Carnegie Mellon University
Computer Science Department
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: ≈ 1 - 2 hours (0.5 - 1 hours for each question)

<table>
<thead>
<tr>
<th>Question</th>
<th>Points</th>
<th>Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>Serializability and 2PL</td>
<td>19</td>
<td></td>
</tr>
<tr>
<td>Deadlock Detection and Prevention</td>
<td>32</td>
<td></td>
</tr>
<tr>
<td>Hierarchical Locking</td>
<td>30</td>
<td></td>
</tr>
<tr>
<td>Optimistic Concurrency Control</td>
<td>19</td>
<td></td>
</tr>
<tr>
<td>Total:</td>
<td>100</td>
<td></td>
</tr>
</tbody>
</table>

Revision: 2020/10/29 23:24
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 serializ-able.
   □ 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.

<table>
<thead>
<tr>
<th>time</th>
<th>t₁</th>
<th>t₂</th>
<th>t₃</th>
<th>t₄</th>
<th>t₅</th>
<th>t₆</th>
<th>t₇</th>
<th>t₈</th>
<th>t₉</th>
<th>t₁₀</th>
<th>t₁₁</th>
</tr>
</thead>
<tbody>
<tr>
<td>T₁</td>
<td>R(A)</td>
<td></td>
<td>R(C)</td>
<td></td>
<td></td>
<td>R(B)</td>
<td></td>
<td>W(C)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>T₂</td>
<td></td>
<td></td>
<td>R(C)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>W(D)</td>
<td>W(E)</td>
<td></td>
</tr>
<tr>
<td>T₃</td>
<td>R(E)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>W(A)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>T₄</td>
<td></td>
<td>W(B)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>R(D)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

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: ‘Tx → Ty with Z on the arrow indicating that there is a conflict on Z where Tx read/wrote on Z before Ty’.
   □ A)  □ Ans B)
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

(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.

<table>
<thead>
<tr>
<th>time</th>
<th>$t_1$</th>
<th>$t_2$</th>
<th>$t_3$</th>
<th>$t_4$</th>
<th>$t_5$</th>
<th>$t_6$</th>
<th>$t_7$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$T_1$</td>
<td>S(A)</td>
<td></td>
<td></td>
<td></td>
<td>S(B)</td>
<td>S(C)</td>
<td></td>
</tr>
<tr>
<td>$T_2$</td>
<td></td>
<td>S(B)</td>
<td>X(C)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$T_3$</td>
<td></td>
<td></td>
<td>X(A)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$LM$</td>
<td>g</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

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 $Z$).

![Wait-for graphs](image_url)
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.

<table>
<thead>
<tr>
<th>time</th>
<th>$t_1$</th>
<th>$t_2$</th>
<th>$t_3$</th>
<th>$t_4$</th>
<th>$t_5$</th>
<th>$t_6$</th>
<th>$t_7$</th>
<th>$t_8$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$T_1$</td>
<td>S(A)</td>
<td></td>
<td>S(B)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$T_2$</td>
<td></td>
<td>X(B)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$T_3$</td>
<td></td>
<td>S(C)</td>
<td>X(D)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$T_4$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>X(A)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$LM$</td>
<td>g</td>
<td>g</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

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.

□ A)

□ B)

Question 2 continues...
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.

\[
\begin{align*}
\text{(a) [1 point] At } t_3: & \quad \square \ g \quad \square \ b \quad \square \ a \quad \square \ - \\
\text{(b) [1 point] At } t_4: & \quad \square \ g \quad \square \ b \quad \square \ a \quad \square \ - \\
\text{(c) [1 point] At } t_5: & \quad \square \ g \quad \square \ b \quad \square \ a \quad \square \ - \\
\text{(d) [1 point] At } t_6: & \quad \square \ g \quad \square \ b \quad \square \ a \quad \square \ - \\
\text{(e) [1 point] At } t_7: & \quad \square \ g \quad \square \ b \quad \square \ a \quad \square \ - \\
\text{(f) [1 point] At } t_8: & \quad \square \ g \quad \square \ b \quad \square \ a \quad \square \ -
\end{align*}
\]

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.

\[
\begin{align*}
\text{(a) [1 point] At } t_3: & \quad \square \ g \quad \square \ b \quad \square \ a \quad \square \ - \\
\text{(b) [1 point] At } t_4: & \quad \square \ g \quad \square \ b \quad \square \ a \quad \square \ - \\
\text{(c) [1 point] At } t_5: & \quad \square \ g \quad \square \ b \quad \square \ a \quad \square \ - \\
\text{(d) [1 point] At } t_6: & \quad \square \ g \quad \square \ b \quad \square \ a \quad \square \ - \\
\text{(e) [1 point] At } t_7: & \quad \square \ g \quad \square \ b \quad \square \ a \quad \square \ - \\
\text{(f) [1 point] At } t_8: & \quad \square \ g \quad \square \ b \quad \square \ a \quad \square \ -
\end{align*}
\]
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 R₁ to R₁₀₀₀
• Artists(id, name, type, area, gender, begin_date_year), spans 50 pages,
  namely A₁ to A₅₀

Further, each page contains 100 records, and we use the notation R₃ : 20 to represent the
20th record on the third page of the Release table. Similarly, A₅ : 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 (R₁ − R₁₀₀₀, A₁ −
A₅₀), (4) record-level (R₁ : 1 − R₁₀₀₀ : 100, A₁ : 1 − A₅₀ : 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(A₂ : 30)” for a request of record-level X lock for the 30th record on the
  second page of the Artists table
• mark “S(A₂ : 30 − A₃ : 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 R₃₄₃.
   □ S(R₃₄₃ : 55)
   □ IS(R₃₄₃), S(R₃₄₃ : 55)
   □ IS(D), IS(R), IS(R₃₄₃), S(R₃₄₃ : 55)
   □ SIX(D), SIX(R), SIX(R₃₄₃), X(R₃₄₃ : 55)

(b) [6 points] Scan all the records on pages R₁ through R₁₀, and modify the record R₂ : 33.
   □ IS(D), SIX(R), IX(R₂), X(R₂ : 33)
   □ IX(D), SIX(R), IX(R₂), X(R₂ : 33)
   □ IX(D), SIX(R), IS(R₂), X(R₂ : 33)
   □ IX(D), IX(R), S(R₁), S(R₃ − R₁₀), SIX(R₂), X(R₂ : 33)

Question 3 continues...
(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.)

```
<table>
<thead>
<tr>
<th>time</th>
<th>T1</th>
<th></th>
<th>T2</th>
<th></th>
<th>T3</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>READ(A)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>WRITE(A)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>READ(C)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>WRITE(C)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>VALIDATE?</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>WRITE?</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>READ(D)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>WRITE?</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>WRITE(D)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>WRITE(B)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>11</td>
<td>VALIDATE?</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>12</td>
<td>WRITE?</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>13</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>14</td>
<td>WRITE(B)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>15</td>
<td>VALIDATE?</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>16</td>
<td>WRITE?</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
```

Figure 1: An execution schedule

(a) **[6 points]** Will T1 abort?
   - Yes
   - No

(b) **[6 points]** Will T2 abort?
   - Yes
   - No

(c) **[6 points]** Will T3 abort?
   - Yes
   - No

(d) **[1 point]** OCC is good to use when there are few conflicts.
   - True
   - False

End of Homework #4