

University of Ruhuna - Faculty of Technology
Bachelor of Information & Communication Technology Degree
Level 3 (Semester 2) Examination
August 2020

Course Unit: ICT3273, Advanced Database Management Systems

Time Allowed: 02 hours

This question paper contains **06 pages** including this instruction page.

IMPORTANT INSTRUCTIONS:

1. The medium of this examination is English.
2. This is a Closed Book examination.
3. This Examination consists of four (04) questions that are given equal marks.
4. You must answer all four (04) questions in this examination.

1. a. Different ways are used to store data records in Database Management Systems.
- Write down one (01) advantage and one (01) disadvantage of the heap file organization. [10 marks]
 - Differentiate *Pile File Method* and *Sorted File Method* when inserting a record to a sequential file organization. [10 marks]
- b. Indexing in databases are defined based on its indexing attributes.
- By using an example briefly describe the clustered index. [10 marks]
 - Differentiate *Dense Primary Index* and *Sparse Primary Index* using suitable diagrams. [20 marks]
- c. Linear hashing is a dynamic hashing technique which offers a lot of flexibility with respect to the timing of bucket splits.
- Use the given Linear Hashing Index snapshot in Figure 01 to answer the questions given below. Assume that a bucket split occurs whenever an overflow page is created

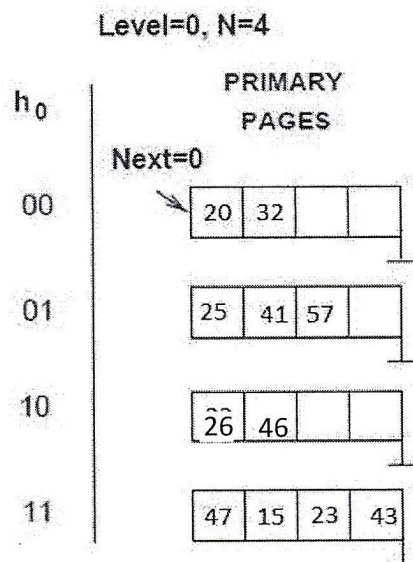


Figure 01

- What is the maximum number of entries that you can enter before you are going to have a split bucket? Give your reasons. [10 marks]
- Show the file after inserting 46 and 59. Clearly show the primary pages, if any overflow pages, changes in Level, N, Next, h_0 etc. [30 marks]
- Write down the value of Level and Next if all four (04) buckets have been split. [10 marks]

2. a. Magnetic disks are used as secondary storage devices to store data. Assume that there is a disk with a sector size of 512 bytes, 1000 tracks per surface, 100 sectors per track, five (05) double sided platters and average seek time of 10 milliseconds.

- i. What is the capacity of the disk? [04 marks]
- ii. No of cylinders the disk has. [04 marks]
- iii. Two (02) valid block sizes and two (02) invalid block sizes. [04 marks]
- iv. If the disk platters rotate at 7200 rpm (revolutions per minute), what is the maximum rotational delay? [04 marks]
- v. what is the transfer rate of the disk if two (02) track of data can be transferred per revolution? [04 marks]

b. B+ tree is a dynamic structure where the height of the tree grows and contracts as records are added and deleted.

- i. Compare a leaf node and a non-leaf node with m entries in a B+ tree by using a suitable diagram. [10 marks]
- ii. Insert 95 into below given B+ tree in Figure 02 and redraw the form of B+ tree after the operation.

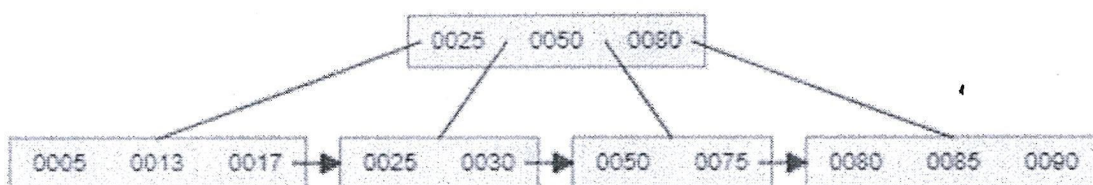


Figure 02

[20 marks]

iii. Delete 50 from above part (b.) (ii.) obtained form and redraw the form of B+ tree after the operation.

[20 marks]

c. Query optimization is the process of selecting the most efficient query evaluation plan.

- i. Use below given relations and SQL query to answer the question given below.
 - lecturer(lec_id, name, designation)
 - subject(sub_id, title, credit)
 - teach(lec_id, sub_id, duration)

```

SELECT S.title, L.name
FROM lecturer L, subject S, teach T
WHERE L.lec_id = T.lec_id AND S.sub_id= T.sub_id
      AND L.designation = "professor" AND T.duration < 30;

```

Construct an initial query tree to represent above SQL query with cartesian product.

[10 marks]

ii. Produce an optimized query tree by applying heuristic transformations to the initial query tree obtained part (b.) (ii.) State the heuristics you have used.

[20 marks]

3. a. Several problems can be occurred when database transactions are executed concurrently.

i. Briefly describe ACID properties of a transaction.

[10 marks]

ii. Use the given below schedule and variable values to answer the questions given.

sum =0, avg =0, x=100, y=200, z=300

T ₁	T ₂
read(x); x := x - 50; write(x);	read(z); sum := sum + z;
read(y); y := y + 50; write(y); commit;	read(x); sum := sum + x; read(y); sum := sum + y; avg := sum/3; commit;

1. Write down the value of average after execution of the given schedule.

[05 marks]

2. Write down the value of average if T₁ executes before T₂.

[05 marks]

3. Write down the value of average if T₁ executes after T₂.

[05 marks]

4. Recognize and briefly describe the problem and the reason that you have

identified for the above part (a.) (ii) schedule in the context of concurrent execution of transactions.

[10 marks]

b. Database is restored to the most recent consistent state just before the time of failure.

i. List down two (02) reasons for transaction failures.

[05 marks]

ii. Briefly describe Deferred Update technique.

[10 marks]

iii. Differentiate current directory and shadow directory in shadow paging.

[10 marks]

c. Consider the four (04) transactions T1, T2, T3 and T4 and the schedule S given below.

T1 : r1(x); w1(x); r1(y); w1(y)

T2 : r2(z); r2(y); w2(x)

T3 : r3(y); r3(z); w3(y); w3(z)

T4 : r4(z); w4(z); r4(x)

S : r3(y); r3(z); r1(x); w1(x); w3(y); w3(z); r2(z); r4(z); w4(z); r1(y); w1(y); r2(y); w2(x); r4(x)

i. Identify and write down two (02) conflict operations and two (02) non conflict operations in the given schedule.

[10 marks]

ii. Construct the precedence directed graph for schedule S and by giving reasons state whether the schedule is serializable or not. If the given schedule is serializable write down the equivalent serial schedule.

[30 marks]

4. a. A distributed database (DDB) is an integrated collection of databases that is physically distributed across sites in a computer network.

i. Write down three (03) advantages of DDB.

[10 marks]

ii. Briefly describe the primary site technique in distributed concurrency control.

[10 marks]

b. Construct a wait for graph for the given schedule with four (04) transactions and identify whether the given transactions are in a deadlock situation.

[20 marks]

T1	T2	T3	T4
l1(A); r1(A)			
		l3(B); r3(B)	
	l2(C); r2(C)		

			l4(D); r1(D)
			l4(A);
		l3(C);	
l1(B);			
	l2(D);		

c. Two phase locking protocols(2PL) are requires both locks and unlocks being done in two phases.

i. Identify and briefly describe the two (02) phases of 2PL.

[10 marks]

ii. Consider the given below schedule with two transactions T1 and T2. Propose a plan using 2PL to ensure a conflict serializable schedule for the given transactions.

T1	T2
r1(X);	
X = X + 10;	
	r2(X);
	X = X + 10;
w1(X);	
	w2(X)
	COMMIT;
r1(Y);	
Y = Y + 10;	
w1(Y);	
COMMIT	

[30 marks]

d. Consider the given below schedule with two transactions T1, T2, T3 and T4. The time stamp values are as follows.

TS(T1) = 20, TS(T2) = 30, TS(T3) = 40, TS(T4) = 45

Read time stamp of X: R_TS(X) = 5

Write time stamp of X: W_TS(X) = 10;

T1	T2	T3	T4
R1(X)			
		R3(X)	
		W3(X)	
	R2(X)		
W1(X)			
			W4(X)

Calculate R_TS(X) and W_TS(X) at the end of this schedule by applying basic time stamp ordering algorithm.

[20 marks]