CS-245 Database System Principles – Winter 2002
Assignment 6

Due at the beginning of class on Tuesday, March 5th

- State all assumptions and show all work.
- Subscribe to cs245@lists.stanford.edu to receive clarifications and changes.
- You can email questions to cs245-staff@cs.stanford.edu

Problem 1 (30 points)

The following is a sequence of undo-log records written by two transactions $T$ and $U$:

$$<\text{START } T>; <T, A, 10>; <\text{START } U>; <U, B, 20>; <T, C, 30>; <U, D, 40>; <\text{COMMIT } U>; <T, E, 50>; <\text{COMMIT } T>$$

Briefly describe the actions of the recovery manager in the order they were performed, including changes to both disk and the log, if there is a crash and the last log record that appears on disk is:

- a) $<\text{START } U>$
- b) $<T, E, 50>$
- c) $<\text{COMMIT } T>$

Problem 2 (15 points)

The following is a sequence of redo-log records written by three transactions $T$, $U$ and $V$:

$$<\text{START } T>; <T, A, 10>; <\text{START } U>; <U, B, 20>; <T, C, 30>; <\text{START CKPT } (T,U)>; <U, D, 40>; <\text{COMMIT } U>; <T, E, 50>; <\text{START } V>; <V, C, 45>; <\text{END CKPT}>; <\text{COMMIT } V>; <T, D, 45>$$

Briefly describe the actions of the recovery manager (changes to disk only), if the system boots after a crash and discovers this log. Write the actions in the order they are performed.

Problems 3-4.

For problems 3 and 4, consider the following transaction log from the start of the run of a database system that is capable of running undo/redo logging with checkpointing:

1. $<\text{START } T1>$
2. $<T1, A, 50, 10>$
3. $<\text{START } T2>$
4. $<T1, B, 130, 10>$
5. $<T1, A, 70, 50>$
6. $<T2, C, 20, 10>$
7. $<T2, D, 30, 10>$
8. $<\text{COMMIT } T1>$
9. $<\text{START } T3>$
10. $<T3, E, 60, 10>$
11. $<T2, D, 40, 30>$
12. <START CKPT (T2,T3)>
13. <T2, C, 70, 20>
14. <COMMIT T2>
15. <START T4>
16. <T4, F, 100, 10>
17. <T4, G, 110, 10>
18. <COMMIT T3>
19. <T4, F, 150, 100>
20. <START T5>
21. <T5, C, 200, 70>
22. <END CKPT>
23. <T4, F, 140, 150>
24. <COMMIT T4>

Assume the log entries are in the format <Tid, Variable, New value, Old value>.

Problem 3 (20 points)

What is the value of the data items A, B, C, D, E, F, and G on disk after recovery…

a) if the system crashes just before line 10 is written to disk?
b) if the system crashes just before line 15 is written to disk?
c) if the system crashes just before line 17 is written to disk?
d) if the system crashes just before line 19 is written to disk?
e) if the system crashes just before line 24 is written to disk?
f) if the system crashes just after line 24 is written to disk?

Format your answer as a table, with rows a-f and a column for each variable A-G. All the cells in the table must be full; i.e., all the information you need is in the problem.

Problem 4 (20 points)

Write down all the possible values each data item can have on disk after each of the crash points in Problem 3. For example, for a), part of your entry should look like:

<table>
<thead>
<tr>
<th>A</th>
<th>B</th>
<th>…</th>
</tr>
</thead>
<tbody>
<tr>
<td>10, 50, 70</td>
<td>10, 130</td>
<td></td>
</tr>
</tbody>
</table>

Problem 5. (15 points)

Prove that for two schedules S1 and S2, if P(S1)=P(S2) and P(S1) is acyclic, then S1 and S2 are conflict equivalent. The proof will be graded for correctness and clarity.