable? If so, what are all the equivalent serial schedules?

3. **[15 points]** (**Exercise 18.2.5 DS:CB**) Say that a transaction T *precedes* a transaction U in a schedule S if every action of T precedes every action of U in S. Note that if T and U are the only transactions in S, then saying T precedes U is the same as saying that S is the serial schedule (T, U). However, if S involves transactions other than T and U, then S might not be serializable, and in fact, because of the effect of other transactions, S might not even be conflict-serializable. Give an example of a schedule S such that:

i. In S, T_1 precedes T_2, and

ii. S is conflict-serializable, but

iii. In every serial schedule conflict-equivalent to S, T_2 precedes T_1.

4. **[15 points]** (**Exercise 18.8.1 DS:CB**) Below are several sequences of events, including *start* events, where st_1; means that transaction T_1 starts. These sequences represent real time, and the timestamp scheduler will allocate timestamps to transactions in the order of their starts. Tell what happens as each executes.

a) st_1; st_2; $r_1(A)$; $r_2(B)$; $w_2(A)$; $w_1(B)$;

b) st_1; $r_1(A)$; st_2; $w_2(B)$; $r_2(A)$; $w_1(B)$;

c) st_1; st_2; st_3; $r_1(A)$; $r_2(B)$; $w_1(C)$; $r_3(B)$; $r_3(C)$; $w_2(B)$; $w_3(A)$;

d) st_1; st_3; st_2; $r_1(A)$; $r_2(B)$; $w_1(C)$; $r_3(B)$; $r_3(C)$; $w_2(B)$; $w_3(A)$;

5. **[15 points]** (**Exercise 18.9.1 DS:CB**) In the following sequence of events, we use $R_i(X)$; to mean "transaction T_i starts, and its read set is the list of database elements X." Also, V_i means "T_i attempts to validate," and $W_i(X)$; means that "T_i finishes, and its write set was X." Tell what happens when each sequence is processed by a validation-based schedular.

a) $R_1(A, B)$; $R_2(B, C)$; V_1; $R_3(C, D)$; V_3; $W_1(A)$; V_2; $W_2(A)$; $W_3(B)$;

b) $R_1(A, B)$; $R_2(B, C)$; V_1; $R_3(C, D)$; V_3; $W_1(A)$; V_2; $W_2(A)$; $W_3(D)$;

c) $R_1(A, B)$; $R_2(B, C)$; V_1; $R_3(C, D)$; V_3; $W_1(C)$; V_2; $W_2(A)$; $W_3(D)$;

6. **[15 points]** **R&G** (pp. 574) Exercise 17.2

7. **[15 points]** **R&G** (pp. 598)

(a) Exercise 18.4

(b) Exercise 18.5