Department of Computer Science and Engineering
Indian Institute of Technology Patna

Course Name: Database
Course Code: CS354
EndSem Paper
Autumn 2020
Date: 23rd Nov
Full Marks: 50+20=70

Question numbers 1 to 10 are of multiple choice type and carrying 5 marks each. For these questions, select only one option from the given options. Full marks will be awarded for finding the correct option and with valid justification only. Only 60% marks will be awarded if there is no justification or justification is invalid. There will be a negative marking (40%) for each wrong selection. There is no option for question number 11 and you need to provide an elaborate answer for this question. Get the necessary instructions for submissions from the instructor. PLEASE DO NOT SHARE YOUR ANSWERS WITH YOUR BATCH MATES

1. A disk spins at 10000 rpm, has 50000 tracks per surface, an average of 120 sectors per track, and four platters; data is stored on only one side of the platters. The average seek time is 5 ms. The sector size is 512 bytes and the block size is 2 sectors. Also assume that the block transfer time depends on the block size, track size, and rotational speed. What is the average time (in ms or miliseconds) needed to find and transfer a block, given its block address?

   (A) (≥ 6 and < 8) ms
   (B) (≥ 8 and < 10) ms
   (C) (≥ 10 and < 12) ms
   (D) (≥ 12 and < 14) ms
   (E) none of these

   `Give justification`

   5 Marks

2. Consider the data and parity block arrangement as shown in Table 1 and Table 2. D_i represents the ith data block; P_j represents the jth parity block. In each Table, parity block is defined over the data blocks of the corresponding row.

Disk1	Disk2	Disk3	Disk4
P1	D2	D3	D4
D5	P6	D7	D8
D9	D10	P11	D12
⋮	⋮	⋮	⋮

Table 1: Arrangement 1

Disk1	Disk2	Disk3	Disk4
P1	D2	D3	D4
P5	D6	D7	D8
P9	D10	D11	D12
⋮	⋮	⋮	⋮

Table 2: Arrangement 2

   Which of the following is TRUE with respect to these arrangements.

   (A) Arrangement 1 provides same benefit as arrangement 2
   (B) Arrangement 1 provides more benefit than arrangement 2
   (C) Arrangement 1 provides less benefit than arrangement 2

(D) nothing can be concluded

`Give Justification`

5 Marks

3. A B+ tree index is to be built on the name attribute of relation Student. Assume that all student names are of length 8 bytes, disk blocks are of 512 bytes, and index pointer are of size 4 bytes. Given this scenario, what would be the best choice of the degree of the B+ tree.

(A) 16

(B) 32

(C) 42

(D) 43

(E) none of these

`Give justification`

5 Marks

4. Consider a relational table r with sufficient number of records, having attributes A_1, A_2, $\cdots$, A_n and let $1 \leq p \leq n$. The two queries Q_1 and Q_2 are given below

$Q_1 : \pi_{A_1,\cdots A_p}(\sigma_{A_p=c}(r))$ where c is a constant
$Q_2 : \pi_{A_1,\cdots,A_p}(\sigma_{c_1 \leq A_p \leq c_2}(r))$ where c_1 and c_2 are constants

The database can be configured to do ordered indexing on A_p or hashing on A_p. Which of the following statements is TRUE?

(A) Ordered indexing will always outperform hashing for both queries

(B) Hashing will always outperform ordered indexing for both queries

(C) Hashing will outperform ordered indexing on Q_1, but not on Q_2

(D) Hashing will outperform ordered indexing on Q_2, but not on Q_1

`Give justification`

5 Marks

5. Suppose the hash function is $h(x) = x\%8$ and each bucket can hold atmost two records. If extendable has structure is used and the following elements are inserted in the order specified
$1, 4, 5, 7, 8, 2, 20$
After inserting these elements, based on the aforementioned hash structure, how many buckets will be there?

(A) 4

(B) 6

(C) 7

(D) 8

(E) none of these

`Give justification`

5 Marks

6. Let r and s be two relations with the following schema

   $r < P, Q, A1, A2, A3 > s < P, Q, B1, B2 >$
   where $< P, Q >$ is the candidate key for both the schemas. Which of the following queries are equivalent?

   (I) $\prod_P(r \bowtie s)$
   (II) $\prod_P(r) \bowtie \prod_P(s)$
   (III) $\prod_P(\prod_{P,Q}(r) \cap \prod_{P,Q}(s))$
   (IV) $\prod_P(\prod_{P,Q}(r) - (\prod_{P,Q}(r) - \prod_{P,Q}(s)))$

   (A) Only I and II
   (B) Only I and III
   (C) Only I, II and III
   (D) Only I, III and IV
   (E) none of these

   `Give justification`

   5 Marks

7. Consider the following schedule S

$T1$	$T2$	$T3$
R(A)		
	R(B)	
		R(B)
	W(B)	
W(A)		

   Schedule S is conflict equivalent to

   (A) $T2 \rightarrow T1 \rightarrow T3$
   (B) $T1 \rightarrow T2 \rightarrow T3$
   (C) $T3 \rightarrow T1 \rightarrow T2$
   (D) none of these

   `Give justification`

   5 Marks

8. Consider the following two schedules - $S1$ and $S2$

   (A) S1 is recoverable but S2 is not
   (B) S2 is recoverable but S1 is not
   (C) Both S1 and S2 are recoverable

$T1$	$T2$
R(x)	
	R(x)
W(x)	
R(y)	
	W(x)
	Commit
W(y)	
Commit	

Table 3: Schedule S1

$T1$	$T2$
R(x)	
W(x)	
	R(x)
R(y)	
	W(x)
	Commit
W(y)	
Commit	

Table 4: Schedule S2

(D) Neither S1 nor S2 are recoverable

`Give justification`

5 Marks

9. Consider the tree protocol in which data items A, B and C are ordered as shown in Figure 1.

   Now consider a schedule S as defined below-

   - $T1$: this transaction needs to lock A and C data items in exclusive mode. $T1$ needs to use each of the data item for 10 time units
   - $T2$: this transaction needs to lock B data item in exclusive mode. $T2$ needs to use B data item for 150 time units.
   - To improve concurrency $T2$ issues exclusive mode lock request on B data item immediately after $T1$ issues exclusive mode lock request on A data item.

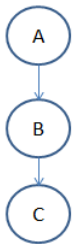

Figure 1: Partial order relationships of data items

Ignore any context switching time and locking/unlocking times, how much time will be required by transaction $T1$ and $T2$ to complete execution using schedule S

(A) T1:20 and T2:150

(B) T1:20 and T2:170

(C) T1:170 and T2:170

(D) T1:170 and T2:150

(E) None of these

`Give justification`

5 Marks

10. In an immediate database modification scheme, consider the following log sequences of two transactions.

```
1. <T1, start>
2. <T1, B, 12000, 10000>
3. <T1, M, 0, 2000>
4. <T1, commit>
5. <T2, start>
6. <T2, B, 10000, 10500>
7. <T2, commit>
```

Suppose the database system crashed just before the log record 7 is written. When the system is restarted, which one statement is true of the recovery procedure?

(A) redo of log record 6 to set B to 10500

(B) undo log record 6 to set B to 10000 and then redo log records 2 and 3

(C) need not redo log records 2 and 3 because transaction T1 has committed

(D) apply redo and undo operations in arbitrary order because they are idempotent

`Give justification`

5 Marks

11. Propose a strategy that can be used to improve the concurrency degree of a tree-based protocol. Create your own example of a schedule to illustrate the proposed strategy.

`The proposed strategy should be illustrated with a suitable example`

20 Marks