Database Management System: Assignment 8

Total Marks : 100

August 28, 2023

Question 1

Assume an `immediate database modification scheme`. Consider the following log records for transactions T1, T2, T3 and T4: *Marks: 2* **MCQ**

steps	Details of log
1	⟨T1,start⟩
2	⟨T2,start⟩
3	⟨T1,A,500,600⟩
4	⟨T1,commit⟩
5	⟨T2,B,200,400⟩
6	⟨checkpoint{T2} ⟩
7	⟨T3,start⟩
8	⟨T2,commit⟩
9	⟨T3,C,200,500⟩
10	⟨T4,start⟩
11	⟨T4,D,600,1000⟩
12	⟨T4,commit⟩

If there is a crash just after step 12 and the recovery of the system is successfully completed, identify the **correct** action for the above scenario.

a) No Action: T1; Redo: T4; Undo: T2, T3

b) No Action: T2; Redo: T4; Undo: T1, T3

c) No Action: T1; Redo: T2, T3; Undo: T4

d) No Action: T1; Redo: T2, T4; Undo: T3

Answer: d)
Explanation: In the `immediate database modification scheme`, during recovery after a crash, a transaction needs to be redone if and only if both ⟨T_i, start⟩, ⟨T_i, commit⟩ are present in the log. Otherwise, undo is required.
Any transactions that are committed before the last checkpoint should be ignored(updates already output to disk due to the checkpoint).
Redo list contains transaction {T2, T4} and undo list contains transaction {T3} and for transaction {T1} no need any action because it is already committed before checkpoint.
Hence, option d) is the answer.

Question 2

Let us consider the following statistics for searching for a condition in a given relation.

Marks:2 **MCQ**

- Number of blocks containing record of the relation (b) = 70
- Time to transfer one block (t_b) = 0.2 milliseconds
- Time for one seek (t_s) = 10 milliseconds

Identify the `cost of selection query` on a `key attribute` using linear search file scan.

a) 10.2 milliseconds

b) 17 milliseconds

c) 24 milliseconds

d) 10 milliseconds

Answer: b)
Explanation: If selection is on a key attribute, cost = (b/2) block transfers + 1 seek
Cost=((70/2)*0.2) + 10 = 17 milliseconds.
For more details refer to Module 38, slide 15.

Question 3

Assume an immediate database modification scheme. Consider the following log records for transactions T1, T2, T3 and T4: *Marks: 2* **MSQ**

steps	Details of log
1	⟨T3,start⟩
2	⟨T3,A,200,400⟩
3	⟨T2,start⟩
4	⟨T2,B,300,600⟩
5	⟨T2,commit⟩
6	⟨checkpoint{T3} ⟩
7	⟨T1,start⟩
8	⟨T1,C,400,800⟩
9	⟨T4,start⟩
10	⟨T3,commit⟩
11	⟨T4,D,300,700⟩

If there is a crash just after step 11 and the recovery of the system is successfully completed, identify the **correct** action for the above scenario.

a) After recovery completion, value of A will be 200.

b) After recovery completion, value of A will be 400.

c) After recovery completion, value of C will be 400.

d) After recovery completion, value of C will be 800.

Answer: b), c)
Explanation: In the immediate database modification scheme, during recovery after a crash, a transaction needs to be redone if and only if both ⟨T_i, start⟩, ⟨T_i, commit⟩ are present in the log. Otherwise, undo is required.
Any transactions that are committed before the last checkpoint should be ignored(updates already output to disk due to the checkpoint).
Redo list contains transaction {T3} and undo list contains transactions {T1, T4} and for transaction {T2} no need any action because it is already committed before checkpoint.
As per the process of transaction recovery, options (b) and (c) are correct.

Question 4

Let us consider the following statistics for two relations `Mountain_Climber` and `Trekk_Schedule`:

Marks: 2 **MCQ**

- Number of records of `Mountain_Climber`: $n_{\texttt{Mountain\_Climber}} = 4000$.
- Number of blocks of `Mountain_Climber`: $b_{\texttt{Mountain\_Climber}} = 400$.
- Number of records of `Trekk_Schedule`: $n_{\texttt{Trekk\_Schedule}} = 2000$.
- Number of blocks of `Trekk_Schedule`: $b_{\texttt{Trekk\_Schedule}} = 200$.

Identify the required number of `block transfers` and `seeks` if the smaller relation fits entirely in memory, use that as the inner relation using `Nested-loop join`.

a) 800400 block transfers and 2 seeks

b) 800200 block transfer and 2 seeks

c) 600 block transfers and 2 seeks

d) 400 block transfers and 2 seeks

Answer: c)
Explanation: If the smaller relations (`Trekk_Schedule`) fit entirely into the memory, it is a must to use that relation as the inner relation in `Nested-loop join`.
As a result, each block will only be read once.
Hence, the total number of block transfers will be: $b_{\texttt{Mountain\_Climber}} + b_{\texttt{Trekk\_Schedule}}$
$=400 + 200$
$= 600$.
Total number of seeks required: 2.
For more details refer to 38.31 of lecture material.

Question 5

Consider the log record of Transaction T1 with one operation instance O1 used in a recovery system with early lock release, B+ tree based concurrency control. *Marks: 2* **MCQ**

Step	Operation
1	⟨T1, start⟩
2	⟨T1, A, 700, 200⟩
3	⟨T1, O1, operation–begin⟩
4	⟨T1, B, 300, 600⟩
5	⟨T1, C, 200, 400⟩
6	⟨T1, O1, operation–end,(B, -300),(C, -200)⟩
7	crash or abort here

Choose the correct set of log entries for the recovery of transactions.

a)

⟨T1,C,200⟩
⟨T1,B,300⟩
⟨T1,O1,operation–abort⟩
⟨T1,A,700⟩
⟨T1,abort⟩

b)

⟨T1,C,400,200⟩
⟨T1,B,600,300⟩
⟨T1,O1,operation–abort⟩
⟨T1,A,700,200⟩
⟨T1,abort⟩

c)

⟨T1,C,200,400⟩
⟨T1,B,300,600⟩
⟨T1,O1,operation–abort⟩
⟨T1,A,700⟩
⟨T1,abort⟩

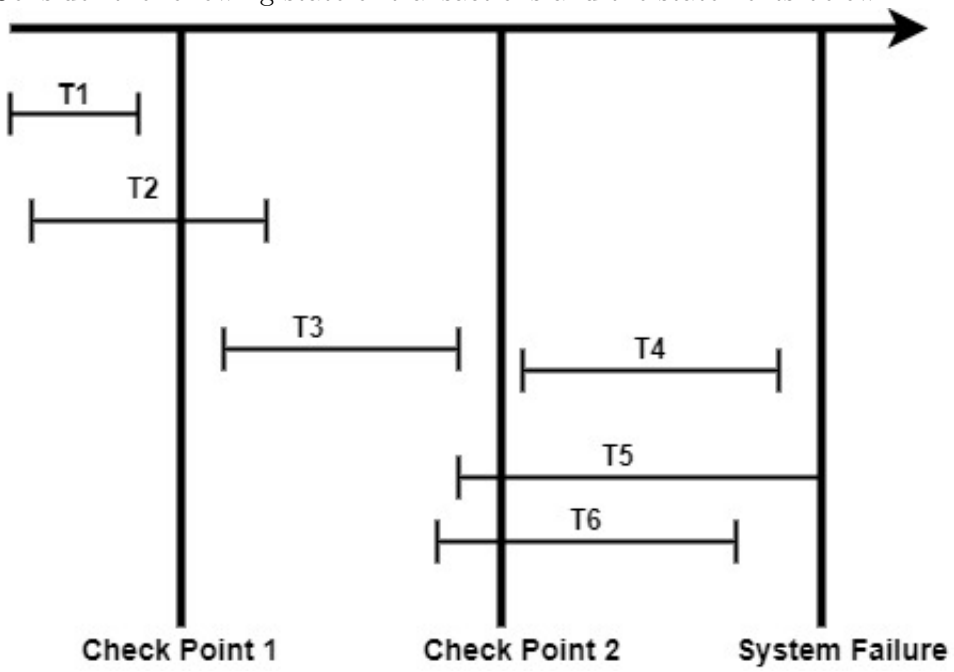
d)

⟨T1,C,400,200⟩
⟨T1,B,600,300⟩
⟨T1,O1,operation–abort⟩
⟨T1,A,700⟩
⟨T1,abort⟩

Answer: d)
Explanation: Step i: Scan the log records backward.
Step ii: For step 6, ⟨T1, O1, operation-end, (B, -300),(C, -200)⟩ log is found; so, logical undo is required for operation O1 on the variables C and B using the information (subtraction of 200 for C variable) and (subtraction of 300 for B variable). That means we have to delete the previous modifications on B and C. So, add the following logs for steps 5, 4 and 3 respectively: ⟨T1,C,400,200⟩, ⟨T1,B,600,300⟩ and ⟨T1,O1,operation–abort⟩.
Step iv: For step 2, add the log ⟨T1, A, 700⟩.
Step v: For step 1, add the log ⟨T1, abort ⟩.
Hence, option (d) is correct.

Question 6

Consider the following state of transactions and the statements below.



1. Only T_1 can be ignored.
2. T_1, T_2 and T_3 can be ignored.
3. T_2, T_3, T_4, and T_6 need to be redone.
4. T_4 and T_6 need to be redone.
5. Only T_5 needs to be undone.

Identify the correct group of statements from the options below. *Marks:2* **MCQ**

a) 1), 2), 3), 5)

b) 1), 3), 4), 5)

c) 1), 3), 5)

d) 2), 4), 5)

Answer: d)
Explanation: Any transaction that is committed before the last checkpoint should be ignored. Therefore, T_1, T_2, and T_3 can be ignored (updates already output to disk due to the last checkpoint).
Any transaction that is committed since the last checkpoint, needs to be redone. Hence, T_4 and T_6 are to be redone.
Any transaction that was running at the time of failure, needs to be undone and restarted. Hence, only T_5 is to be undone.
Hence, option (d) is correct.

Question 7

Consider the following relational schema: *Marks: 2* **MCQ**

```
Mountain_Climber(MCid, MC_name, Address, EmailID)
Trekk_Details(Tname, Altitude, Diff_Level)
Trekk_Schedule(MCid, Tname, Start_date, End_date)
```

Two query trees are given below.

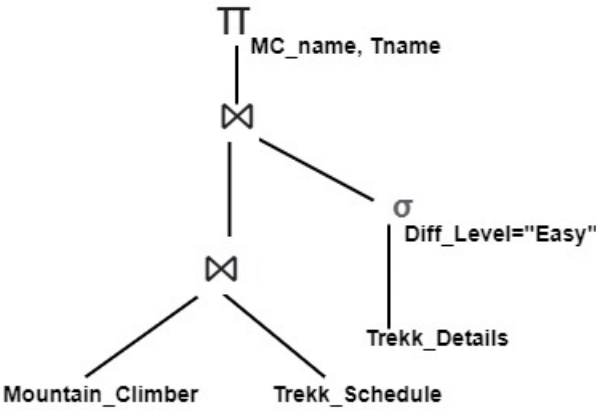


Figure 1:

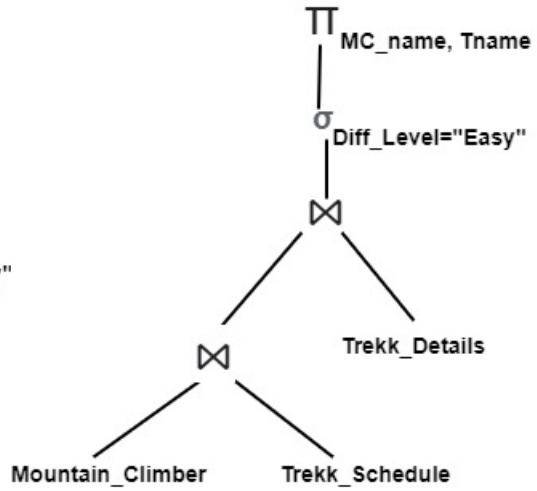


Figure 2:

Identify the correct statement for the above two query trees.

a) Two query trees are equivalent as identical operations (irrespective of their positions) are used in both trees.

b) Two query trees are not equivalent as selection or projection operation cannot be carried out before or after the natural join operation.

c) Two query trees are equivalent and the query tree of Figure 1 will lead to more efficient query processing.

d) Two query trees are equivalent and the query tree of Figure 2 will lead to more efficient query processing.

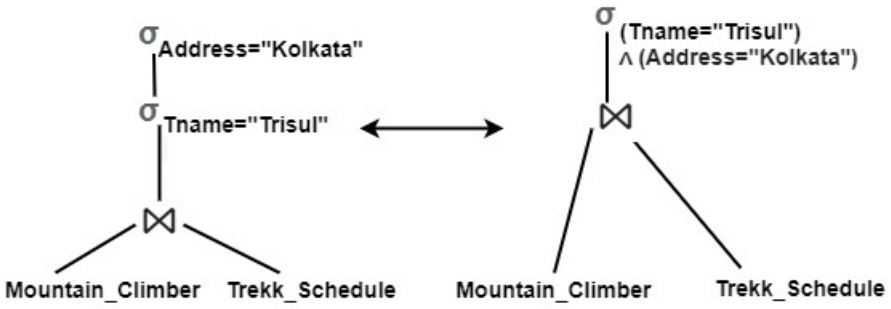
Answer: c)
Explanation: Two query trees are equivalent, and Figure 1 will lead to more efficient query processing because performing the selection operation as early as possible reduces the size of the relation to be joined.
Hence, option (c) is correct.

Question 8

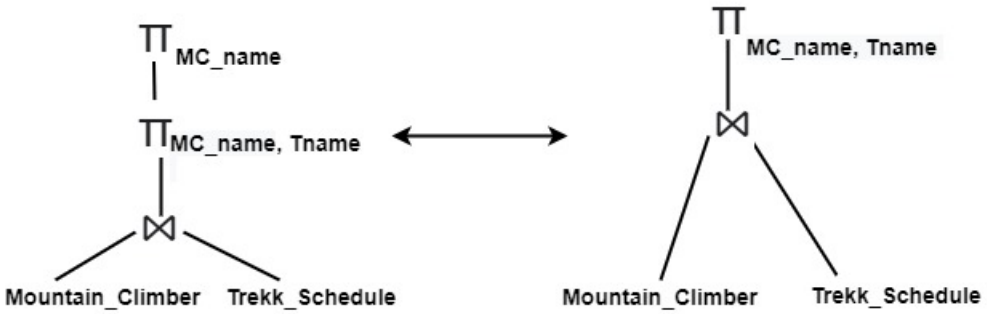
Consider the following relational schema: *Marks: 2* **MSQ**

Mountain_Climber(MCid, MC_name, Address, EmailID)
Trekk_Details(Tname, Altitude, Diff_Level)
Trekk_Schedule(MCid, Tname, Start_date, end_date)
Identify the option(s) that represent equivalent Query trees.

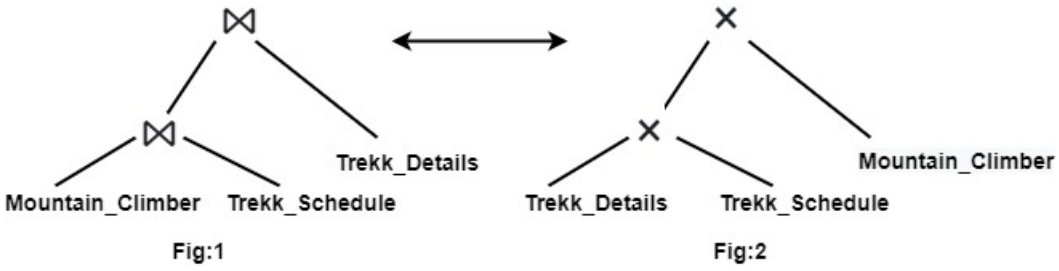
a)



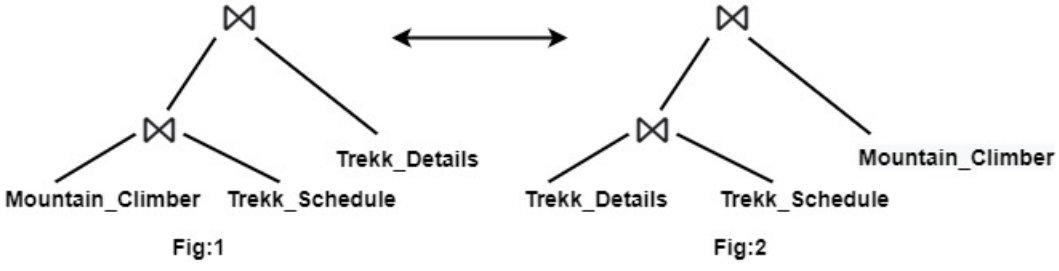
b)



c)



d)



Answer: a), d)
Explanation: As per the equivalence rules of relational algebra expressions, options a) and d) are correct. For more details refer to Module 39, slides 11,12.

Question 9

Identify the `cost estimation` of a `query evaluation plan`, if 500 blocks are required to be transferred from the disk and the required number of `disk seeks` is 20.

- Time to transfer one block: $t_T = 2$ milliseconds.
- Time for one seek: $t_S = 0.2$ seconds.

Marks:2 **MCQ**

a) 1.2 Seconds

b) 2.2 Seconds

c) 5 Seconds

d) 5.2 Seconds

Answer: c)
Explanation: Cost for b block transfers plus S seeks will be $(b * t_T + S * t_S)$ seconds
$=(500 * 2 * 10^{-3}) + (20 * 0.2)$ seconds
$= (1 + 4)$ Seconds
$= 5$ Seconds
For more details refer to 38.12 of lecture material.
Hence, option c) is the answer.

Question 10

Which of the following statements is (are) **true**? *Marks: 2* **MSQ**

a) `Physical blocks` are those blocks residing on the disk.

b) `System buffer blocks` are those blocks which reside temporarily in the main memory.

c) `System buffer blocks` are those blocks residing on the disk.

d) The `log` is a sequence of `schedules`, which maintains information about topological orderings of each schedule.

Answer: a), b)
Explanation: The `log` is a sequence of log records, which maintains information about `update activities` on the database, kept on any stable storage.
Physical blocks are those blocks residing on the disk.
System buffer blocks are the blocks residing temporarily in main memory.
Hence, options (a) and (b) are the correct answers.
For more details refer to Module 36.