

CARNEGIE MELLON UNIVERSITY
COMPUTER SCIENCE DEPARTMENT
15-445/645 – DATABASE SYSTEMS (FALL 2020)
PROF. ANDY PAVLO

Homework #4 (by Gautam Jain)
Due: **Sunday Nov 8, 2020 @ 11:59pm**

IMPORTANT:

- **Upload this PDF** with your answers to **Gradescope by 11:59pm on Sunday Nov 8, 2020.**
- **Plagiarism:** Homework may be discussed with other students, but all homework is to be completed **individually.**
- **You have to use this PDF for all of your answers.**

For your information:

- Graded out of **100** points; **4** questions total
- Rough time estimate: \approx 1 - 2 hours (0.5 - 1 hours for each question)

Revision : 2020/10/29 23:24

Question	Points	Score
Serializability and 2PL	19	
Deadlock Detection and Prevention	32	
Hierarchical Locking	30	
Optimistic Concurrency Control	19	
Total:	100	

Question 1: Serializability and 2PL.....[19 points]

(a) Yes/No questions:

- i. [2 points] A conflict serializable schedule need not always be view serializable.
 Yes No
- ii. [2 points] There could be schedules under 2PL (not rigorous) that are not serializable.
 Yes No
- iii. [2 points] A view serializable schedule may contain a cycle in its precedence graph.
 Yes No
- iv. [2 points] It is not possible to have a deadlock in rigorous 2PL.
 Yes No
- v. [2 points] You will never have unrepeatable reads in rigorous 2PL.
 Yes No

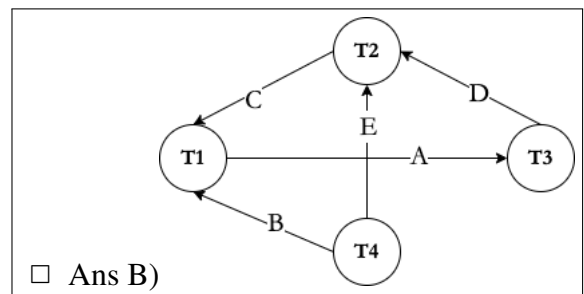
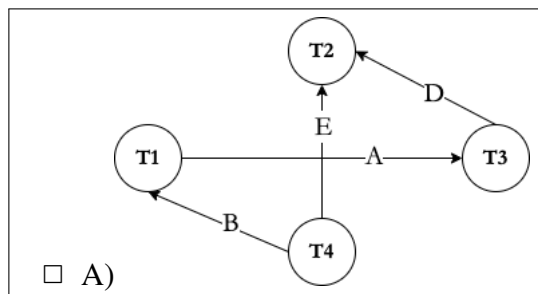
(b) Serializability:

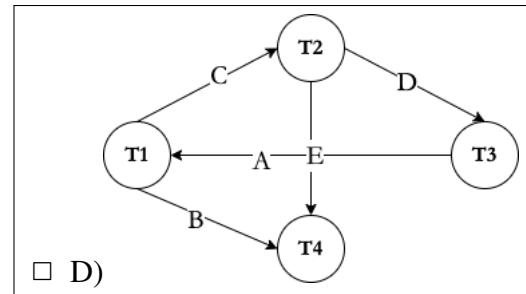
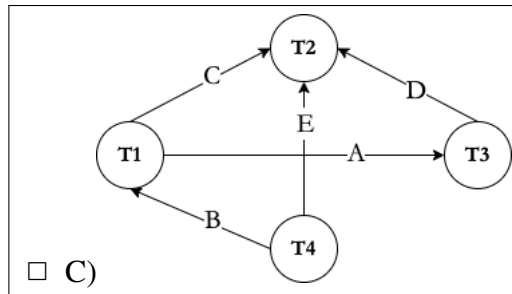
Consider the schedule given below in Table 1. R(·) and W(·) stand for ‘Read’ and ‘Write’, respectively.

time	t_1	t_2	t_3	t_4	t_5	t_6	t_7	t_8	t_9	t_{10}	t_{11}
T_1	R(A)			R(C)			R(B)		W(C)		
T_2						R(C)				W(D)	W(E)
T_3		R(E)			W(A)						
T_4			W(B)					R(D)			

Table 1: A schedule with 4 transactions

- i. [1 point] Is this schedule serial?
 Yes No
- ii. [2 points] Choose the correct dependency graph of the schedule given above. Each edge in the dependency graph looks like this: ‘ $T_x \rightarrow T_y$ with Z on the arrow indicating that there is a conflict on Z where T_x read/wrote on Z before T_y ’.





- iii. **[1 point]** Is this schedule conflict serializable?
 Yes No
- iv. **[3 points]** Mark all the transactions that can be removed from the schedule that can make it serializable.
 T1 T2 T3 T4 Original schedule is also serializable
- v. **[2 points]** Is this schedule possible under 2PL?
 Yes No

Question 2: Deadlock Detection and Prevention.....[32 points]

(a) Deadlock Detection:

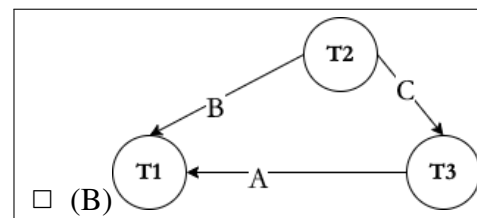
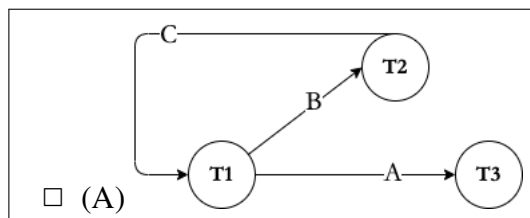
Consider the following lock requests in Table 2. And note that

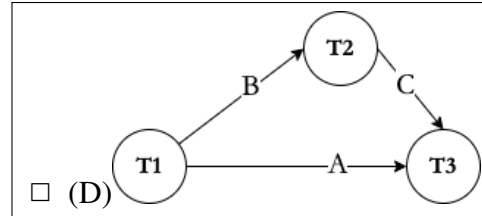
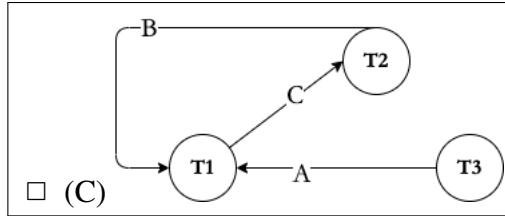
- S(·) and X(·) stand for ‘shared lock’ and ‘exclusive lock’, respectively.
- $T_1, T_2,$ and T_3 represent three transactions.
- LM stands for ‘lock manager’.
- Transactions will never release a granted lock.

time	t_1	t_2	t_3	t_4	t_5	t_6	t_7
T_1	S(A)				S(B)	S(C)	
T_2		S(B)		X(C)			X(B)
T_3			X(A)				
LM	g						

Table 2: Lock requests of three transactions

- i. For the lock requests in Table 2, determine which lock will be granted or blocked by the lock manager. Please write ‘g’ in the LM row to indicate the lock is granted and ‘b’ to indicate the lock is blocked or the transaction has already been blocked by a former lock request. For example, in the table, the first lock (S(A) at time t_1) is marked as granted.
- (a) [1 point] At t_2 : g b
 - (b) [1 point] At t_3 : g b
 - (c) [1 point] At t_4 : g b
 - (d) [1 point] At t_5 : g b
 - (e) [1 point] At t_6 : g b
 - (f) [1 point] At t_7 : g b
- ii. [2 points] Mark the correct wait-for graph for the lock requests in Table 2. Each edge in the wait-for graph looks like this: $T_x \rightarrow T_y$ because of Z . Z is denoted in the arrow in the figure. (i.e., T_x is waiting for T_y to release its lock on resource Z).





iii. **[2 points]** Determine whether there exists a deadlock in the lock requests in Table 2. Mark all that apply.

- There is no deadlock
- Cycle $(T_3 \rightarrow T_1 \rightarrow T_3)$ exists and schedule deadlocks
- There is no cycle
- Cycle $(T_1 \rightarrow T_2 \rightarrow T_1)$ exists and schedule deadlocks

(g) Deadlock Prevention:

Consider the following lock requests in Table 3.

Like before,

- $S(\cdot)$ and $X(\cdot)$ stand for ‘shared lock’ and ‘exclusive lock’, respectively.
- T_1, T_2, T_3, T_4 , and T_5 represent five transactions.
- LM represents a ‘lock manager’.
- Transactions will never release a granted lock.

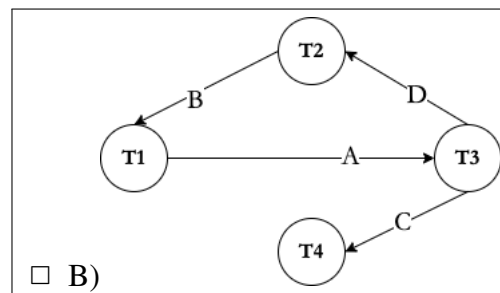
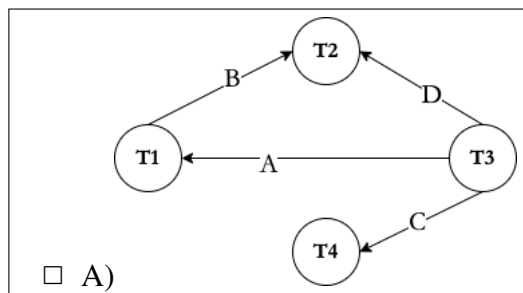
time	t_1	t_2	t_3	t_4	t_5	t_6	t_7	t_8
T_1	S(A)		S(B)					
T_2		X(B)						X(D)
T_3				S(C)	X(D)		X(A)	
T_4						X(C)		
LM	g	g						

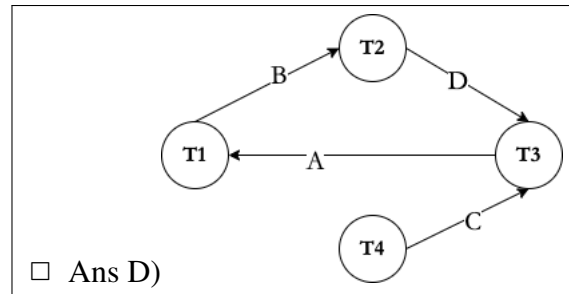
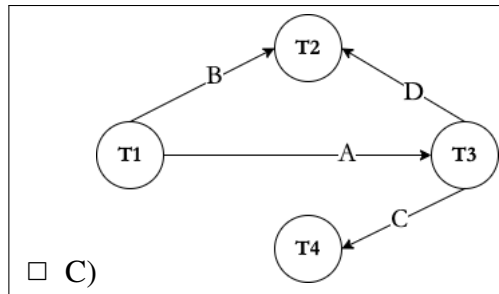
Table 3: Lock requests of four transactions

- i. For the lock requests in Table 3, determine which lock request will be granted, blocked or aborted by the lock manager (LM), if it has no deadlock prevention policy. Please mark ‘g’ for grant, ‘b’ for block (or the transaction is already blocked), ‘a’ for abort, and ‘-’ if the transaction has already died.

- (a) [1 point] At t_3 : g b a -
- (b) [1 point] At t_4 : g b a -
- (c) [1 point] At t_5 : g b a -
- (d) [1 point] At t_6 : g b a -
- (e) [1 point] At t_7 : g b a -
- (f) [1 point] At t_8 : g b a -

- ii. [2 points] Mark the correct wait-for graph for the lock requests in Table 3. Each edge in the wait-for graph looks like this: $T_x \rightarrow T_y$ because of Z . Z is denoted in the arrow in the figure. (i.e., T_x is waiting for T_y to release its lock on resource Z). If a lock request is not proposed because the transaction is blocked before making that lock request, you can ignore that lock request, as if it does not exist.





iii. [2 points] Determine whether there exists a deadlock in the lock requests in Table 3. Mark all that apply.

- There is no deadlock
 Cycle $(T_1 \rightarrow T_2 \rightarrow T_3 \rightarrow T_1)$ exists and schedule deadlocks
 There is no cycle
 Cycle $(T_3 \rightarrow T_2 \rightarrow T_1 \rightarrow T_3)$ exists and schedule deadlocks

iv. To prevent deadlock, we use the lock manager (LM) that adopts the Wait-Die policy. We assume that in terms of priority: $T_1 > T_2 > T_3 > T_4$. Here, $\overline{T_1} > \overline{T_2}$ because T_1 is older than T_2 (i.e., older transactions have higher priority). Determine whether the lock request is granted ('g'), blocked ('b'), aborted ('a'), or already dead ('-'). Follow the same format as the previous question.

- (a) [1 point] At t_3 : g b a -
 (b) [1 point] At t_4 : g b a -
 (c) [1 point] At t_5 : g b a -
 (d) [1 point] At t_6 : g b a -
 (e) [1 point] At t_7 : g b a -
 (f) [1 point] At t_8 : g b a -

v. Now we use the lock manager (LM) that adopts the Wound-Wait policy. We assume that in terms of priority: $T_1 > T_2 > T_3 > T_4$. Here, $\overline{T_1} > \overline{T_2}$ because T_1 is older than T_2 (i.e., older transactions have higher priority). Determine whether the lock request is granted ('g'), blocked ('b'), granted by aborting another transaction ('a'), or the requester is already dead ('-'). Follow the same format as the previous question.

- (a) [1 point] At t_3 : g b a -
 (b) [1 point] At t_4 : g b a -
 (c) [1 point] At t_5 : g b a -
 (d) [1 point] At t_6 : g b a -
 (e) [1 point] At t_7 : g b a -
 (f) [1 point] At t_8 : g b a -

Question 3: Hierarchical Locking [30 points]

Consider a database (D) consisting of two tables, Release (R) and Artists (A). Specifically,

- Release(rid, name, artist_credit, language, status, genre, year, number_sold), spans 1000 pages, namely R_1 to R_{1000}
- Artists(id, name, type, area, gender, begin_date_year), spans 50 pages, namely A_1 to A_{50}

Further, **each page contains 100 records**, and we use the notation $R_3 : 20$ to represent the 20th record on the third page of the Release table. Similarly, $A_5 : 10$ represents the 10th record on the fifth page of the Artists table.

We use Multiple-granularity locking, with **S**, **X**, **IS**, **IX** and **SIX** locks, and **four levels of granularity**: (1) *database-level* (D), (2) *table-level* (R, A), (3) *page-level* ($R_1 - R_{1000}$, $A_1 - A_{50}$), (4) *record-level* ($R_1 : 1 - R_{1000} : 100$, $A_1 : 1 - A_{50} : 100$).

For each of the following operations on the database, check all the sequence of lock requests based on intention locks that should be generated by a transaction that wants to efficiently carry out these operations by maximizing concurrency. Please take care of efficiency for eg. do not take an exclusive lock when a shared lock is sufficient.

Please follow the format of the examples listed below:

- mark “**IS(D)**” for a request of **database-level IS lock**
 - mark “**X($A_2 : 30$)**” for a request of **record-level X lock for the 30th record on the second page of the Artists table**
 - mark “**S($A_2 : 30 - A_3 : 100$)**” for a request of **record-level S lock from the 30th record on the second page of the Artists table to the 100th record on the third page of the Artists table.**
- (a) **[6 points]** Fetch the 55th record on page R_{343} .
- S($R_{343} : 55$)
 - IS(R_{343}), S($R_{343} : 55$)
 - IS(D), IS(R), IS(R_{343}), S($R_{343} : 55$)
 - SIX(D), SIX(R), SIX(R_{343}), X($R_{343} : 55$)
- (b) **[6 points]** Scan all the records on pages R_1 through R_{10} , and modify the record $R_2 : 33$.
- IS(D), SIX(R), IX(R_2), X($R_2 : 33$)
 - IX(D), SIX(R), IX(R_2), X($R_2 : 33$)
 - IX(D), SIX(R), IS(R_2), X($R_2 : 33$)
 - IX(D), IX(R), S(R_1), S($R_3 - R_{10}$), SIX(R_2), X($R_2 : 33$)

- (c) **[6 points]** Count the number of releases with 'year' > 2014.
- IS(D), S(R)
 - S(R)
 - X(R)
 - IX(D), X(R)
- (d) **[6 points]** Increase the number_sold of all release by 2701.
- IS(D), S(R)
 - S(R)
 - X(R)
 - IX(D), X(R)
- (e) **[6 points]** Increase the artist_credit in release and id in artist by 1 for all the tuples in the respective tables.
- X(D)
 - S(D)
 - IS(D), S(R), S(A)
 - IX(D), X(R), X(A)

Question 4: Optimistic Concurrency Control [19 points]

Consider the following set of transactions accessing a database with object A , B , C , D . The questions below assume that the transaction manager is using **optimistic concurrency control** (OCC). Assume that a transaction switches from the READ phase immediately into the VALIDATION phase after its last operation executes.

Note: VALIDATION may or may not succeed for each transaction. If validation fails, the transaction will get immediately aborted.

You can assume that the DBMS is using the serial validation protocol discussed in class where only one transaction can be in the validation phase at a time, and each transaction is doing forward validation (i.e. Each transaction, when validating, checks whether it intersects its read/write sets with any active transactions that have not yet committed.)

time	T_1	T_2	T_3
1	READ(A)		
2			READ(B)
3	WRITE(A)		
4	READ(C)		
5	WRITE(C)		
6	VALIDATE?		
7		READ(D)	
8	WRITE?		
9		WRITE(D)	
10		WRITE(B)	
11		VALIDATE?	
12			READ(A)
13		WRITE?	
14			WRITE(B)
15			VALIDATE?
16			WRITE?

Figure 1: An execution schedule

- (a) [6 points] Will T_1 abort?
 Yes
 No
- (b) [6 points] Will T_2 abort?
 Yes
 No
- (c) [6 points] Will T_3 abort?
 Yes
 No
- (d) [1 point] OCC is good to use when there are few conflicts.
 True
 False