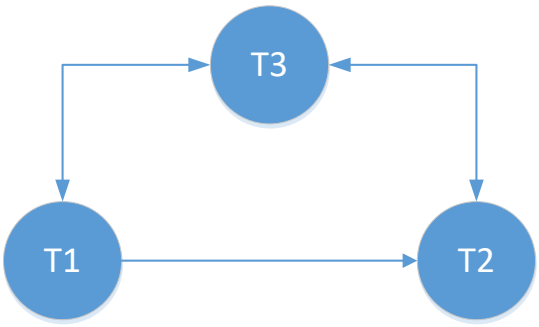
Assignment 2

Managing Big Data

Galdan MOULINNEUF – 2927686

Question 2

1) Here is a precedence graph for the schedule:

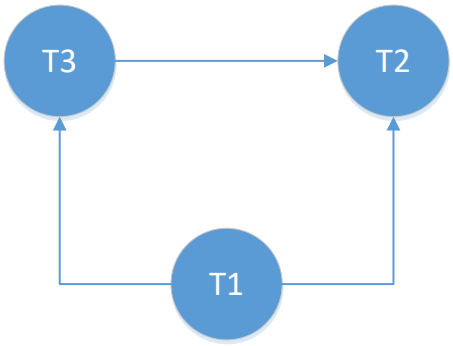


2) The schedule is not conflict serializable because the precedence graph isn't cyclic.

3) Here are the instances:

$$cycle\ X(T2 \rightarrow T3), Y(T3 \rightarrow T2)$$
$$cycle\ X(T2 \rightarrow T3), YZ(T3 \rightarrow T1), Y(T1 \rightarrow T2)$$

4) The new schedule gives the following precedence graph:



This schedule is conflict serializable because:

$$T1 \rightarrow T3 \rightarrow T2$$

Question 3

Since it's not stated in the subject, I will assume that tables evocated here (Dirty Page Table and the Transaction Table) were both empty before the log started. Also, TT records will be represented as (transID, lastLSN) and DPT records as (pageID, recLSN).

a) Analysis

First, the analysis phase tells us that the last begin checkpoint is LSN-00 and ends at LSN-10. Then, this phase will run until LSN 110 by doing the following:

LSN	ACTION
20	Adds (T1, 20) to TT; Adds (P5, 20) to DPT
30	Adds (T2, 30) to TT; Adds (P3, 30) to DPT
40	Change status of T2 to C from U
50	Deletes entry to T2 from TT
60	Adds (T3, 60) to TT; Does not change P3 entry in DPT
70	Changes (T1, 20) to (T1, 70)
80	Adds (T4, 80) to TT; Adds (P6, 80) to DPT
90	Change (T4, 80) to (T4, 90); Adds (P7, 90)
100	Change status of T4 to C from U
110	Deletes entry for T4 from TT

In the end, here are the status of the two tables:

TT has 2 entries: (T1, 70); (T3, 60)
DPT has 4 entries: (P5, 20); (P3,30); (P6, 80); (P7,90).

b) Redo

The redo phase starts at LSN 20 because this is the smallest recLSN in DPT.
The phase will also run until LSN 110 by doing the following:

LSN	ACTION
20	Change to P5 redone.
30	P3 is retrieved and its pageLSN is checked. If the page had been written to disk before the crash (i.e. if pageLSN >= 30), nothing is re-done otherwise the changes are re-done.
40	No action
50	No action
60	Changes to P3 are redone
70	No action
80	Change to P6 redone.
90	P7 is retrieved and its pageLSN is checked. If the page had been written to disk before the crash (i.e. if pageLSN >= 90), nothing is re-done otherwise the changes are re-done.
100	No action
110	No action

c) Undo

The undo phase starts at LSN 110 because this is the highest lastLSN in TT.

The Loser Set are LSN 70 and LSN 60.

LSN 70: Adds LSN 20 to the Loser Set.

Loser Set = (60, 20)

LSN 60: Undoes the change on P3 and adds a CLR indicating this Undo.

Loser Set = (20)

LSN 20: Undoes the change on P5 and adds a CLR indicating this Undo.